



L'INFORMAZIONE STATISTICA MICROTERRITORIALE

L'ESPERIENZA DEL REGISTRO
STATISTICO DI BASE DEI LUOGHI



L'INFORMAZIONE STATISTICA MICROTERRITORIALE: L'ESPERIENZA DEL REGISTRO STATISTICO DI BASE DEI LUOGHI

A cura di: Marco Di Zio, Angela Ferruzza.

Attività editoriali: Nadia Mignolli (coordinamento), Claudio Bava, Alfredina Della Branca, Marco Farinacci, Alessandro Franzò e Manuela Marrone.

Responsabile per la grafica: Sofia Barletta.

ISBN 978-88-458-2152-3

© 2024

Istituto nazionale di statistica
Via Cesare Balbo, 16 - Roma



Salvo diversa indicazione, tutti i contenuti pubblicati sono soggetti alla licenza Creative Commons - Attribuzione - versione 4.0.
<https://creativecommons.org/licenses/by/4.0/deed.it>

È dunque possibile riprodurre, distribuire, trasmettere e adattare liberamente dati e analisi dell'Istituto nazionale di statistica, anche a scopi commerciali, a condizione che venga citata la fonte.

Immagini, loghi (compreso il logo dell'Istat), marchi registrati e altri contenuti di proprietà di terzi appartengono ai rispettivi proprietari e non possono essere riprodotti senza il loro consenso.



INDICE

	Pag.
Prefazione del Presidente dell'Istat	7
1. Il Sistema integrato dei Registri	9
1.1 Introduzione	9
1.2 I registri del SIR	9
2. Il Registro statistico di base dei luoghi - RSBL	11
2.1 RSBL, le sue componenti e le geografie: dal dato amministrativo al dato statistico	11
2.2 RSBL: uno strumento per un linguaggio statistico comune e integrato	13
2.3 RSBL sviluppi futuri e potenzialità: dal globale al locale e dal locale al globale	14
3. L'architettura e il processo	19
3.1 Le funzioni di RSBL nel contesto del SIR	19
3.2 Il processo di aggiornamento del registro	20
3.3 La struttura semantica delle informazioni modellate dall'ontologia	21
3.4 L' <i>Ontology-based Data Management</i> applicata a RSBL	24
4. SITUAS: il Sistema informativo territoriale delle unità amministrative e statistiche per l'analisi della struttura e la dinamica del territorio	27
4.1 SITUAS in RSBL e il portale per l'utenza	27
4.2 Unità territoriali amministrative: comuni, province/Città metropolitane e regioni	28
4.3 Unità territoriali statistiche	30
4.4 Unità territoriali estere	31
4.5 Caratteristiche e classificazioni delle unità territoriali	31
4.6 Architettura del Sistema	32
4.6.1 <i>I moduli SITUAS per la diffusione</i>	32
4.6.2 <i>Il modulo SITUAS per gestione delle unità territoriali e le classificazioni</i>	33
5. Basi territoriali e microzonizzazione	35
5.1 Caratteristiche principali del Progetto Sezioni e Microzone 2021 e finalità	35
5.2 Regole generali	36
5.3 Flusso di lavoro per l'aggiornamento delle Basi territoriali 2021	37
5.3.1 <i>Definizione della proposta dello strato geografico delle sezioni 2021</i>	37

	Pag.
5.3.2 <i>Indagine censuaria per la validazione delle BT2021</i>	39
5.3.3 <i>Risultati dell'indagine e validazione della proposta Istat delle BT 2021</i>	40
5.3.4 <i>Strumenti GIS per l'aggiornamento delle BT2021</i>	42
5.4 Principali risultati	43
6. Integrazione dati geografici e analisi spaziale per la realizzazione delle nuove basi territoriali	47
6.1 Uno strumento di analisi: le sezioni di Censimento 2021	47
6.2 Integrazione SdC 2021 e altri GeoDB	47
6.2.1 <i>Integrazione con dati di copertura/uso del suolo</i>	47
6.2.2 <i>Analisi spaziali e indici radiometrici di vegetazione</i>	49
6.3 Versatilità della cartografia dell'Istat	50
7. La componente indirizzi in RSBL	51
7.1 Struttura, composizione e aggiornamento del Registro	51
7.2 L'indirizzo come elemento di geocodifica dell'unità statistica nel SIR	58
8. Registro edifici abitazioni e altre unità immobiliari: il processo di integrazione e localizzazione dei dati amministrativi	61
8.1 Il nuovo Registro edifici e alloggi: discontinuità con il passato e nuove possibilità	61
8.2 La cartografia catastale e l'integrazione delle altre fonti cartografiche	62
8.2.1 <i>Il processo di acquisizione dei dati geografici</i>	62
8.2.2 <i>Il processo di georiferimento automatico: ASIR</i>	62
8.2.3 <i>L'integrazione con altre fonti per la geocodifica alla sezione di Censimento e l'arricchimento della base informativa</i>	63
8.3 Abitazioni e altre unità immobiliari: elementi di integrazione con il Registro di popolazione e il Registro delle unità economiche	64
9. RSBL nel Sistema integrato dei registri	69
9.1 Metodologia di integrazione tra i registri RSBL e RBI	69
9.2 Integrazione tra i registri RSBL e RBI	70
9.3 Definizione dei livelli di link e dei gruppi di famiglie residenti	71
9.4 Prima metodologia deterministica	72
9.5 Calcolo della qualità delle associazioni famiglia-abitazione	73
9.6 Seconda metodologia deterministica	75
10. Rendere interoperabili i dati geospaziali di RSBL	77
10.1 Il WebGIS e i GeoPortali	77
10.2 La piattaforma WebGIS dell'Istat	78
11. Il processo tecnologico	81
11.1 L'industrializzazione	81
11.2 Il controllo di processo per la componente indirizzi	82
11.3 La reportistica	85

	Pag.
12. La griglia di popolazione	89
12.1 Definizione e uso della griglia di popolazione	89
12.2 La griglia di popolazione attraverso il metodo <i>top-down</i>	90
12.3 Iter realizzativi per la produzione della griglia di popolazione finale: il metodo <i>bottom-up</i>	91
13. Il Censimento della Popolazione e delle abitazioni e il Registro RSBL	95
13.1 Il Censimento permanente della Popolazione e delle abitazioni	95
13.1.1 Cambiamenti metodologici rispetto ai Censimenti tradizionali	95
13.1.2 Il processo di produzione dei dati sulle abitazioni	96
13.1.3 Allocazione delle unità statistiche a livello subcomunale	96
13.2 Distribuzione della popolazione e delle abitazioni occupate all'interno del Comune di Roma	97
14. Studio sulle periferie urbane dei comuni capoluogo delle Città metropolitane: primi risultati	101
14.1 Finalità e obiettivi	101
14.2 Periferie urbane di alcuni comuni capoluogo	102
14.2.1 Indice di vecchiaia e quota di stranieri residenti	103
14.2.2 Bassa istruzione e quota di giovani che non studiano e non lavorano	104
14.2.3 Famiglie con potenziale disagio economico e famiglie che non vivono in casa di proprietà	106
14.3 Sviluppi futuri	108
Glossario	111
Riferimenti bibliografici	115

PREFAZIONE

In un'epoca nella quale i dati guidano le decisioni a ogni livello della società, dalla pianificazione urbana all'azione per il clima, dallo sviluppo economico alla salute pubblica, la dimensione geografica delle informazioni risulta sempre più strategica e vitale. L'integrazione della localizzazione spaziale e delle informazioni territoriali nelle analisi statistiche correnti garantisce una lettura dei dati essenziale per analizzare, interpretare e affrontare opportunamente le sfide sociali, economiche, ambientali a livello locale e globale.

La collocazione territoriale delle informazioni non si limita soltanto a dire dove si trovano le unità statistiche ma risulta anche un'informazione indispensabile per comprendere come le relazioni spaziali e i contesti geografici influenzino la vita delle persone, l'efficienza delle imprese, le dinamiche delle economie e lo stato di salute del nostro ambiente. Localizzando il più precisamente possibile le unità statistiche, siano esse individui, famiglie, imprese o risorse naturali, si possono evidenziare comportamenti specifici, seguire i cambiamenti nel tempo e identificare le correlazioni che spesso sono nascoste all'interno dei dati non georeferenziati.

Negli ultimi anni, l'Istituto nazionale di statistica ha sviluppato un sistema integrato di informazioni con un elevato livello di dettaglio, il Sistema integrato dei registri, che rappresenta una importante infrastruttura di base per i processi di produzione statistica. Tra i vari registri, è stato sviluppato il Registro statistico base dei luoghi (RSBL), il quale fornisce strumenti e informazioni necessarie alla rappresentazione spaziale dei luoghi e, tramite processi di integrazione, permette di aggiungere questa componente alle unità statistiche presenti negli altri registri del sistema integrato.

RSBL è molto complesso e rappresenta, anche nel panorama internazionale, un'esperienza significativa. È stato, infatti, costruito sulla base di molteplici componenti e numerosi archivi amministrativi che sono stati utilizzati per produrre un registro coerente e consistente, nel pieno rispetto delle normative vigenti per la tutela della riservatezza. La sua costruzione costituisce una notevole sfida concettuale, metodologica, di processo e ha comportato un grosso impegno di risorse, coinvolgendo e valorizzando le numerose competenze tecniche presenti in Istituto.

Questo ebook rappresenta il prodotto finale ottenuto da un lavoro congiunto e descrive il processo di costruzione di RSBL e delle sue componenti, l'architettura che lo governa, e alcune applicazioni importanti come l'integrazione con altri registri. Nelle pagine che seguono si darà conto, con estremo dettaglio, anche delle potenzialità di questo registro nel rispondere alla domanda di analisi socio-economiche con una dimensione microterritoriale, che negli ultimi anni è divenuta sempre più pressante e articolata. A tale proposito voglio ricordare, ma solo a titolo di esempio, tutto il materiale che l'Istat ha prodotto per rispondere alle esigenze della Commissione parlamentare di inchiesta sulle condizioni di sicurezza e sullo stato di degrado delle città e delle loro periferie, lavoro che si è avvalso del potenziale informativo del Registro statistico di base dei luoghi e che ha consentito di descrivere in profondità le realtà dei 14 comuni capoluogo delle Città metropolitane.

Oltre a una approfondita descrizione delle componenti, delle metodologie di integrazione, degli strumenti informatici e di analisi geografica nell'ebook c'è molto di più e, in particolare, troviamo la dimostrazione di come il territorio e la geografia possono rappresentare criteri e strumenti di integrazione delle informazioni, consentendo al risultato finale di risultare maggiore della somma delle parti.

Questi sono elementi essenziali per comprenderne l'usabilità e rappresentare le potenzialità presenti in questa nuova infrastruttura informativa. RSBL rappresenta, per l'Istat e per tutti gli utenti finali, una miniera che ancora nasconde tesori da scoprire e ritengo, quindi, indispensabile proseguire con convinzione su questo sentiero di produzione, di studio e di analisi intrapreso.

Leggendo questo ebook invito tutti a riflettere sull'importanza dello spazio e del luogo per una nuova stagione nella produzione di statistica ufficiale e che rappresentano elementi fondamentali per dare forma alla nostra comprensione della realtà e del territorio che ci circonda.

Infine, a tutti gli attori che hanno contribuito con impegno a questo importante risultato raggiunto dal nostro Istituto va il mio più sincero ringraziamento.

Il Presidente dell'Istituto Nazionale di Statistica - Istat

Francesco Maria Chelli

1. IL SISTEMA INTEGRATO DEI REGISTRI¹

1.1 Introduzione

Il Sistema integrato dei registri (SIR) è una delle principali infrastrutture statistiche introdotte dal programma di modernizzazione dell'Istituto avviato nel 2016 (Istat 2016a). Nel processo di produzione statistico in questo contesto, i registri formano la base strutturale sulla quale si innestano indagini campionarie e informazioni da nuove fonti di dati, al fine di produrre le statistiche di interesse secondo i principi di qualità della statistica ufficiale. L'obiettivo del SIR è quello di definire un ambiente informativo unico, strutturato e coerente che sia vincolo e supporto ai vari processi produttivi statistici.

La rilevanza di questo sistema e il suo ruolo centrale derivano dal fatto che le informazioni presenti nei registri costituiscono la base di riferimento per tutti i processi statistici dell'Istat. Nei registri vengono raccolte le differenti popolazioni statistiche di interesse e un insieme di informazioni principali (per esempio, data di nascita e genere nel Registro degli individui) che diventano il riferimento per tutti i processi che di esse si occupano. L'integrazione dei registri, ovvero il SIR, permette la produzione di quadri informativi più ampi, capaci di rispondere alla domanda con un approccio multifunzionale che garantisce una gestione unitaria delle diverse tematiche (sociali, ambientali, economiche, eccetera) e una integrazione concettuale e statistica, oltre che fisica, tra unità statistiche di diverso tipo.

Dal punto di vista architettonico, il SIR è costituito da registri alimentati da fonti amministrative e da indagini. I registri e le indagini lavorano in maniera sinergica nella produzione di quadri informativi, per i quali si controlla la coerenza temporale e tematica. Il processo che porta all'integrazione informativa nel SIR si basa sulla modellazione di ontologie di dominio per l'integrazione concettuale e fisica, e su uno specifico quadro metodologico in cui ricadono le operazioni di trattamento dei dati e di valutazione della qualità.

1.2 I registri del SIR

Il Sistema integrato dei registri si compone di registri che fanno riferimento a tre contesti tematici principali: popolazione; unità produttive e istituzioni; territorio. Le relazioni tra tali tematismi permettono di connotarlo come sistema.

Nel disegno generale il SIR è composto da tre tipologie di registro: i Registri statistici di base (RSB), i Registri statistici tematici (RST) e i Registri statistici estesi (RSE).

I Registri statistici di base identificano le tipologie di unità statistiche e popolazioni di base di fondamentale importanza per la statistica ufficiale. Come esempi possiamo citare alcuni casi importanti (alcuni già sviluppati, altri in via di costruzione):

¹ Il Capitolo è stato redatto da Marco Di Zio.



- il Registro di base degli individui, delle famiglie e delle convivenze (RBI);
- il Registro di base delle unità economiche (RSUE);
- il Registro di base dei luoghi (RSBL);
- il Registro di base delle attività, un registro logico, che mette in relazione le unità statistiche dei diversi registri base in funzione di una attività che caratterizza la loro relazione.

I Registri statistici di base sono identificati dall'insieme delle unità statistiche che afferiscono a uno stesso dominio e possono contenere più tipologie di unità statistiche (come indirizzi, unità amministrative, edifici, eccetera), connesse tra di loro. Ad esempio, nel caso di RBI possono essere definite più popolazioni di riferimento, come residenti, abitualmente dimoranti, popolazione insistente sul territorio, eccetera. Ciascuna unità del registro può fare riferimento a più popolazioni statistiche di interesse.

I Registri statistici di base sono costituiti da unità statistiche che hanno le seguenti caratteristiche:

- codice unico statistico, stabile nel tempo, che identifica la specifica unità statistica;
- tutte le proprietà utili a identificare più popolazioni statistiche a supporto delle esigenze della produzione statistica ufficiale;
- tutte le proprietà utili per supportare la produzione di liste.

I Registri statistici estesi (RSE) ampliano l'informazione disponibile per una specifica popolazione di un RSB, con l'integrazione statistica di altre variabili, anche desunte da indagini statistiche, con il fine di supportare uno o più processi di produzione. Un esempio è il Registro esteso pubblica amministrazione (REPA).

I Registri statistici tematici identificano sistemi di informazioni a supporto di più processi statistici, attraverso il trattamento omogeneo e condiviso di alcune variabili definibili come *core*. Gli RST non sono, a differenza degli RSE, vincolati a specifiche popolazioni, ma hanno l'obiettivo di supportare analisi che fanno riferimento a specifiche tematiche, indipendentemente dell'appartenenza delle unità statistiche a specifiche popolazioni. Un esempio è il Registro tematico del Lavoro (RTL).

In questo contesto il ruolo del Registro statistico base dei luoghi (RSBL), in qualità di registro base, contiene tutte le unità statistiche che descrivono il territorio, sia dal punto di vista amministrativo sia dal punto di vista geografico, e consente quindi di integrare le altre unità statistiche al territorio attraverso l'architettura del SIR.

2. IL REGISTRO STATISTICO DI BASE DEI LUOGHI - RSBL¹

2.1 RSBL, le sue componenti e le geografie: dal dato amministrativo al dato statistico

Il Registro statistico di base dei luoghi (RSBL) può essere considerato il fulcro del Sistema integrato dei registri (SIR) e l'Istat ha modificato nel tempo i suoi processi, finalizzandoli anche alla costruzione e all'utilizzo di questo sia in produzione, sia in diffusione.

RSBL è costituito da diverse componenti, delle quali si tratterà estesamente nei capitoli dedicati:

- SITUAS², il registro relativo al Sistema di unità territoriali amministrative (7.904 comuni, province/città metropolitane e regioni), di Unità territoriali statistiche (Sistemi locali del lavoro, FUA) e di Unità territoriali estere. Tali unità sono indispensabili per l'analisi della struttura e la dinamica del territorio (cfr. Capitolo 4);
- le Basi territoriali, per le quali l'utilizzo di numerosi archivi di dati geografici ha dato luogo a 756 mila sezioni georeferenziate (erano 403 mila nel 2011), fondate sulla identificazione di 1,1 milioni di microzone che consentono l'individuazione anche di aree verdi e infrastrutture (cfr. Capitoli 5 e 6);
- il registro relativo agli indirizzi e alle coordinate geografiche, basato su numerosi archivi amministrativi, che hanno dato luogo, attualmente, a 30,2 milioni di Codice univoco di indirizzo (CUI) (cfr. Capitolo 7);
- il registro relativo agli edifici (a uso residenziale e non) e alle abitazioni, i cui dati sono derivati dall'uso congiunto del Catasto e di archivi desumibili da agenzie geografiche e dati geografici *open source*, che hanno identificato attualmente 14,3 milioni di edifici residenziali (anno 2021) e oltre 35 milioni di abitazioni (anno 2021) (cfr. Capitolo 8).

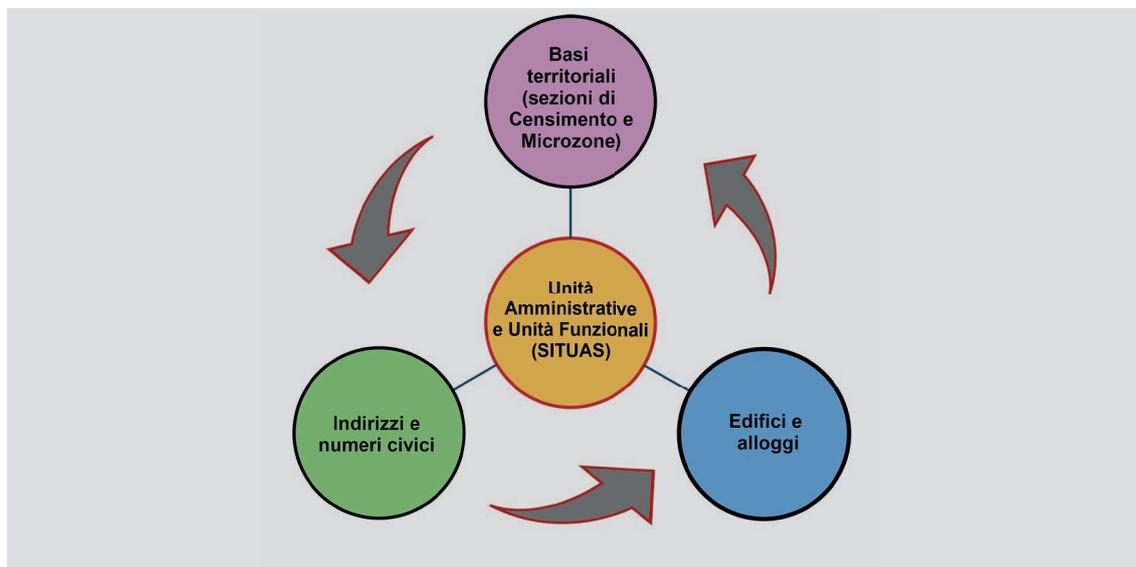
RSBL è, quindi, un sistema complesso costituito da diverse componenti (Figura 2.1), per ognuna delle quali le variabili sono state costruite dettagliando diverse caratteristiche delle entità sotto analisi e considerandone la qualità, basandosi sulla trasformazione del dato amministrativo in informazione statistica. La sfida è costruire informazione spaziale capace di rispondere alla domanda informativa pressante di statistiche territorialmente dettagliate, integrando le diverse componenti. Al fine di migliorare la produzione dei dati e le connesse analisi statistiche è, quindi, essenziale il raggiungimento di obiettivi che necessitano l'attenzione alle diverse geografie, la necessità di averne una per le unità statistiche di tutti gli altri registri sociali ed economici, per i dati ambientali, nonché per i dati relative ad alcune indagini.

¹ Il Capitolo è stato redatto da Angela Ferruzza.

² Il Sistema SITUAS è consultabile all'indirizzo: <https://situas.istat.it>.



Figura 2.1 - Le componenti di RSBL



Fonte: Elaborazioni su dati Istat

Il processo di costruzione e integrazione ha visto differenti metodi applicati a diverse entità per ottenere la più elevata qualità possibile nei risultati (cfr. Capitolo 3, Capitolo 9, Capitolo 10, Capitolo 11). Tale processo ha dovuto affrontare diverse questioni, alcune delle quali elencate a titolo esemplificativo ed evidenziate nei capitoli seguenti:

- un numero molto elevato di oggetti interessati;
- processi di aggiornamento continuo di numerosi archivi amministrativi dei quali è necessario considerare e controllare la qualità;
- architetture innovative basate su integrazione concettuale, fisica e statistica e sulle ontologie;
- processi tecnologici rilevanti;
- integrazione di componenti basate su fonti differenti indipendenti le une dalle altre e costruite con processi indipendenti tra loro;
- approcci metodologici che considerano *interlinkage* e strategie raffinate di integrazione;
- attenzione alle questioni connesse alla riservatezza dei dati nel rispetto delle normative vigenti;
- l'integrazione necessaria del Registro statistico di base degli individui (RSBI) e dei dati censuari con RSBL;
- l'integrazione avviata attualmente in via sperimentale dei registri economici con RSBL.

L'uso di dati amministrativi e di registri statistici è, senz'altro, un elemento essenziale della produzione statistica, ma è una grande sfida per ragioni istituzionali, metodologiche, procedurali, connesse anche alle tematiche relative alla riservatezza. RSBL, integrandosi con gli altri registri, assume un ruolo centrale nell'ambito del Sistemi integrato dei registri. RSBL, infatti, colloca sul territorio tutte le unità statistiche presenti nel SIR, consentendo il completamento geospaziale con diversi livelli di approssimazione: ad esempio, una famiglia risiede in una abitazione e una unità locale commerciale è presente in un negozio, questi sono collocabili in una sezione censimento di un determinato comune. RSBL consente, anche, di considerare la dinamica nel tempo delle unità territoriali, le interrelazioni tra queste e, conseguentemente, anche le diverse geografie nel tempo.

2.2 RSBL: uno strumento per un linguaggio statistico comune e integrato

RSBL, in quanto elemento di integrazione per i diversi registri, è la manifestazione concreta di come il territorio e le diverse geografie siano fattore di integrazione tra le statistiche sociali, ambientali ed economiche, consentendo di rispondere in maniera incrementale alla domanda, non sempre sufficientemente corrisposta, di informazione statistica di dettaglio territoriale elevato. Essendosi modificate le modalità di esecuzione dei censimenti, RSBL risulta essere indispensabile per la produzione dei dati censuari subcomunali e per sezione di censimento.

RSBL rispetta, infatti, il primo principio del GSGF, *Global Statistical Geospatial Framework* (UNSD 2019) (cfr. Capitolo 10), che sottolinea la necessità dell'adozione di un approccio comune e coerente per collocare ogni unità statistica di un *dataset* nel tempo e nello spazio, utilizzando infrastrutture geospaziali di riferimento.

In RSBL sono considerati luoghi diversi e geografie diverse. Potenzialmente, ogni tipo di informazione con coordinate geografiche può essere integrata: RSBL può consentire di georeferenziare informazioni statistiche o di individuare le aree interessate, garantendo l'implementazione di output flessibili e, quindi, di considerare Unità areali modificabili utilissime, ad esempio, in contesti di definizioni di strategie e azioni.

La qualità dell'informazione statistica e la potenza informativa degli indicatori e delle misure statistiche dipendono, in prima istanza, non solo dal modello statistico dei dati, ma anche da una adeguata disaggregazione geografica che può evidenziare alcune variazioni dei fenomeni.

In seconda istanza, le infrastrutture di dati spaziali offrono la possibilità di avere riferimenti geografici standardizzati per localizzare le unità statistiche. L'assegnazione di un identificatore univoco per ciascun luogo, o area, rende possibile i *linkage* sia di dati statistici e geografici, sia di dati statistici tra loro e l'associazione allo stesso spazio geografico.

La disponibilità di queste geografie comuni permette, quindi, la generazione di dati geostatistici, tramite griglie cartografiche o unità statistiche connesse a confini amministrativi o a delimitazioni statistiche (Figura 2.2). I dati statistici possono essere aggregati o disaggregati a differenti livelli ai fini della loro integrazione. È possibile avere geografie comuni tra diversi elementi statistici e geografie variabili: le geografie comuni sono essenziali per la costruzione di un linguaggio unitario e lo sono anche per l'integrazione tra i diversi domini statistici o fenomenologici. La geocodifica di dati statistici espande, quindi, considerevolmente le possibilità analitiche, consentendo le integrazioni di indicatori e di dati, nonché l'analisi degli stessi da un punto di vista geografico. RSBL, è quindi, tramite le geografie comuni, uno strumento potente per un linguaggio statistico comune e integrato.

Figura 2.2 - I diversi livelli informativi di RSBL nell'ambito di un comune. Sezioni 2021, edifici, indirizzi



Fonte: Elaborazioni su dati Istat

2.3 RSBL sviluppi futuri e potenzialità: dal globale al locale e dal locale al globale

Come si vedrà nel prosieguo del volume (cfr. Capitoli 12, 13 e 14) le prime analisi e applicazioni sono già disponibili: si può, infatti, fare riferimento alla popolazione per griglia di un chilometro quadrato, all'utilizzo di RSBL nell'ambito del processo di produzione, diffusione e analisi dell'informazione statistica del Censimento della Popolazione e delle abitazioni, anche fino al livello subcomunale, allo studio delle periferie urbane (Commissione parlamentare 2024). Si tratta di prime analisi che certamente avranno ulteriori sviluppi e approfondimenti. Sono in corso anche le attività relative all'integrazione di dati economici utilizzando RSBL, con riferimento attualmente ai dati del Registro ASIA unità locali. Queste, senz'altro, forniranno ricche opportunità di produzione di informazione statistica di qualità.

Le informazioni geografiche di RSBL hanno, inoltre, un crescente potenziale se si considerano le misure statistiche e gli indicatori relativi ai cambiamenti climatici³, agli eventi estremi⁴ e alla sostenibilità⁵. Se si parla di sostenibilità, facendo riferimento al *framework* di indicatori definito dall'Agenda 2030 dall'*Inter-agency and Expert Group on SDG Indicators* (IAEG-SDGs), uno degli elementi innovativi di maggior rilievo (Istat 2024c) è quello di

3 L'Accordo di Parigi adottato alla 21ª Conferenza delle parti della convenzione (COP21), richiede ai Paesi di accelerare e intensificare gli sforzi per combattere i cambiamenti climatici attraverso contributi determinati a livello nazionale (NDC), inclusi piani di adattamento nazionali (PAN).

4 Il Quadro di Sendai sancisce il passaggio dalla gestione delle catastrofi alla gestione del rischio di catastrofi, attraverso il ruolo di primo piano riconosciuto alle attività di prevenzione, con 7 obiettivi globali da raggiungere entro il 2030:

1. riduzione del numero di vittime causate da disastri;
2. riduzione del numero di persone colpite da disastri;
3. riduzione della perdita economica diretta;
4. riduzione del danno prodotto dalle catastrofi sulle infrastrutture critiche e sui servizi di base;
5. aumento del numero di Paesi con strategie di riduzione del rischio di disastri;
6. potenziamento della cooperazione internazionale rivolta ai Paesi in via di sviluppo;
7. aumento della disponibilità e dell'accesso ai sistemi di allerta rapida multirischio.

L'informazione statistica assume un ruolo essenziale in questi contesti.

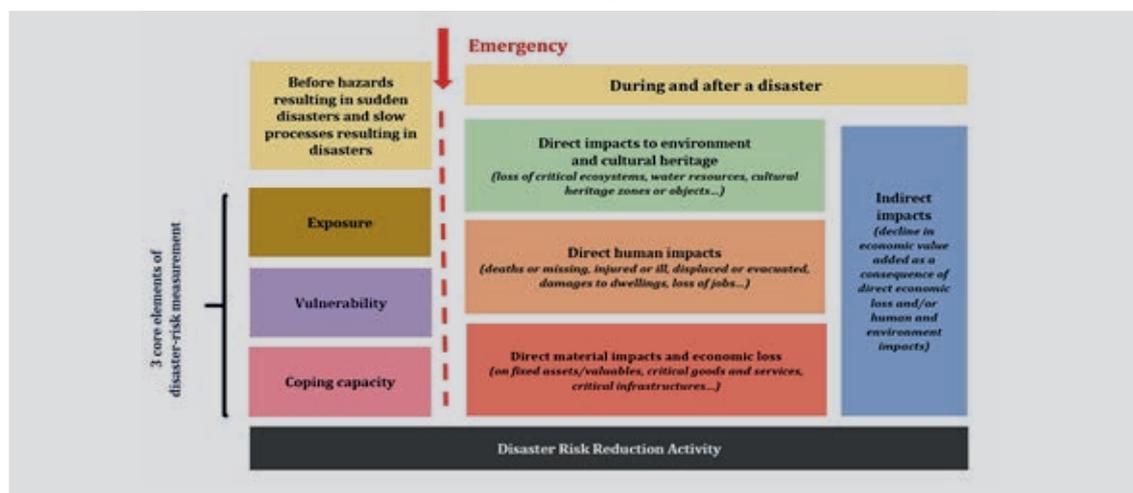
5 L'Agenda 2030 per lo Sviluppo sostenibile raccoglie in una visione olistica tutti i fenomeni economici, sociali, ambientali e istituzionali attraverso 17 obiettivi integrati e indivisibili.

2. Il Registro statistico di base dei luoghi - RSBL

considerare la necessità dell'integrazione dell'informazione statistica economica, sociale e ambientale con la necessità di avere le misure statistiche più disaggregate possibili anche a livello territoriale, per garantire di non lasciare nessuno indietro. Per tale ragione nell'ambito delle attività UN, è stata evidenziata (Istat 2023), l'importanza dell'informazione geografica per la costruzione di alcuni indicatori nel *Geo White Paper on disaggregation by geographic location* e nella *SDGs Geospatial Roadmap*⁶, così come è stata sottolineata l'importanza di sviluppare analisi geostatistiche che consentano l'integrazione dei dati statistici e geografici con la disponibilità di dati georeferenziati utilizzabili per geografie variabili e misure statistiche fortemente disaggregate.

L'approccio geostatistico è, quindi, una delle chiavi essenziali per le analisi relative alla sostenibilità, utile per garantire l'armonizzazione delle informazioni, promuovere confronti e analisi di tendenze che non possono essere sviluppate considerando dati aggregati, facilitare una più dettagliata visualizzazione dell'informazione, consentendo di considerare l'informazione statistica anche per aree geografiche variabili e modificabili, indispensabili per il monitoraggio delle azioni. Nell'ambito delle misure statistiche scelte per il monitoraggio della Strategia nazionale di sviluppo sostenibile sono, infatti, risultate essere di particolare rilievo le questioni legate alla dimensione territoriale, essenziale per potere delineare le Strategie regionali per lo sviluppo sostenibile⁷. I dati desumibili da RSBL o disponibili per integrazioni realizzabili tramite l'utilizzo di RSBL possono essere indispensabili per analisi e azioni con riferimento, sia ai cambiamenti climatici, sia agli eventi estremi che ai disastri prima, durante e dopo l'evento estremo, per azioni di prevenzione, di adattamento, di mitigazione, che è opportuno monitorare tramite le informazioni statistiche (Figura 2.3).

Figura 2.3 - Le fasi e gli elementi relativi al *disaster risk management* (UNECE)



Fonte: UNECE

- L'IAEG-SDGs Working Group on Geospatial Information ha predisposto il *Geo White Paper on disaggregation by geographic location* e la *Statistical Commission* ha adottato la *SDGs Geospatial Roadmap* nel 2022. In entrambi i documenti si ribadisce che la disaggregazione territoriale, da sola o insieme alle altre possibili disaggregazioni degli indicatori, consente di mettere in luce le disuguaglianze sociali e di effettuare analisi relative a segmenti di popolazioni vulnerabili e alle aree che presentano situazioni di marginalità. Cfr. <https://ggim.un.org/UNGGIM-wg6/>.
- Si è stabilito di utilizzare misure statistiche desumibili dalla piattaforma Istat-Sistan, preferibilmente identiche a indicatori UN-IAEG-SDGs e coerenti con gli indicatori di Benessere equo e sostenibile (Bes), anche per permettere che le misure rispettino i requisiti di ammissibilità statistica. Sono utilizzati i criteri di parsimonia, fattibilità, tempestività, estensione e frequenza delle serie temporali, sensibilità alle politiche pubbliche, dimensione territoriale, focalizzando l'attenzione su misure statistiche che avessero la migliore disaggregazione territoriale possibile. <https://www.mase.gov.it/pagina/la-snsvs>.

Anche considerando le misure statistiche relative ai cambiamenti climatici (UNECE 2024) e gli indicatori relativi agli eventi estremi e ai disastri⁸, è fortemente raccomandato da UNECE (UNECE 2020, UNECE 2021a, UNECE 2023) che i dati statistici siano riferiti alla scala geografica più accurata e dettagliata possibile, fino alla coordinata geografica e che si lavori in termini di dati geografici e georeferenziati.

Le analisi ottenibili sulla base delle risultanze di RSBL, a titolo puramente esemplificativo, potrebbero fare riferimento a:

- produzione di un *toolset* di indicatori statistici – sociali, economici e ambientali – sempre pronti e disponibili in caso di eventi estremi;
- analisi ex ante di aree che presentano un elevato rischio di incendi, alluvioni, frane o terremoti oppure di aree contaminate;
- analisi ex post di aree colpite da disastri naturali;
- individuazione di edifici a una specifica distanza da fiumi, laghi, mare oppure vicini a industrie o ad aree contaminate, ai fini di azioni di prevenzione;
- individuazione di edifici in aree colpite da incendi, alluvioni, frane, terremoti, eventi meteorologici estremi oppure di aree contaminate, ai fini di azioni di recupero;
- individuazione di edifici o aree interessate da gravi ondate di calore nei grandi comuni;
- informazioni statistiche legate al grado di copertura del verde nei maggiori centri urbani, utile per produrre indicatori relativi al cambiamento climatico e alla sostenibilità;
- analisi di inquinamento atmosferico o terrestre considerando aree territoriali dettagliate e la popolazione esposta;
- analisi del consumo di suolo;
- analisi legate ai diversi ecosistemi, analisi delle aree protette;
- analisi legate al consumo di energia utilizzando i dati degli edifici residenziali e non, soprattutto in aree urbane.

I dati relativi alla popolazione legale 2021, diffusi sulla griglia regolare con celle di un 1 km², e la possibilità di visualizzarli, sono oggetto di analisi dettagliata nei Capitoli 10 e 12 ma di seguito (cfr. Figure 2.4, 2.5, 2.6) si riportano alcuni utilizzi per le inondazioni avvenute in Emilia Romagna nel 2023 e nel 2024 e relativamente agli incendi che hanno avuto luogo nel 2023 in Sicilia.

L'Istat assicura già misure statistiche disaggregate territorialmente e ne ha incrementato nel tempo la disponibilità. Grazie all'utilizzo del Registro statistico di base dei luoghi, la diffusione di informazioni statistiche di elevato dettaglio territoriale e geografico sarà ulteriormente incrementata, rendendo disponibile un mosaico statistico e informativo via via più ricco, nel pieno rispetto di quanto stabilito e auspicato in ambito internazionale e delle normative nazionali in tema di riservatezza dell'informazione statistica.

Il linguaggio comune creato dalle misure statistiche garantisce una cultura condivisa. L'informazione statistica, costruita grazie all'utilizzo di RSBL, e gli sviluppi che tale registro può garantire costituiscono lo strumento indispensabile per ampliare l'utilizzo anche di geografie comuni finalizzate alla costruzione della necessaria consapevolezza statistica. Gli sviluppi futuri garantiranno il pieno sfruttamento delle potenzialità di RSBL per costruire

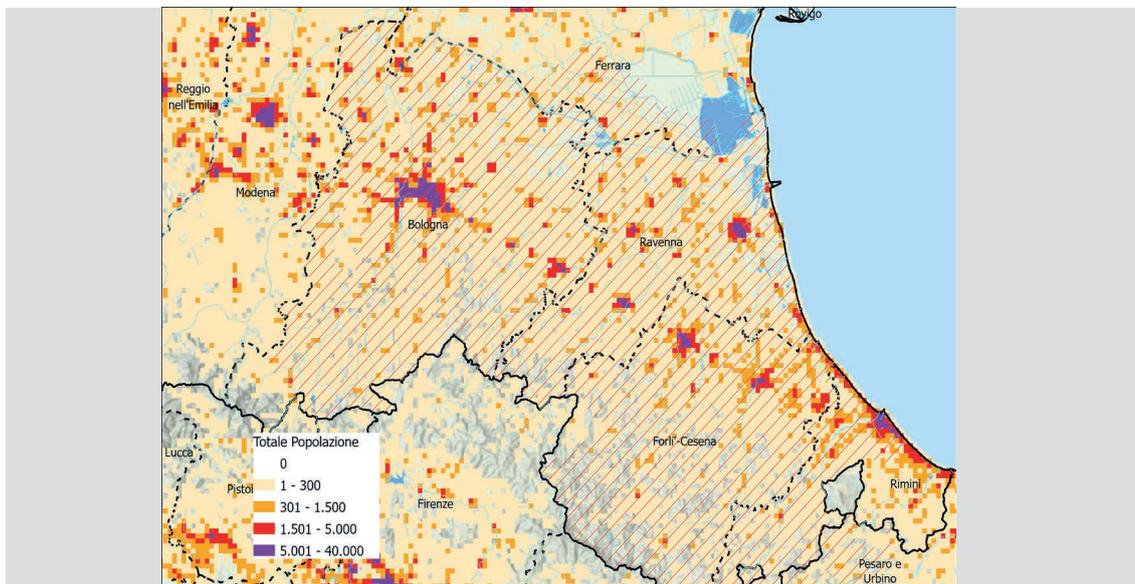
8 In ambito UNECE e UNSD sono state sviluppate linee guida e definiti *set* di indicatori chiave, selezionati secondo criteri di rilevanza, disponibilità del dato, confrontabilità tra Paesi, sia per le statistiche connesse ai cambiamenti climatici, sia per quelle connesse al rischio da disastri, organizzati per aree e argomenti:

- UNECE ha definito un *set* di 44 indicatori per le statistiche connesse ai cambiamenti climatici (CCRS), per le 5 aree: determinanti, emissioni, impatti, mitigazione, adattamento;
- UNSD un *set* di 158 indicatori per le statistiche connesse ai cambiamenti climatici (CCRS) per le 5 aree: determinanti, impatti, vulnerabilità, mitigazione, adattamento;
- UNECE ha sviluppato anche un *set* di 54 indicatori per il rischio da disastri (DRS).

2. Il Registro statistico di base dei luoghi - RSBL

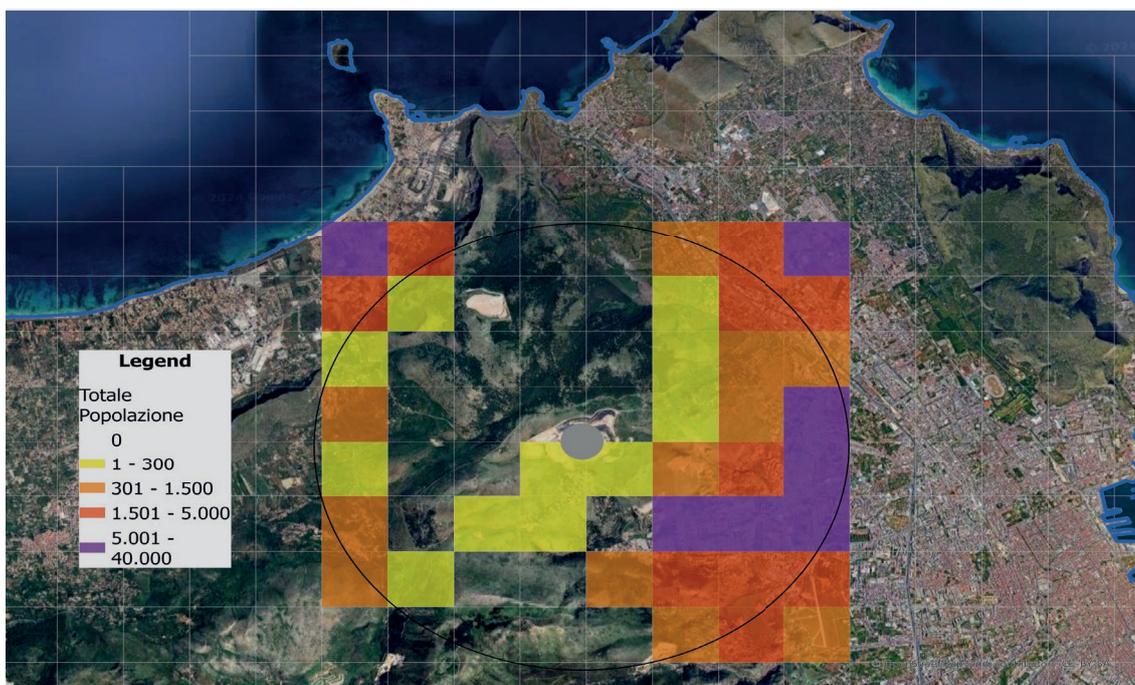
informazione geostatistica che agevolmente consenta di effettuare analisi dal locale al globale e dal globale al locale.

Figura 2.4 - Griglia di popolazione in Emilia Romagna le aree delle inondazioni



Fonte: Elaborazioni su dati Istat

Figura 2.5 - Griglia di popolazione con riferimento agli incendi sviluppatasi attorno al Comune di Palermo nel luglio 2023



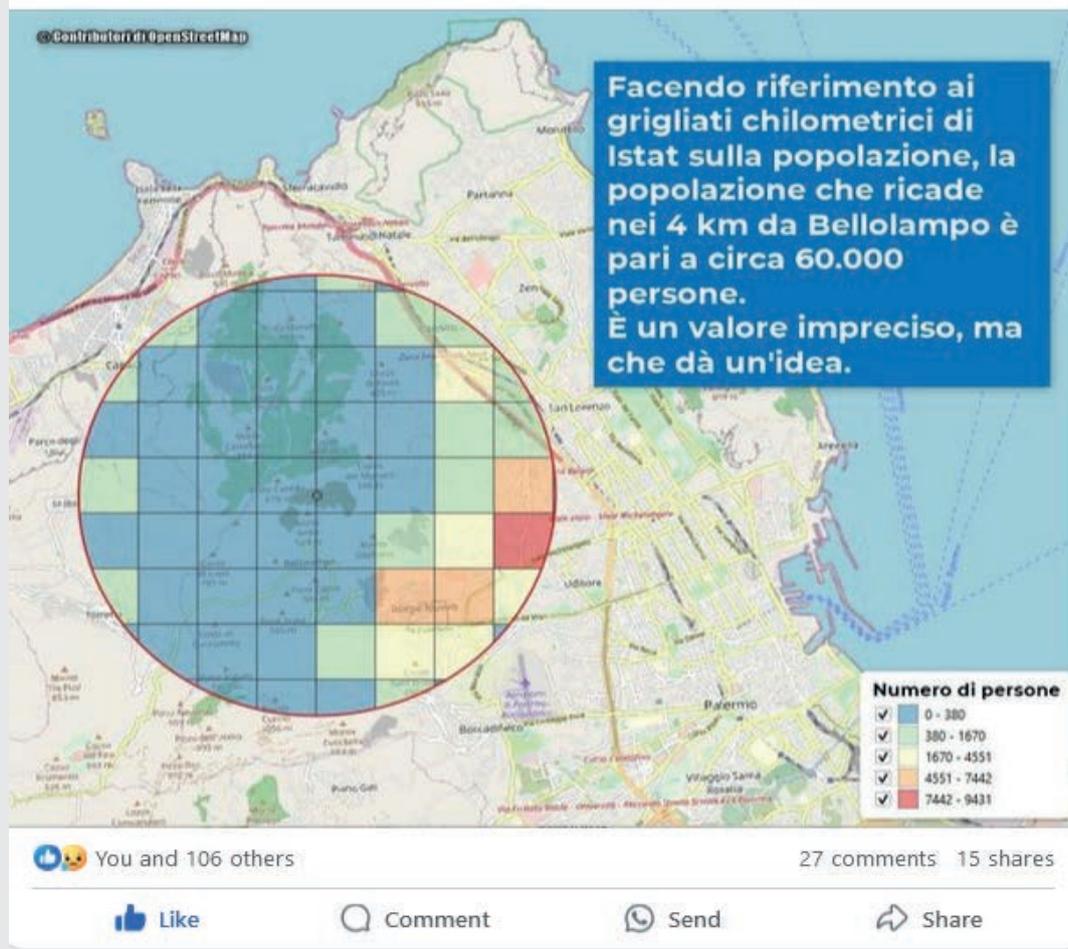
Fonte: Elaborazioni su dati Istat

Figura 2.6 - Utilizzo della griglia di popolazione con riferimento agli incendi sviluppatasi attorno al Comune di Palermo nel luglio 2023

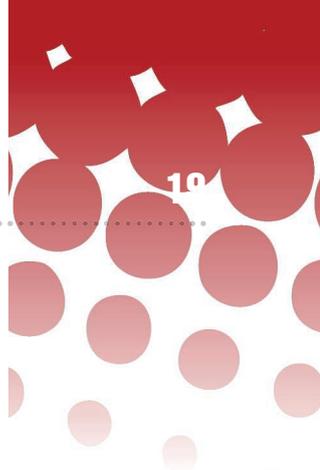
Ho fatto una rapida stima della popolazione che ricade nei 4 km citati nell'ordinanza sulla diossina di ieri del sindaco [Roberto Lagalla](#).

Secondo i grigliati chilometrici di Istat sulla popolazione, si tratta di circa 60.000 persone.

Se riesco, pubblico dati e mappa interattiva (ma sono giornate per me complicate, non sarò veloce)



Fonte: Elaborazioni su dati Istat



3. L'ARCHITETTURA E IL PROCESSO¹

3.1 Le funzioni di RSBL nel contesto del SIR

Il Registro dei luoghi è definito come uno dei registri base del SIR (Sistema integrato dei registri) (Istat 2016*b*). La sua peculiarità, nell'ambito del SIR, è rappresentata da un triplice ruolo di:

- produzione di statistiche sulla base delle caratteristiche geofisiche e ambientali caratterizzate dai suoi contenuti, ma anche da contenuti socio ed economici ottenuti dall'integrazione con altri registri del SIR;
- geolocalizzazione o di attribuzione, a specifici territori, delle unità degli altri registri;
- identificazione delle unità statistiche che alimentano le diverse componenti del registro e di supporto all'identificazione delle unità statistiche di altri registri (per esempio, unità locali di ASIA).

La funzione di localizzazione delle unità statistiche del SIR comporta non solo la predisposizione di tutti i processi di scambio delle informazioni di localizzazione con tutti gli altri registri del sistema (cfr. Capitolo 9) ma implica anche la progettazione e la programmazione di tutte le attività di aggiornamento e riallineamento dei vari sistemi (ad esempio il SIM) e di tutte le componenti di RSBL (per gli indirizzi cfr. Capitolo 7, per gli edifici e gli immobili cfr. Capitolo 8).

Il Registro dei luoghi, per la sua complessità di contenuti, è stato organizzato in componenti secondo tematismi specifici, a loro volta collegati da chiavi uniche che li mettono in relazione, come un ecosistema di registri nel RSBL.

I diversi elementi delle componenti (cfr. Capitolo 2) sono: gli indirizzi, le unità amministrative, le Basi territoriali, le Griglie regolari, Grafi stradali, le coordinate geografiche, gli edifici e le abitazioni e le partizioni amministrative o statistiche del territorio.

L'architettura del registro ha previsto, per ognuna di queste componenti, un processo standard per la loro definizione e caricamento. Per ognuna di esse sono stati previsti il processo di costruzione delle relazioni con le altre componenti e il processo di aggiornamento dei dati nel tempo.

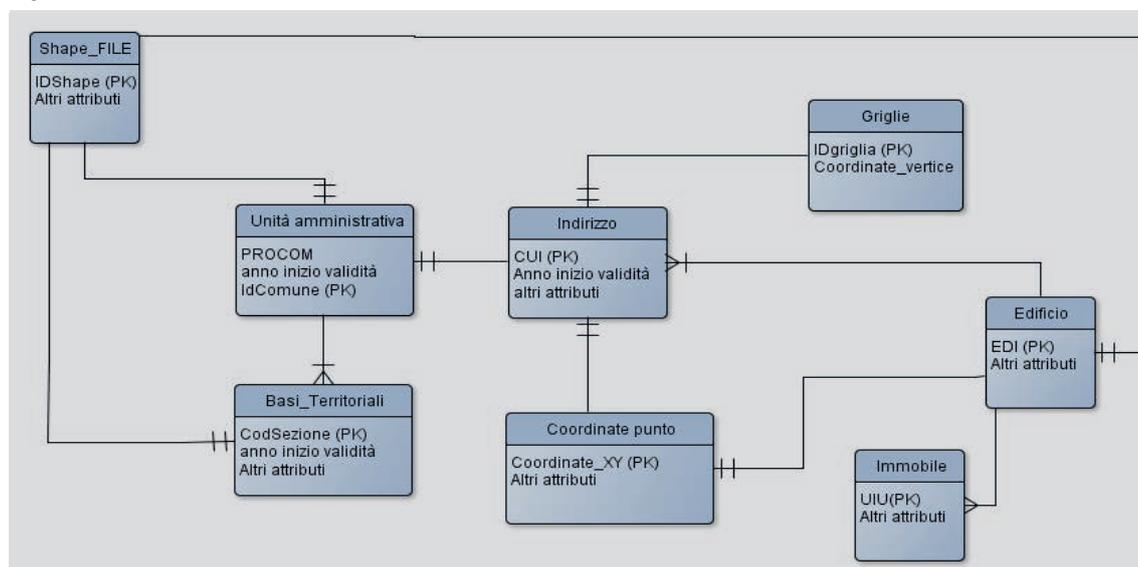
La Figura 3.1 mostra come sono modellate le componenti del registro, le chiavi identificative delle unità statistiche e le loro relazioni.

In ottemperanza ai principi della governance del SIR, definiti nel 2016 (Istat 2016), in cui è previsto che “le variabili dei registri sono riportate in uno solo di essi assicurando la coesione tra le diverse elaborazioni ottenute”, si è proceduto, per ogni unità statistica, alla definizione di tutte le informazioni a esse attribuite oltre che i codici identificativi. Nei registri che si integrano con RSBL sono presenti soltanto i codici identificativi, ad esempio nel Registro base degli individui (RBI) l'informazione relativa all'indirizzo di residenza non è esplicitata con l'indirizzo completo, ma soltanto con il codice CUI (Codice unico di indirizzo) e così per il comune di residenza identificato dal codice del comune (Procom).

¹ Il Capitolo è stato redatto da Raffaella Maria Aracri e Roberta Radini.



Figura 3.1 - Schema delle componenti di RSBL e delle loro relazioni



Fonte: Elaborazioni su dati Istat

3.2 Il processo di aggiornamento del registro

Il registro, ormai a regime dal 2018, è dotato di un processo di caricamento che realizza gli aggiornamenti annuali in funzione dei dati amministrativi che lo alimentano. Ognuna delle componenti dalle quali è composto, è alimentata da specifiche fonti amministrative e da processi che mettono in relazione le singole componenti di RSBL, rendendole coerenti. Quindi, annualmente vengono acquisite le fonti amministrative che al momento sono: le Liste anagrafiche comunali (LAC); l'Anagrafe tributaria; il Catasto. Sono in fase di studio altre fonti che alimenteranno la componente Registro degli indirizzi (cfr. Capitolo 7).

Le fonti amministrative vengono pretrattate dalle funzioni di raccolta dati, per distinguere le unità già analizzate negli anni precedenti e, quindi, già identificate, da quelle nuove o che compaiono per la prima volta nelle fonti amministrative. Questo delta di dati viene elaborato dal registro RSBL per essere identificato in termini di indirizzo, immobile e, alla fine, di edificio. Il dettaglio dell'implementazione del processo di caricamento viene descritto nei Capitoli 7 e 8.

In funzione della qualità dei dati le fonti vengono pretrattate, sia per migliorarne la qualità e massimizzare le funzioni di identificazione, sia per uniformare e standardizzare il formato dei dati spesso non uniforme da una fonte a un'altra.

Terminato il processo di riconoscimento, le singole unità statistiche vengono caricate nel registro e arricchite da indicatori di qualità che vengono messi a disposizione dell'utente finale.

Terminata la funzione di caricamento vengono avviate le attività di *linkage* delle unità statistiche delle diverse componenti del registro. Vengono collegate, ad esempio, le evoluzioni territoriali dei comuni agli indirizzi, oppure gli indirizzi agli edifici e gli edifici agli immobili.

Una volta rese coerenti e interrelate, le unità statistiche delle diverse componenti del registro vengono avviate alle attività di *linkage* con le unità statistiche degli altri registri del SIR, come descritto nel Capitolo 9, ad esempio, per l'integrazione di RBI e gli indirizzi di residenza.

3.3 La struttura semantica delle informazioni modellate dall'ontologia

Per il registro, come per tutti i registri del SIR, sono stati individuati criteri, metodologie e processi che garantiscono una integrazione coerente tra gli oggetti e le proprietà provenienti da più fonti informative.

Il sistema di integrazione adottato è modulato in integrazione concettuale, fisica e statistica.

L'integrazione concettuale consente di:

- identificare i metadati, in termini di definizioni, classificazioni di oggetti e proprietà delle unità statistiche del registro;
- identificare la coerenza e/o le modificazioni che avvengono negli oggetti e nelle proprietà nel tempo;
- identificare il livello di coerenza che esiste tra gli oggetti e le proprietà presenti nelle diverse componenti del registro e in altri registri.

L'integrazione fisica si basa su l'identificazione di:

- uno stesso oggetto presente in fonti diverse con un numero ID univoco o codice stabile nel tempo;
- la definizione, per ogni oggetto, delle relazioni logiche e fisiche, nel tempo e nello spazio, tra le informazioni disponibili da fonti diverse;
- l'identificazione di relazioni logiche e fisiche che si determinano tra oggetti che si riferiscono a una stessa tipologia (ad esempio relazioni tra comuni che variano nel tempo) o tra oggetti che afferiscono a differenti tipologie (ad esempio relazioni tra indirizzi ed edifici).

L'integrazione statistica consente di:

- mettere insieme oggetti e proprietà di più registri in modo da potere identificare popolazioni e sottopopolazioni statistiche, integrare variabili, derivare variabili;
- definire la coerenza tra informazioni desumibili dal registro e informazioni da indagini statistiche, al fine di correggere e/o valutare le informazioni statistiche direttamente derivabili dai registri.

L'integrazione fisica è descritta nel dettaglio nel Capitolo 11, mentre quella statistica nel Capitolo 13, dove viene illustrato come l'integrazione dei registri RBI e RSBL e l'indagine del Censimento ha consentito di realizzare i diversi prodotti statistici di popolazione per abitazione e per sezione di censimento.

Il sistema di integrazione concettuale è stato modellato e progettato come un vero e proprio sistema di gestione e rappresentazione, sia dei metadati sia dei dati, assolvendo anche alle funzioni di controllo di qualità e di gestione coerente delle evoluzioni temporali.

L'approccio adottato è l'*Ontology-based Data Management*, basato sulla definizione di ontologie che vengono mappate sui dati.

Alla base di questo *framework* è la definizione dell'ontologia del dominio del registro che consente di garantire l'integrazione concettuale non solo delle componenti dello stesso ma, soprattutto, del registro all'interno del SIR.

Infatti, le ontologie sono una rappresentazione formale, condivisa ed esplicita di una concettualizzazione di un dominio di interesse. In generale, la progettazione di una ontologia prevede: l'identificazione del dominio di interesse e gli scopi dell'ontologia; definizione dei concetti e degli individui (istanze di concetti); definizione delle proprietà intrinseche (proprie di un concetto) ed estrinseche dei concetti (relazioni tra concetti) (De Giacomo *et al.* 2018).

In particolare, l'ontologia del Registro statistico dei luoghi componente indirizzi è stata sviluppata per tutte le componenti dello stesso. Per brevità si riportano di seguito l'elenco di alcuni dei principali concetti dell'ontologia:

- sezione di censimento: unità territoriale minima di rilevazione del comune, sulla quale è organizzata la rilevazione censuaria. È costituita da un solo corpo delimitato da una linea spezzata chiusa. A partire dalle sezioni di Censimento sono ricostruibili, per somma, le entità geografiche e amministrative di livello superiore (località abitate, aree subcomunali, collegi elettorali e altre). Ciascuna sezione di Censimento deve essere completamente contenuta all'interno di una e una sola località. Il territorio comunale deve essere esaustivamente suddiviso in sezioni di Censimento; la somma di tutte le sezioni di Censimento ricostruisce l'intero territorio nazionale;
 - accesso esterno: rappresenta l'accesso esterno da un'area di circolazione a una unità ecografica semplice (abitazioni, uffici, locali in cui sono ubicate attività economiche, eccetera), l'accesso esterno si specifica in: accesso esterno non georeferenziato; accesso esterno georeferenziato da sezione Censimento; accesso esterno georeferenziato da coordinate;
 - indirizzo: rappresenta l'indirizzo nel suo complesso (odonomo e numerazione civica);
 - odonomo: denominazione, costituita dalla DUG (Denominazione urbanistica generica: via, piazza, largo, eccetera) e dalla DUF (Denominazione ufficiale), che serve a identificare univocamente un'area di circolazione;
 - comune: ente territoriale locale rappresentativo di una collettività stanziata su un determinato territorio, dotato di propria autonomia statutaria, normativa e amministrativa, secondo i principi fissati dalla Costituzione all' articolo 114 e dalla normativa nazionale (Testo unico degli enti locali);
 - coordinata geografica: identifica la posizione (x, y, h) di un accesso esterno sulla superficie terrestre.
- Di seguito l'elenco di alcune delle principali proprietà estrinseche (relazioni tra concetti)²:
- *haIndirizzo*: relazione che lega un indirizzo al suo stato temporale;
 - *haOdonimo*: relazione che lega un Indirizzo con il corrispondente Odonimo (DUG + DUF);
 - *ha_sezione_di_censimento*: relazione che lega una sezione di Censimento al suo stato temporale;
 - *censito_in*: relazione che lega un accesso esterno al suo stato temporale;
 - *ha_tipo_coordinata_geografica*: relazione che lega una coordinata geografica con la tipologia di coordinata;
- Sono state, infine, definite le seguenti proprietà intrinseche dei concetti:
- per indirizzo: specificità, valore_metrico, CAP, esponente, Numero_civico, Id_indirizzo.
- Nella Figura 3.2 viene mostrata la rappresentazione grafica dell'ontologia RSBL – componente indirizzi, espressa nel linguaggio *Graphol* (Aracri *et al.* 2017) (Console *et al.* 2014) conforme al linguaggio *OWL2* (vedi link).

² La nomenclatura delle relazioni prevede un verbo come *label* e l'oggetto al quale si applica. Per esempio, il suffisso *ha* per indicare una relazione di attribuzione, oppure si usano suffissi più pertinenti come *in*.

3.4 L'*Ontology-based Data Management* applicata a RSBL

L'utilizzo delle ontologie abilita l'adozione di una metodologia per la gestione del dato basato sulle ontologie, ovvero l'*Ontology-based Data Management* (OBDM).

L'OBDM (Lenzerini 2011; Lenzerini 2018) è un paradigma introdotto e promosso dal Dipartimento di ingegneria informatica automatica e gestionale "Antonio Ruberti" di Sapienza Università di Roma e da OBDA Systems, e ha come obiettivo l'integrazione e il governo dei dati memorizzati in un sistema informativo di una organizzazione attraverso una ontologia. Il fine di questo approccio è di creare un unico punto di accesso concettuale al patrimonio informativo dell'organizzazione stessa, per permettere di realizzare a livello teorico tutti i servizi di *data governance* in un sistema complesso.

L'OBDM può, in effetti, essere visto come una forma di *data integration* virtuale, con la differenza che esso si basa sull'idea di sostituire lo schema globale, ovvero la vista unificata del dominio, con una rappresentazione concettuale e formale dello stesso, formulata attraverso una ontologia espressa in un linguaggio logico.

Questa scelta fa sì che la vista integrata offerta da un sistema OBDM non si limiti a una struttura che accomoda i dati alle sorgenti, ma è una descrizione semanticamente ricca dei concetti rilevanti del dominio di interesse e delle relazioni che intercorrono tra di essi.

Così come avviene nei sistemi di *data integration*, vengono utilizzate delle relazioni concettuali, o *mapping*, per definire le corrispondenze semantiche tra lo schema globale e i dati nelle sorgenti, in questo caso la sorgente è il registro RSBL. Quando si vuole avere una risposta a una esigenza informativa, si esprime una *query* sull'ontologia (e non sulle basi di dati del sistema informativo) e le corrispondenze create tra i dati e i concetti dell'ontologia consentiranno al motore di ragionamento ontologico [ref. *MASTRO*] di calcolarne la risposta. Questo libera l'utente dall'obbligo di conoscere gli aspetti tecnici di memorizzazione dei dati e i dettagli riguardanti dove sono memorizzati e con quali meccanismi sono organizzati. Analogamente, quando si vuole svolgere un compito di *data governance* (*quality assessment*, reingegnerizzazione, *data cleaning*, eccetera), non si accede direttamente alle sorgenti informative, ma si svolgono le opportune funzioni attraverso l'ontologia di dominio.

Quest'ultima, oltre a descrivere formalmente il modello informativo del registro, diventa anche un mezzo per avvicinarsi a un approccio dichiarativo della *data governance*. Attraverso l'esplicitazione della rappresentazione del dominio, si acquista, infatti, riusabilità della conoscenza acquisita, cosa che invece non avviene quando lo schema globale è semplicemente una descrizione unificata delle variabili di un registro o di un sistema di registri.

L'ontologia di dominio di RSBL è in relazione con quella di altri registri come ad esempio di RBI. Quindi, tutti i principi dell'OBDM illustrati sono applicabili non solo al registro ma anche alla sua integrazione con altri registri. Inoltre, questa funzione è abilitante all'uso dei dati del registro nell'ambito delle applicazioni di Intelligenza artificiale (IA) sia per il controllo della qualità attraverso le funzioni di *reasoning* comprese nel OBDM, sia per le funzioni di inferenza che consentono di recuperare informazioni non esplicitamente modellate.

Il modello OBDM implementato consente di dare accesso ai dati anche in forma limitata ossia soltanto ai dati ai quali l'utente è autorizzato attraverso la predisposizione di profili di accesso.

Non ultimo consente l'estrazione di dati in formato RDF (*Resource Description Framework*) che è alla base della codifica, dello scambio e del riutilizzo di metadati strutturati e consente l'interoperabilità semantica, oltre che la pubblicazione in formato *Linked Open Data*. Il formato dei dati RDF è un linguaggio standard machine readable che favorisce l'uso

3. L'architettura e il processo

dei dati del registro per alimentare le applicazioni basate su modelli IA come ad esempio i *Large Language Model* (LLM), dove i dati del registro per la loro elevata qualità e affidabilità rappresentano *Ground truth*.

Quest'ultima funzione può essere sfruttata per la realizzazione degli *high-value data* (HVD) come previsto dalla Commissione europea nel Regolamento di esecuzione (regolamento UE n. 138/2023).

4. SITUAS: IL SISTEMA INFORMATIVO TERRITORIALE DELLE UNITÀ AMMINISTRATIVE E STATISTICHE PER L'ANALISI DELLA STRUTTURA E LA DINAMICA DEL TERRITORIO¹

4.1 SITUAS in RSBL e il portale per l'utenza

Il Sistema informativo territoriale delle unità amministrative e statistiche (SITUAS) raccoglie in maniera organica e coerente il patrimonio dell'Istituto Nazionale di Statistica (Istat) su codici, denominazioni, classificazioni e variazioni degli enti territoriali amministrativi e delle unità statistiche e costituisce un imprescindibile strumento di consultazione, interrogazione e analisi per la conoscenza territoriale. SITUAS da un lato, in qualità di componente del Registro statistico di base dei luoghi (RSBL), costituisce la fonte per gli altri registri dei dati finalizzati alla gestione degli assetti amministrativi del territorio, dall'altro, in qualità di sistema informativo, si avvale del portale <https://situas.istat.it>, in continua evoluzione, per il rilascio di informazione all'utenza. Le funzionalità studiate e implementate in SITUAS sono pensate affinché il sistema possa consentire la gestione di sempre nuove geografie e classificazioni territoriali di elevata rilevanza provenienti anche da altri enti e istituzioni del Sistema statistico nazionale.

Il sistema informativo documenta tutti gli enti territoriali riconosciuti dall'ordinamento amministrativo del Regno d'Italia prima e, successivamente, da quello costituzionale della Repubblica italiana, nonché le aree statistiche di interesse per le *policy* nazionali ed europee. Le unità territoriali descritte sono definite genericamente con il termine di "unità amministrativa" (Ua) o "unità statistica" (Us).

Il contenuto informativo offerto da SITUAS consente di analizzare l'evoluzione dell'organizzazione amministrativa sul territorio nazionale, di ricostruire l'assetto (composizioni e relazioni di appartenenza) delle partizioni amministrative o statistiche di interesse istituzionale individuate nel tempo, relazionando il dato numerico al reale territorio di riferimento, ovvero al territorio dell'epoca. Consente anche di ricostruire o confrontare le unità territoriali e le relative codifiche in virtù dei cambiamenti intervenuti nel tempo.

Nello specifico, impostando la variabile tempo (una singola data o un periodo), l'ambito geografico (amministrativo o statistico) e la tipologia di unità di interesse (per esempio, comuni, province, eccetera), il portale consente all'utente di scegliere tra una serie di *report* tematici.

Questi *report* offrono informazioni, su:

- demografia, dimensione, caratteristiche e classificazioni delle unità territoriali a una data di interesse;
- tipologie di variazioni intervenute nel periodo ricompreso tra due date a scelta (costituzioni, estinzioni, cambi denominazione, eccetera);
- approfondimenti su dettagli ed estremi normativi recanti le variazioni.

¹ Il Capitolo è stato redatto da Tiziana Clary (paragrafi 4.1, 4.2 e 4.4), Luisa Franconi (paragrafo 4.3), Giorgia Bonamano (paragrafo 4.5), Giuseppe Busanello (paragrafi 4.6 e 4.6.1) e Daniela Giacomini (paragrafo 4.6.2).

Il portale, inoltre, prevede funzionalità e servizi, che via via si arricchiranno nel tempo, dedicati alla confrontabilità delle geografie amministrative² o statistiche nel tempo e alla ricostruzione delle relative codifiche in un dato periodo di interesse, oltre alla possibilità di ricercare codici e denominazioni delle unità territoriali amministrative³ con approfondimenti sul relativo periodo di validità.

Tutti i termini e le classificazioni utilizzati nel sistema sono consultabili nel *Geoglossario* del portale⁴.

4.2 Unità territoriali amministrative: comuni, province/Città metropolitane e regioni

Uno degli elementi costitutivi dello Stato è il territorio, su cui si articolano enti che concorrono alla realizzazione delle sue funzioni. Le regioni, le province e i comuni sono enti pubblici⁵, il cui elemento costitutivo è il territorio. La suddivisione territoriale del Paese costituisce una base di riferimento dei dati di fondamentale importanza per l'utilizzo degli stessi ai fini statistici.

È in tale contesto che si pone il Sistema informativo territoriale delle unità amministrative e statistiche (SITUAS) quale strumento di governo per una corretta rappresentazione del territorio e delle relative classificazioni ai fini della produzione e diffusione delle statistiche ufficiali nazionali e internazionali.

Attraverso la gestione delle variazioni intervenute nel tempo e documentate dagli oltre 4.200 provvedimenti (leggi, decreti, eccetera) identificati e registrati nel proprio database, SITUAS traccia la storia di tutte le unità amministrative territoriali esistite ed esistenti, con particolare riferimento alla ricostruzione delle codifiche statistiche assegnate a ognuna a partire dal livello minimo di riferimento, il comune, fino a risalire al livello più alto, la regione.

Le variazioni, che per le loro differenti caratteristiche incidono in diverso modo e grado nel disegno della geografia amministrativa, possono riguardare la tipologia, la numerosità, i confini, la denominazione e le relazioni di appartenenza. Tali eventi sono acquisiti attraverso una ricognizione quadrimestrale condotta dall'Istat presso le regioni e le province autonome, alimentando un patrimonio informativo storico consistente (dall'Unità d'Italia a oggi), prodotto a partire dalla documentazione statistica detenuta dall'Istat e integrata da una sistematica attività di recupero e analisi delle fonti storiche dedicate.

In risposta all'evoluzione dell'organizzazione amministrativa del territorio nazionale, SITUAS consente di elaborare livelli territoriali, quali: comuni, province o unità territoriali sovracomunali⁶ (dal 2015) e regioni (dal 1948). A questi si aggiungono partizioni "storiche" non più esistenti come i circondari (1861-1927), i distretti (1861-1915) e i compartimenti (1861-1947).

2 Per consultare i contenuti si rimanda alla pagina <https://situas.istat.it/web/#/ricostruzione-codici>.

3 Per la funzionalità di ricerca del codice si rimanda alla pagina <https://situas.istat.it/web/#/ricerca-codice>.

4 Il *Geoglossario* è consultabile alla pagina <https://situas.istat.it/web/#/glossario>.

5 Dal 2015, la fisionomia del territorio amministrativo italiano ha assunto nuovi contorni e nuovi significati, in particolare a livello provinciale. Nelle regioni a statuto ordinario per effetto della legge n. 56/2014 hanno trovato attuazione le città metropolitane, mentre nelle regioni a statuto speciale, dove la disciplina è demandata alle leggi regionali, le azioni di riordino sono state più radicali prevedendo, oltre le città metropolitane, anche i liberi consorzi dei comuni (Sicilia) o la soppressione degli enti di secondo livello come nel caso delle province del Friuli-Venezia Giulia e quindi determinando un vuoto nel completamento della partizione territoriale.

6 L'Istat, in risposta alle leggi di riforma degli assetti amministrativi del territorio e a supporto dell'informazione statistica per il Paese, ha introdotto la definizione di "unità territoriale sovracomunale", in cui includere oltre alle province, anche le città metropolitane, i liberi consorzi comunali e le ex province friulane come "unità non amministrative".

4. SITUAS: il Sistema informativo territoriale delle unità amministrative e statistiche per l'analisi della struttura e la dinamica del territorio

A partire dall'Unità d'Italia, sono oltre 10.800 le unità territoriali presenti nel sistema, di cui 10.342 sono comuni; unità che nel tempo hanno richiesto l'emissione e la gestione di oltre 14 mila codici statistici per una identificazione univoca rispetto al tempo e allo spazio in virtù delle variazioni avvenute (Tavola 4.1).

Tavola 4.1 - Numero di comuni e codici comunali per ripartizione geografica. Anni 1861-2024 (valori assoluti)

Ripartizioni geografiche	Comuni (a)	di cui attivi (b)	Codici (c)	Comuni con:								Codici denominazione storico (d)	Totale codici (c) (d)
				Un codice	di cui attivi	Due codici	di cui attivi	Tre codici	di cui attivi	Quattro codici	di cui attivi		
Nord-ovest	4.359	2.990	5.590	3.212	2.022	1.063	897	84	71	-	-	1.074	6.664
Nord-est	2.148	1.387	3.606	1.074	893	690	359	384	135	-	-	586	4.192
Centro	1.128	968	1.602	723	590	336	310	69	68	-	-	320	1.922
Sud	1.894	1.783	2.384	1.406	1.306	486	475	2	2	-	-	535	2.919
Isole	813	768	1.363	467	424	205	203	78	78	63	63	152	1.515
ITALIA	10.342	7.896	14.545	6.882	5.235	2.780	2.244	617	354	63	63	2.667	17.212

Fonte: Istat, Sistema informativo territoriale delle unità amministrative e statistiche (SITUAS)

(a) Numero di comuni esistiti ed esistenti nel periodo 1861-2024. Il computo dei comuni è ricondotto all'ultima ripartizione geografica di appartenenza rispetto alla data del 22 gennaio 2024.

(b) Numero di comuni esistenti alla data del 22 gennaio 2024.

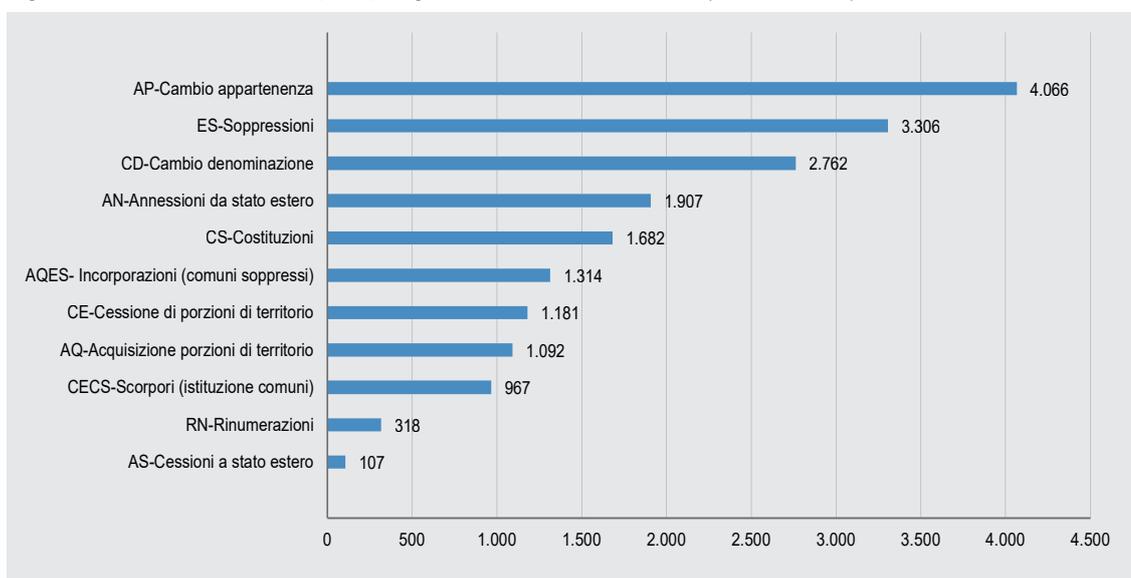
(c) Totale dei codici statistici emessi nel periodo tra il 1861 e il 22 gennaio 2024. Il conteggio è dato dalla somma dei codici assegnati a ciascun comune esistente ed esistito in virtù delle variazioni subite.

(d) Totale degli identificativi assegnati alle denominazioni precedenti dei comuni esistiti ed esistenti nel periodo 1861-2020.

Al numero dei codici ufficiali assegnati alle unità amministrative si aggiungono i "codici denominazione storico", ovvero le codifiche associate alle denominazioni non più vigenti a seguito di eventi di cambio nome (dal 2020 non più oggetto di emissione) per un totale di 17.212 identificativi (Tavola 4.1).

Attraverso la ricostruzione e la documentazione della storia amministrativa dei livelli territoriali descritti, il sistema permette di rilevare come nell'arco di oltre 160 anni di storia del Paese si siano verificate circa 19 mila eventi di variazione, di cui oltre 18 mila a livello comunale (Figura 4.1).

Figura 4.1 - Variazioni comunali per tipologia di evento. Anni 1861-2024 (valori assoluti)



Fonte: Istat, Sistema informativo territoriale delle unità amministrative e statistiche (SITUAS)



Il primato spetta agli eventi di cambio di appartenenza (21,7 per cento), risultato dovuto alla nati-mortalità delle province o alla modifica dei loro confini per scambio di comuni, occorrenze dalle quali discendono le attribuzioni di nuovi codici per l'identificazione del comune nell'ambito della nuova provincia di appartenenza. Sono consistenti anche gli eventi di soppressione dei comuni, ovvero di cessazione della validità del relativo codice, pari al 17,7 per cento che salgono al 18,2 per cento se vengono considerate anche le cessioni a stato estero, e infine i cambi denominazione pari al 14,8 per cento. Le costituzioni di nuovi comuni rappresentano il 9,0 per cento del totale Italia (ovvero 18.702 eventi complessivi), che salgono al 19,2 per cento se calcolate anche le occorrenze di acquisizione di nuove unità da stato estero.

4.3 Unità territoriali statistiche

Le unità territoriali statistiche (Us), generalmente definite da aggregazione di comuni, hanno confini determinati da processi di organizzazione del territorio di tipo istituzionale e/o socioeconomico e/o da principi di omogeneità rispetto a caratteristiche ambientali o morfologiche. Possono rappresentare aree di competenza istituzionale di amministrazioni o istituzioni che sono acquisite dalla statistica ufficiale come elemento predefinito e la cui demografia è soggetta ad atti, normative, regolamenti, eccetera, oppure sono risultanti da elaborazioni statistiche (condotte, ad esempio, dall'Istat, dal Sistema statistico nazionale o europeo o secondo metodologie definite a livello internazionale) per la definizione di domini di studio o sottopopolazioni di particolare interesse, alle quali riferire l'informazione statistica. Nel sistema informativo SITUAS la definizione e gestione delle unità statistiche sono caratterizzate da estrema flessibilità.

Questa si esplicita sia in termini di unità componente (ovvero la possibilità di creare relazioni logiche non solo con i comuni, ma anche tra le diverse unità statistiche), sia in termini di definizione delle associazioni rispetto al tempo. In particolare, il sistema consente la gestione del tempo in maniera flessibile attraverso due modalità. Da un lato, permettendo di elaborare le relazioni secondo periodi di tempo personalizzabili e, dall'altro, di "versionare" una geografia per mantenere la confrontabilità in serie storiche nella composizione delle Us.

Attualmente in SITUAS è presente la composizione delle seguenti unità territoriali statistiche:

- Sistemi locali del lavoro⁷ (2001 e 2011), Istat (2015a): unità essenziali nello studio di fenomeni socio-economici e utilizzate per *policy* specifiche⁸;
- Distretti industriali⁹ (2011), Istat (2015b): unità che rappresentano l'elemento portante nella struttura produttiva del Paese;
- Zone urbane funzionali (*Functional urban areas* - FUA, 2011)¹⁰ (Dijkstra *et al.* 2019): unità per il confronto del fenomeno urbano a livello internazionale;

7 Insieme di comuni contigui legati fra loro da flussi di pendolarismo. Cfr. <https://situas.istat.it/web/#/glossario>.

8 Ulteriori dettagli al link: <https://www.istat.it/it/informazioni-territoriali-e-cartografiche/sistemi-locali-del-lavoro>.

9 Sistemi locali caratterizzati dalla presenza di piccola o media impresa, con una elevata concentrazione territoriale di occupazione manifatturiera focalizzata in una industria principale. Per approfondimenti cfr. <https://situas.istat.it/web/#/glossario>.

10 Raggruppamento di comuni contigui costituito dalla city (comune con elevata densità di popolazione e con almeno 50 mila abitanti) e dai comuni circostanti (zona di pendolarismo) Regolamento (UE) 2017/2391, Tercet; cfr. <https://situas.istat.it/web/#/glossario>. Per maggiori informazioni cfr. <https://www.istat.it/it/informazioni-territoriali-e-cartografiche/cities-and-functional-urban-areas>.

4. SITUAS: il Sistema informativo territoriale delle unità amministrative e statistiche per l'analisi della struttura e la dinamica del territorio

- Nuts (Nomenclatura delle unità territoriali statistiche)¹¹: Le regioni della statistica ufficiale europea.

Il sistema è predisposto per una gestione da parte di più utenti delle Us in modo da permettere l'acquisizione e manutenzione di geografie dei più disparati ambiti tematici da affidare a specialisti settoriali.

4.4 Unità territoriali estere

I paesi esteri rappresentano una ulteriore componente territoriale di SITUAS, al momento in corso di implementazione, allo scopo di renderne fruibili i dati secondo modalità di interrogazione personalizzabili e dinamiche nel tempo.

Tra i compiti istituzionali dell'Istat, ai fini della produzione e diffusione delle statistiche ufficiali nazionali ed europee, vi è infatti anche la gestione dei paesi esteri, distinti tra stati e territori. Questi ultimi sono corrispondenti ad aree geografiche soggette alla sovranità di uno Stato nazionale e riconosciuti come luoghi di interesse per le finalità statistiche dettate da regolamenti europei.

L'Istat assegna un codice statistico ai paesi esteri, ne aggiorna le denominazioni e registra i cambiamenti che avvengono nella configurazione geopolitica mondiale. Gli elenchi finora prodotti¹² consentono l'elaborazione di informazioni che spaziano dalle tipologie di cambiamenti intervenuti nel tempo, alla composizione delle unità a una data di riferimento con le relative caratteristiche e nomenclature internazionali associate, quali ad esempio: i codici catastali (codice nazionale emesso dall'Agenzia delle entrate) e classificazioni internazionali¹³.

4.5 Caratteristiche e classificazioni delle unità territoriali

Con riferimento alle principali unità territoriali di interesse per la statistica ufficiale e, in particolare, per i comuni il sistema documenta:

- le classificazioni elaborate dall'Istat che trovano il loro fondamento su caratteri geomorfologici e ambientali, misurati a soli fini statistici come la litoraneità, le zone altimetriche, le ecoregioni o la superficie territoriale;
- talune caratteristiche di fonte amministrativa come i codici fiscali, i codici catastali, la sigla automobilistica, il comune capoluogo di provincia o regione, eccetera;
- alcune classificazioni di elevata rilevanza nazionale predisposte da altri enti o organismi pubblici a fini di *policy* come, ad esempio, le Aree interne¹⁴ (Istat e Presidenza del Consiglio dei Ministri) o le Aree soggette a vincoli naturali¹⁵ (in collaborazione con il Crea-Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria, utilizzabile per soli fini statistici);

11 Regolamento (CE) n. 1059/2003; si rimanda alla pagina <https://ec.europa.eu/eurostat/web/nuts/history>. Di prossima pubblicazione nel portale SITUAS insieme alle relative classificazioni.

12 Disponibili all'indirizzo <https://www.istat.it/it/archivio/6747> con aggiornamento annuale alla data del 31 dicembre.

13 Classificazioni ISO 3166-1 Alpha 2 e 3 e UNSD_M49 Standard country or area codes for statistical use (M49).

14 Classificazione dei comuni in base alla loro distanza da tre tipologie di servizio (salute, istruzione e mobilità). Cfr. <https://situas.istat.it/web/#/glossario>.

15 Aree dove gli agricoltori affrontano maggiori costi di produzione e possono beneficiare di pagamenti compensativi calcolati in base ai costi aggiuntivi sostenuti e al mancato guadagno. Cfr. <https://situas.istat.it/web/#/glossario>.



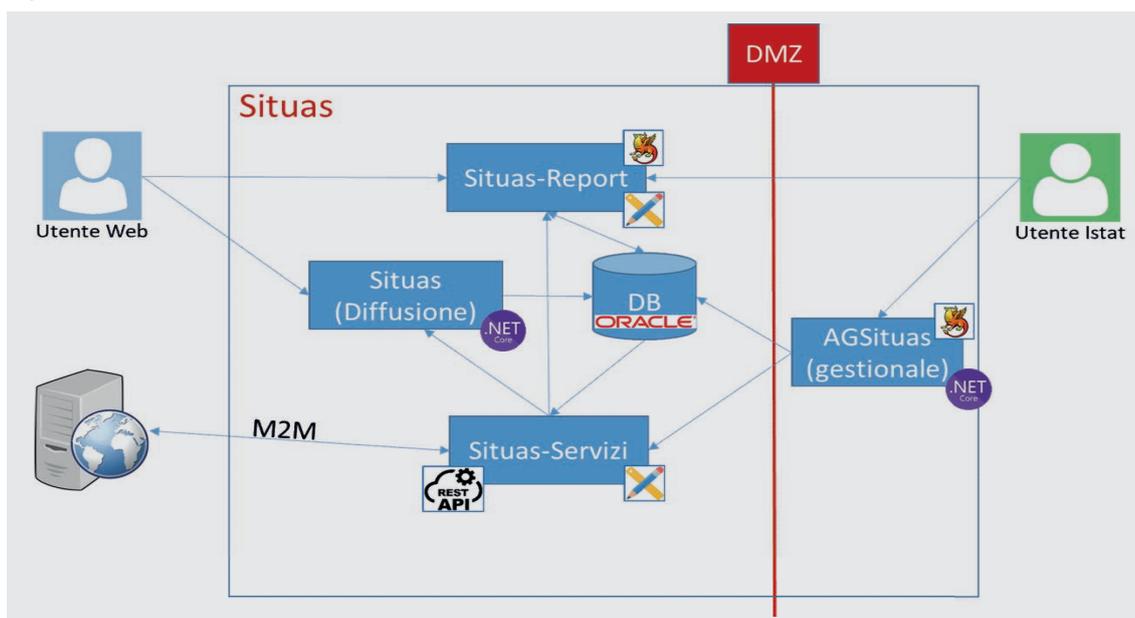
- le nomenclature delle regioni europee Nuts (Nomenclatura delle Unità territoriali statistiche) e le classificazioni derivanti da regolamenti europei quali ad esempio il grado di urbanizzazione dei comuni¹⁶, le zone costiere¹⁷.

Anche per le unità territoriali statistiche il Sistema documenta classificazioni sperimentali¹⁸ o derivanti da regolamenti europei¹⁹. Le classificazioni sono aggiornate correntemente nel sistema in funzione delle eventuali variazioni che occorrono nel tempo e rese fruibili nel portale mediante *report* tematici dedicati (ad esempio, Aree di *policy*, *report* su Dimensione, Caratteristiche, eccetera) e consultabili a una data di interesse.

4.6 Architettura del Sistema

Il Sistema informativo territoriale delle unità statistiche territoriali (SITUAS) è costituito da quattro moduli applicativi, di cui tre dedicati all'esposizione sul web e uno per la gestione interna Istat, area protetta da *firewall* (DMZ) (Figura 4.2).

Figura 4.2 - L'architettura di SITUAS e i flussi coinvolti



Fonte: Istat

4.6.1 I moduli SITUAS per la diffusione

SITUAS - Diffusione è il portale per la pubblicazione di dati aggiornati su unità amministrative e statistiche. Il *back end* è realizzato in .NET 6.0 mentre il *front end* fa uso della libreria *Javascript React*, utilizzando *Redux* come componente per la gestione dello stato dell'applicazione. Per il rilascio dei dati è prevista l'esposizione di *API Rest*. Quelle utilizzate

¹⁶ Classificazione armonizzata internazionale dei comuni sulla base di combinazioni di densità e soglia di popolazione minima di celle contigue di 1 km² della griglia regolare Geostat. Cfr. <https://situas.istat.it/web/#/glossario>.

¹⁷ Classificazione europea dei comuni secondo il grado di vicinanza alla costa. Cfr. <https://situas.istat.it/web/#/glossario>.

¹⁸ La visualizzazione delle classificazioni sperimentali dei SLL è disponibile in Gistat.

¹⁹ Classificazioni del Regolamento Tercet di prossima pubblicazione sul portale di SITUAS.

4. SITUAS: il Sistema informativo territoriale delle unità amministrative e statistiche per l'analisi della struttura e la dinamica del territorio

da *SITUAS - Diffusione* sono di libero accesso, mentre è prevista la pubblicazione di *API Rest* specifiche per consentire a terze applicazioni di recuperare dati personalizzati non necessariamente diffusi o diffondibili al pubblico.

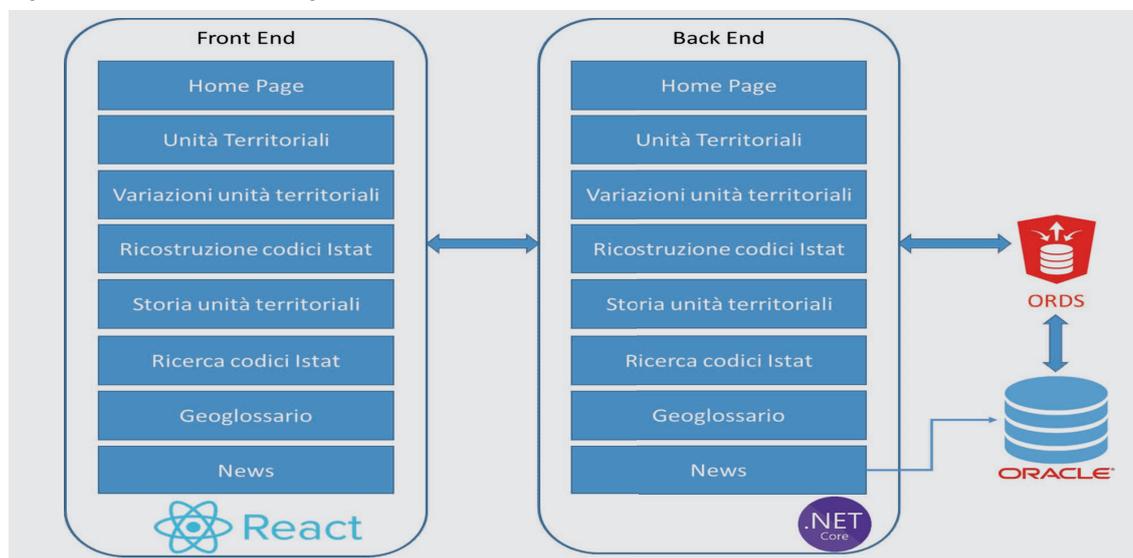
SITUAS - Report è il sito per la gestione (creazione, modifica e cancellazione) della reportistica. Tale modulo applicativo è realizzato per consentire agli utenti Istat abilitati, in funzione del ruolo e delle autorizzazioni assegnati, di profilare dati, metadati e modalità di consultazione dei contenuti e dei campi per tipologia di utente. Contestualmente alla definizione di un nuovo *report* il software genera in automatico l'indirizzo *https* per l'invocazione dell'*API Rest* per rendere fruibili a terzi quanto profilato. Per *SITUAS - Report* è prevista la possibilità di abilitare l'autenticazione mediante Spid e CIE.

SITUAS - Servizi è un modulo realizzato in *Oracle APEX*, mediante *API Rest* consente la creazione e la pubblicazione dei *report* e ne controlla gli accessi. Per la verifica dell'autorizzazione all'invocazione di un microservizio sono state implementate due soluzioni alternative: la prima prevede l'inserimento dell'applicativo *client* e del rispettivo *server* in una *whitelist* (utenti interni Istat); la seconda implementa un meccanismo di autenticazione a due fattori, che prevede lo scambio di credenziali con l'applicazione *client* al fine di generare un *token* da scambiare nel corso delle richieste all'interno della sezione (per utenti esterni).

4.6.2 Il modulo SITUAS per gestione delle unità territoriali e le classificazioni

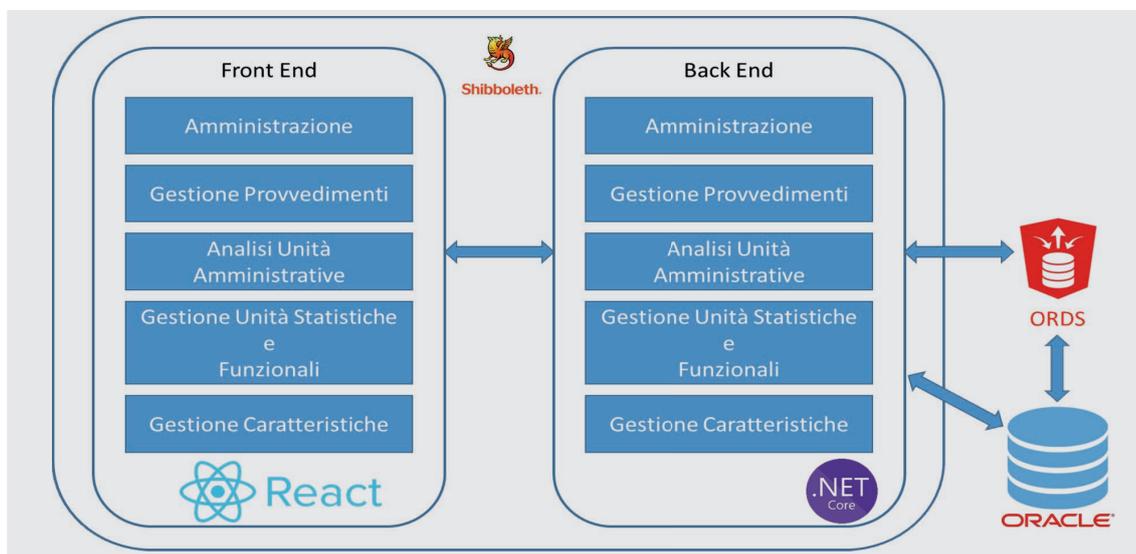
AGSituas è un'applicazione interna per la gestione dei dati. Il sistema consente agli utenti identificati quali responsabili dei dati di gestire l'inserimento e l'aggiornamento: dei provvedimenti amministrativi con impatto sulle unità amministrative; delle relazioni tra i livelli territoriali (Ua); dei caratteri delle unità amministrative e delle unità statistiche; della demografia e della composizione delle unità statistiche; delle relazioni tra livelli territoriali (Us). È prevista la gestione dei paesi esteri che verrà rilasciata in fasi successive. Il sistema consente la gestione degli utenti ai quali deve essere assegnato un ruolo collegato a specifiche funzionalità. Ogni utente dovrà poi autenticarsi all'applicativo attraverso il *Single Sign On* adottato da Istat.

Figura 4.3 - Rappresentazione grafica di *SITUAS - Diffusione*



Fonte: Istat



Figura 4.4 - Rappresentazione grafica di *SITUAS - Gestionale*

Fonte: Istat

5. BASI TERRITORIALI E MICROZONIZZAZIONE¹

5.1 Caratteristiche principali del Progetto Sezioni e Microzone 2021 e finalità

Uno dei pilastri del Registro statistico di base dei luoghi (RSBL) sono le Basi territoriali 2021 (BT2021), aggiornate dal progetto “Sezioni e Microzone 2021”². Le BT 2021 sono una infrastruttura poligonale costruita per contenere e diffondere le informazioni provenienti dal Censimento permanente della Popolazione e delle abitazioni e, più in generale, dal Sistema Integrato dei Registri. In questo Capitolo sarà descritto il flusso di lavoro che ha generato le BT2021.

Gli adempimenti finalizzati alla realizzazione delle basi territoriali, sono previsti dall’articolo 9 della legge sull’ordinamento delle anagrafi della popolazione residente (legge 24 dicembre 1954, n. 1228) e dagli articoli 38, 39 e 40 del Capo VII del nuovo regolamento anagrafico della popolazione residente emanato con Decreto del Presidente della Repubblica 30 maggio 1989, n. 223 (Istat 1992).

Le attività di aggiornamento delle BT 2021 sono state organizzate secondo le disposizioni contenute nel Piano generale del Censimento approvato in data 8 ottobre 2020³.

Gli oggetti delle BT sono, in ordine gerarchico decrescente: i limiti comunali; le località abitate distinte in centri, nuclei abitati; le località produttive; il resto del territorio comunale denominato case sparse; le sezioni di Censimento.

L’obiettivo delle BT è di definire una suddivisione del territorio aderente alla realtà. Il cambiamento della strategia censuaria ha modificato sostanzialmente la funzione assegnata alla base territoriale comunale. Infatti, le BT in passato venivano utilizzate principalmente per la raccolta dei dati delle unità di rilevazione; in occasione del Censimento permanente, e con il ricorso al campionamento delle unità, le sezioni di Censimento sono, invece, impiegate solo parzialmente per la raccolta dei dati e non esaustivamente per tutti i comuni italiani.

Questa nuova edizione delle basi territoriali è finalizzata alla diffusione dei dati al minimo livello territoriale. Le sezioni di Censimento, infatti, continuano a distinguere in modo esclusivo la fase di diffusione dei dati statistici, pure preservando il vincolo di riservatezza delle informazioni statistiche sul territorio.

È bene premettere che la copertura geografica delle sezioni 2021 eredita regole e oggetti geometrici delle Basi territoriali del 2011 ma, per gli scopi della diffusione dei dati subcomunali, è stata significativamente migliorata la qualità del disegno e aumentata l’o-

1 Il Capitolo è stato redatto da Fabio Lipizzi e Alberto Sabbi (paragrafi 5.1 e 5.2), Marina Arcasenza (paragrafi 5.3, 5.3.1 e 5.3.2), Germana Endennani (paragrafi 5.3.3 e 5.3.4) e Assunta Sera (paragrafo 5.4).

2 Nella nuova base territoriale, il livello primario è quello delle “microzone”. Questa base è stata costruita enucleando parti di territorio omogenee per uso e copertura del suolo all’interno delle sezioni di Censimento. Allo scopo di contenere l’onere operativo che i comuni hanno affrontato nell’attività di validazione delle basi territoriali, l’informazione geografica è stata generalizzata semplificando la base territoriale e limitando il numero di sezioni da revisionare. In questo contesto si parlerà solo di sezioni 2021, indicando le microzone come un ulteriore elemento di diffusione successivo alle attività connesse alle sezioni di Censimento.

3 Cfr. www.istat.it/it/censimenti/popolazione-e-abitazioni/documentazione e <https://www.istat.it/files//2022/03/Piano-Generale-di-Censimento.pdf>.

mogeneità interna delle sezioni. Peraltro, già in occasione del Censimento del 1981 si profilava l'idea di considerare le basi territoriali non solo ai fini puramente strumentali delle operazioni censuarie e si sottolineava l'esigenza di favorire le loro "caratteristiche omogenee sotto i profili demografico, socioeconomico, urbanistico, ambientale e simili". Questa affermazione è un primo importante principio che rafforza la possibilità di un utilizzo analitico della suddivisione del territorio collegata al dato statistico censuario. La caratteristica di omogeneità delle basi territoriali è il presupposto per l'utilizzo di questo enorme patrimonio informativo, unico nel panorama nazionale, in una ottica mirata a una più ampia conoscenza del territorio subcomunale.

Sul piano tecnologico le innovazioni recepite nel progetto hanno permesso alcune modifiche strutturali di processo e di prodotto rispetto al 2011, attuando una notevole semplificazione delle singole attività. Le innovazioni tecnologiche introdotte hanno consentito un sensibile aumento di produttività nella formulazione della proposta delle BT, sia nella produzione degli elaborati cartografici sia nelle fasi successive del processo di acquisizione dei dati geografici.

L'insieme di questi elementi ha consentito di progettare e realizzare l'intero processo di aggiornamento delle basi territoriali.

5.2 Regole generali

Il processo di generazione della nuova base territoriale è stato elaborato aggiungendo alle sezioni di Censimento del 2011 ulteriori fonti cartografiche, che sono state sovrapposte geograficamente tra loro, ricontrollate e validate geometricamente. Lo scopo è di caratterizzare ulteriormente le sezioni di Censimento in corrispondenza di un disegno più preciso. Il territorio antropizzato è stato oggetto di una operazione di fotointerpretazione, verificando l'espansione delle località del 2011 (centri e nuclei abitati e località produttive) sulla base delle foto aeree (ortofoto) di recente acquisizione (Figura 5.1).

Figura 5.1 - Un esempio del disegno delle nuove sezioni 2021



Fonte: Elaborazioni su dati Istat

La regola generale per il disegno delle nuove sezioni 2021 prevede che siano delimitati nuovi poligoni solo all'interno di una sezione di Censimento preesistente nel 2011.

Ciascuna nuova sezione di Censimento è disegnata seguendo un criterio di omogeneità interna. A ognuna di questa viene assegnato un codice di tipo speciale (*Cod_Tipo_S*) che caratterizza l'area in termini di uso e copertura del suolo, come si dirà anche nel Capitolo successivo. Per limitare l'eccessiva proliferazione sul territorio delle nuove sezioni di Censimento, le aree omogenee sono definite dal numero minimo di sezioni. Generalmente queste operazioni sono state effettuate senza influenzare la corrispondenza con le sezioni residenziali del 2011, ma piuttosto enucleando la parte funzionale del territorio.

5.3 Flusso di lavoro per l'aggiornamento delle Basi territoriali 2021

Le fasi di lavoro per l'aggiornamento delle Basi territoriali 2021, sono simili a quelle già realizzate nelle precedenti edizioni delle BT:

1. definizione della proposta e pubblicazione dello strato geografico delle sezioni 2021;
2. validazione comunale della proposta;
3. interlocuzione con le amministrazioni comunali e revisione delle proposte comunali nel rispetto degli standard cartografici individuati dall'Istat;
4. controllo, correzione e validazione dei dati geografici.

Tuttavia, il processo di produzione del dato all'interno delle singole fasi di lavoro è stato sostanzialmente modificato.

5.3.1 Definizione della proposta dello strato geografico delle sezioni 2021

È stata formulata dall'Istat una proposta di aggiornamento delle BT2021. La proposta, già sperimentata anche in passato, consente una riduzione dell'onere operativo delle amministrazioni comunali nelle attività di controllo delle basi stesse. Queste hanno il solo compito di revisionare e convalidare la proposta. Se necessario, possono modificare o integrare le Basi territoriali inviate. Tale operazione consente di ottenere un disegno realizzato con le stesse regole e omogeneo per tutto il Paese.

Nel periodo intercorso dalla precedente definizione delle BT2011, i progressi tecnologici (Laaribi *et al.* 2019), nell'ambito dei sistemi informativi geografici, si sono rivelati estremamente importanti sia sul piano della gestione, sia sul lato dell'aggiornamento dei dati geografici. L'esperienza maturata, supportata anche da quella internazionale (UNECE 2021*b*; UN 2009), ha consentito di estendere la formazione tecnica, altresì, al personale delle sedi territoriali Istat che ha partecipato attivamente alla realizzazione di questa fase di lavoro.

In particolare le sedi territoriali Istat (UU.TT.), che hanno collaborato al progetto sono: Marche-Abruzzo e Puglia (RFT), Sicilia e Sardegna (RTI), Lazio-Calabria e Molise (RTG)⁴.

La definizione della proposta ha riguardato, principalmente:

- **l'espansione delle località abitate:** si è proceduto al disegno di nuove sezioni se sulle ortofoto sono stati rilevati almeno sette nuovi edifici o sono state ampliate le sezioni, quando gli edifici erano in numero inferiore (Figura 5.2a. La nuova sezione con

⁴ Va qui ricordato che la nuova organizzazione dell'Istat è entrata in vigore prima dell'avvio del progetto "sezioni e microzone 2021".

- tratteggio rosso si trova a sud della località simboleggiata in arancio);
- **la fusione di località abitate 2011**: a causa dell'espansione urbanistica si è verificato spesso il caso nel quale le distanze tra località si sono ridotte al di sotto della soglia prevista (30 metri per i nuclei e 70 per i centri abitati). Ciò ha portato alla fusione di alcune località in una sola (Figura 5.2b);
 - **le nuove località**: sono state individuate in presenza di almeno 15 nuovi edifici sull'ortofoto, distanti meno di 30 metri dalla casa più vicina del più prossimo centro o nucleo abitato (Figura 5.2c);
 - **la correzione del disegno delle località abitate**: sono state riclassificate o ridisegnate le località del 2011 qualora si mostrassero prive di insediamenti abitativi, escludendo aree classificate "abitate" e/o provvedendo a una loro frammentazione (Figura 5.2d).

Figura 5.2a - Espansione di una località



Figura 5.2b - Fusione di due località



Figura 5.2c - Nuova località



Figura 5.2d - Frammentazione di località



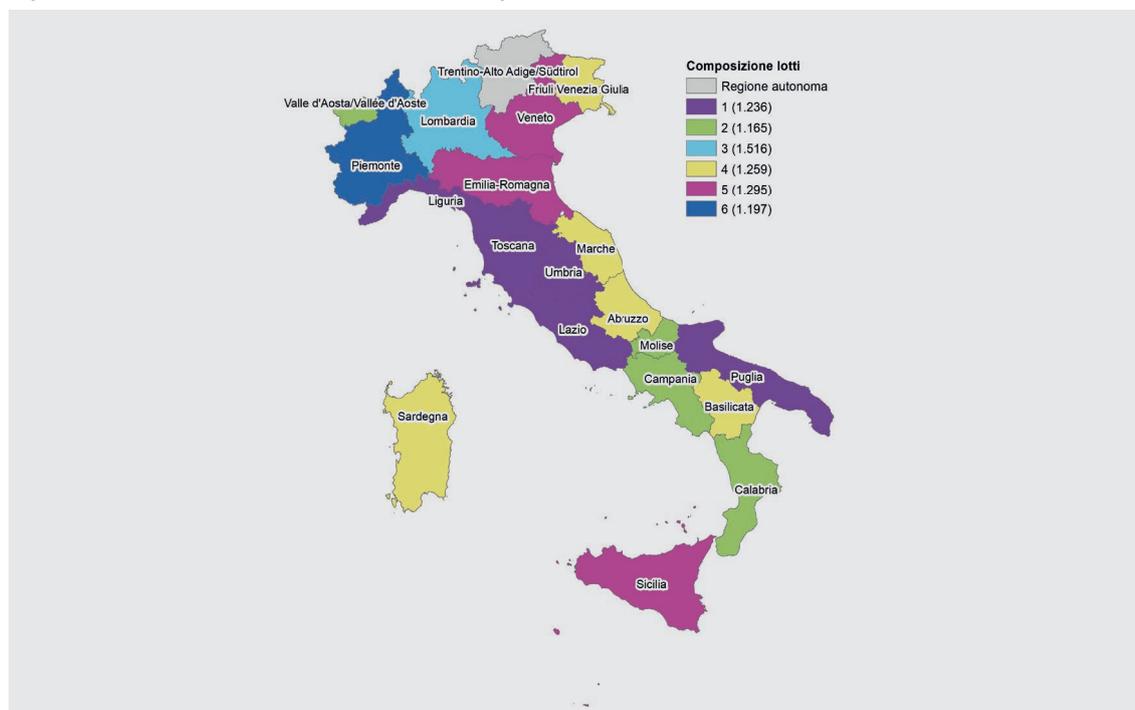
5. Basi territoriali e microzonizzazione

5.3.2 Indagine censuaria per la validazione delle BT2021

La fase 2, è stata realizzata dai comuni. Le amministrazioni comunali sono anche gli enti depositari delle BT, come indicato nel regolamento anagrafico e sono gli unici in grado di avere una conoscenza puntuale del territorio da loro governato.

L'avvio delle attività di validazione delle nuove sezioni di Censimento con i comuni italiani è iniziata con l'invio di una prima lettera a maggio del 2021 a tutti, con l'esclusione di quelli del Trentino-Alto Adige/*Südtirol*. Questa esclusione è dovuta al fatto che le Province autonome di Trento e Bolzano/*Bozen* hanno proceduto autonomamente all'aggiornamento delle BT e alla conduzione dell'indagine. La lettera aveva come oggetto "Adempimenti dei comuni relativi all'aggiornamento delle Basi territoriali, progetto sezioni di Censimento 2021" e invitava i comuni a prepararsi per le specifiche attività di revisione delle BT.

Figura 5.3 - Composizione dei lotti di lavoro per regione



Fonte: Elaborazioni su dati Istat

Ai fini della gestione della lavorazione si è ritenuto utile suddividere in otto lotti le regioni italiane. I lotti sono stati definiti cercando di rispettare, per quanto possibile, una suddivisione omogenea del territorio in termini di superficie e numero di comuni (Figura 5.3). Sono stati individuati anche due ulteriori lotti: il primo per le due Province autonome (lotto 8); il secondo (lotto 7), relativo ai 14 comuni capoluogo delle Città metropolitane (Milano, Torino, Venezia, Genova, Bologna, Firenze, Roma, Napoli, Bari, Reggio di Calabria, Messina, Catania, Palermo e Cagliari).

La restituzione dei dati era prevista con tempi diversi, secondo il numero di sezioni 2011 presenti nel territorio comunale.

Successivamente è stata inviata una seconda lettera circolare avente per oggetto "Circolare adempimenti dei comuni relativi all'aggiornamento delle Basi territoriali, progetto Microzone e sezioni di Censimento 2021".

Insieme alla descrizione della documentazione invitata e alle istruzioni operative per la restituzione delle nuove basi validate, la circolare conteneva anche due allegati diversi per ogni lotto⁵.

Dopo circa un anno di attività il sito per la validazione delle Basi territoriali è stato chiuso definitivamente il 18 luglio del 2022.

5.3.3 Risultati dell'indagine e validazione della proposta Istat delle BT 2021

L'interscambio informativo con i comuni e la gestione delle fasi operative della validazione sono state gestite attraverso un portale dedicato, accessibile dai responsabili comunali.

Tramite questo sono stati condivisi i dati geografici e non, per tutti i comuni nel formato vettoriale (*shapefile*) e *kmz*, nel più diffuso *pdf* e in quello tabellare. È stata rilasciata anche una grande quantità di materiale formativo e informativo, tra l'Istat e i responsabili comunali. Anche in questa fase l'apporto delle UU.TT. Istat si è rivelato decisivo per gestire le operazioni di validazione con i singoli comuni.

Per avere un quadro completo dell'andamento dell'indagine è necessario ricordare che non tutte le regioni sono state seguite direttamente dal personale delle UU.TT. Dal 2 febbraio 2022 è subentrata, nella conduzione dell'indagine, la direzione della raccolta dati.

L'intervento della Direzione centrale della raccolta dati, ha consentito anche per queste regioni, un invio massivo di solleciti oltre le scadenze indicate nella circolare riportate in appendice. Questa attività ha incrementato le risposte dei comuni di circa il 20 per cento rispetto agli arrivi previsti inizialmente.

Alla chiusura dell'indagine, prorogata fino al 18 luglio 2022, hanno risposto 6.698 comuni, ossia circa l'85 per cento (Tavola 5.1) e tra questi tutti i 14 grandi comuni. In nove regioni hanno risposto più del 90 per cento dei comuni, in sette regioni l'80 per cento e in quattro regioni il 70 per cento. Chiudono le ultime due regioni, Lazio e Calabria, rispettivamente con il 68,4 e il 48,9 per cento dei comuni che hanno risposto all'indagine. In termini di superficie tali dati corrispondono a circa l'85 per cento della superficie italiana e i valori rilevati, a livello regionale, non si discostano molto da quelli appena descritti. La sola eccezione riguarda la Valle d'Aosta/*Vallée d'Aoste* che dall'82,4 per cento passa, in termini di superficie, al 78 per cento. Considerando la popolazione, invece, si nota che nei comuni aderenti all'indagine, la corrispondente popolazione è prossima al 90 per cento di quella italiana: comuni di maggiore dimensione demografica sono anche quelli che hanno risposto più frequentemente. In 12 regioni e nelle due Province autonome i valori superano l'88 per cento della corrispondente popolazione, in quattro il 70 per cento e in due il 50 per cento. Tra queste la Calabria è in coda con il 53,3 per cento della popolazione regionale.

I comuni con dimensioni demografiche medio grandi hanno spesso lavorato con gli *shapefile* distribuiti sul sito di scambio. Questo ha semplificato il lavoro svolto "centralmente" dall'Istat, sia nella fase di verifica delle BT, potendo usufruire di un dato elaborabile digitalmente, sia nella successiva fase di integrazione delle informazioni pervenute.

⁵ Il primo Allegato conteneva l'elenco dei comuni per lotto di appartenenza e le relative date di restituzione del materiale cartografico; il secondo conteneva l'elenco dei referenti delle sedi territoriali e della sede centrale dell'Istat.

5. Basi territoriali e microzonizzazione

Tavola 5.1 - Numero di comuni che hanno risposto all'indagine al 18 luglio 2022, relativa superficie (in km²) e popolazione anagrafica al 31/12/2020 per regione (valori assoluti e percentuali)

Regioni	Numero di comuni			Superficie (in km ²)			Popolazione anagrafica al 31/12/2020		
	VA	%	Totale	VA	%	Totale	VA	%	Totale
Piemonte	1.137	96,4	1.180	24.446,4	96,8	25.256,6	3.321.895	97,2	3.416.740
Valle d'Aosta/ <i>Vallée d'Aoste</i>	61	82,4	74	2.544,2	78,0	3.260,9	107.850	86,9	124.089
Liguria	224	96,1	233	4.967,7	96,0	5.175,9	849.582	89,2	952.085
Lombardia	1.301	86,4	1.505	21.087,6	89,0	23.681,4	7.577.080	88,0	8.606.972
Trentino- Alto Adige/ <i>Südtirol</i>	261	92,6	282	12.692,3	93,3	13.604,8	1.012.648	94,0	1.077.078
Bolzano/ <i>Bozen</i>	116	100,0	116	7.397,9	100,0	7.397,9	534.912	100,0	534.912
Trento	145	87,3	166	5.294,4	85,3	6.206,9	477.736	88,1	542.166
Veneto	519	92,3	562	16.327,9	91,1	17.929,5	4.331.633	93,9	4.613.747
Friuli-Venezia Giulia	190	88,4	215	7.076,3	89,2	7.932,5	1.084.288	90,2	1.201.510
Emilia-Romagna	296	90,5	327	20.290,9	91,0	22.303,7	3.855.768	95,3	4.047.251
Toscana	270	99,3	272	22.841,2	99,8	22.885,1	3.309.245	99,5	3.324.446
Umbria	89	96,7	92	7.925,8	93,6	8.464,2	857.158	99,0	865.452
Marche	224	98,7	227	9.265,0	98,6	9.401,2	1.486.880	99,2	1.498.236
Lazio	258	68,4	377	10.772,3	67,6	15.944,5	2.143.383	72,4	2.960.173
Abruzzo	251	82,3	305	9.290,9	85,8	10.831,5	1.131.267	88,3	1.281.012
Molise	100	73,5	136	3.277,3	73,5	4.460,4	192.880	65,5	294.294
Campania	410	74,7	549	10.168,3	75,0	13.551,7	3.608.943	76,8	4.702.166
Puglia	184	71,9	256	13.817,1	71,1	19.423,1	2.722.696	75,3	3.616.572
Basilicata	119	90,8	131	9.434,6	93,7	10.073,1	513.290	94,2	545.130
Calabria	197	48,9	403	6.940,2	46,3	14.982,6	898.675	53,3	1.687.575
Sicilia	275	71,1	387	19.064,3	75,4	25.275,4	2.909.167	79,2	3.673.135
Sardegna	318	84,6	376	20.461,6	85,2	24.014,9	1.251.741	86,9	1.440.472
Grandi comuni	14	100,0	14	3.615,5	100,0	3.615,5	9.308.078	100,0	9.308.078
ITALIA	6.698	84,8	7.903	256.307,1	84,9	302.068,7	52.474.147	88,6	59.236.213

Fonte: Elaborazioni su dati Istat

I 14 comuni capoluogo delle Città metropolitane hanno ricevuto la proposta delle Basi territoriali solo in formato vettoriale, allo scopo di evitare il proliferare degli elaborati cartografici. Questi comuni, infatti, generalmente possiedono sistemi informativi geografici realizzati con finalità più generali di governo della città. Nelle Province autonome di Trento e Bolzano/*Bozen* gli uffici di statistica hanno sostituito l'Istat procedendo autonomamente all'aggiornamento delle BT con i comuni delle province⁶.

L'adozione di una strategia modulare indirizzata a diverse unità di rilevazione, in parte centralizzata e in parte decentralizzata, ha consentito di massimizzare i vantaggi, in termini di tempestività e qualità dei risultati, e ridurre gli svantaggi dell'intero processo.

I comuni che sono intervenuti sul file *pdf* della proposta Istat hanno inserito delle note, utilizzando un software *open source*, fornito da Istat (Figura 5.4).

Tutte le operazioni di trasmissione ed elaborazione sono avvenute utilizzando esclusivamente supporti digitali. Questo ha consentito di ridurre notevolmente i tempi di spedizione/ricezione dei materiali cartografici, quelli di elaborazione/disegno e acquisizione dei dati, nonché gli ingenti costi dovuti alla stampa dell'elaborato cartografico in un formato A0.

⁶ Ai comuni della Provincia autonoma di Bolzano/*Bozen* è stata inviata una lettera circolare bilingue (Prot. n. 61996/22 del 13/01/2022).



Figura 5.4 - Elaborato cartografico in formato pdf oggetto di revisione da parte del comune



Fonte: Elaborazioni su dati Istat

5.3.4 Strumenti GIS per l'aggiornamento delle BT2021

Le operazioni di aggiornamento hanno previsto varie fasi di *editing* geografico per la produzione delle BT2021 (proposta ai comuni, verifica e acquisizione del materiale restituito all'Istat, validazione finale), pertanto è stato necessario ingegnerizzare le varie fasi del processo e implementare strumenti personalizzati per la produzione.

Per semplificare le operazioni di acquisizione dei dati geografici è stato standardizzato l'ambiente di lavoro su progetti con estensione regionale⁷. Nella fase di acquisizione dei dati geografici sono state inserite alcune regole per la localizzazione degli errori topologici⁸, personalizzando la barra degli strumenti⁹ del software GIS, utilizzato per l'aggiornamento delle BT 2021, *ArcGIS* della Esri. Ciò ha consentito di mantenere l'integrità delle relazioni spaziali correggendo eventuali *gap* (poligoni non adiacenti) e *overlap* (poligoni sovrapposti).

Generalizzando le fasi di una indagine statistica, i pulsanti di ausilio all'*editing* topologico (Figura 5.5) possono configurarsi come uno strumento di imputazione assistita dove, invece di inserire dati testuali o numerici di un questionario, sono acquisiti o rettificati i dati geografici in funzione delle modifiche indicate dai comuni. Dal punto di vista geografico si presenta una digitalizzazione a video (*head up*) di oggetti vettoriali. Si ricorda che le operazioni di *editing* di un dato vettoriale non sono paragonabili in complessità a quelle di un dato alfanumerico, perché quest'ultimo non risponde a regole geometriche e topologiche. Per questo è stato necessario inserire un passaggio di controllo e correzione nel processo di acquisizione dei dati geografici.

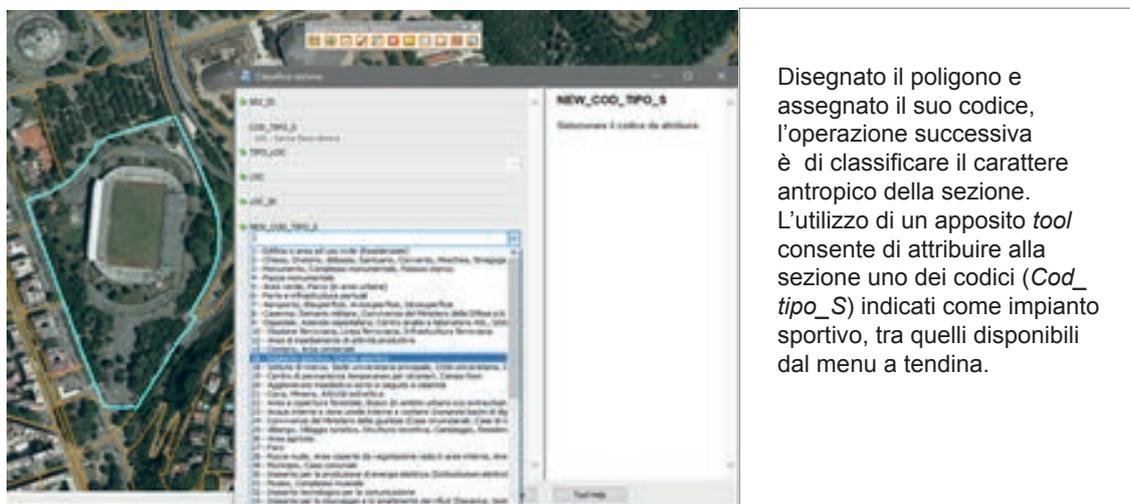
A tale scopo sono state sviluppate due ulteriori *toolbar*, per assicurare l'accuratezza tematica dei dati di tipo deterministico. In particolare per i controlli di coerenza, per quelli dei *range*, per la validità degli attributi relativi al dato geografico e la coerenza spaziale

⁷ I dati geografici nazionali risiedono su un *geodatabase Oracle Enterprise*, ma per l'aggiornamento si è operato su un sottoinsieme di dati estratti a livello regionale, in formato file *geodatabase*.

⁸ Si veda a tale proposito una generalizzazione della topologia applicata ai GIS cfr. *Egenhofer et al.* 1991.

⁹ In particolare sono state sviluppate delle *toolbar* personalizzate in linguaggio *Python* per *ArcGIS*.

Figura 5.5 - *Toolbar* aggiornamento sezioni



Fonte: Elaborazioni su dati Istat

delle località contigue¹⁰. Gli algoritmi di localizzazione degli errori implementati, si sono rivelati particolarmente efficienti per individuare le anomalie inserite dall'operatore nella fase di acquisizione dei dati.

Le procedure di controllo sono di tipo "micro", dove tutte le osservazioni sono sottoposte a verifica. In caso di errori rilevati, la procedura crea una reportistica per ogni *tool* eseguito che l'operatore dovrà esaminare per procedere alla eventuale correzione.

È necessario qui ricordare che la peculiarità del dato geografico non consente l'utilizzo di procedure probabilistiche per la correzione degli errori, se non in alcuni casi del tutto particolari. Ciò è strettamente connesso alla inevitabile necessità di una visualizzazione completa e dettagliata delle informazioni rilevate sul territorio. Così, ad esempio, una sezione disegnata lungo un'asta fluviale può avere il solo attributo di "acqua" con una probabilità certa e non può essere confusa in altro modo. Da qui nasce la naturale necessità di un utilizzo parsimonioso di procedure probabilistiche per i dati rilevati al massimo dettaglio territoriale, cui non si è fatto ricorso in questo lavoro.

5.4 Principali risultati

Le BT, quindi, sono oggetti geografici che descrivono il mosaico insediativo dell'Italia e la sua evoluzione nel medio e lungo periodo, sebbene presentino qualche limite causato dalle finalità censuarie.

Nella Tavola 5.2 sono riportati i dati definitivi delle sezioni di Censimento 2021 e 2011, raggruppati per regione. I dati del 2021 sono valorizzati prima e dopo il processo di validazione comunale.

Come si nota dalla tavola, questa operazione ha incrementato notevolmente il numero dei poligoni rispetto a quello delle sezioni di Censimento 2011. In estrema sintesi, le sezioni di Censimento 2021 sono 756.259 (erano 402.677 nel 2011) e sono aumentate di circa l'87 per cento. Complessivamente nelle località abitate e produttive le sezioni 2021

¹⁰ Due o più località contigue, tra comuni diversi, possono essere solo dello stesso tipo (centro-centro, nucleo-nucleo, località produttiva-località produttiva).

sono 527.569 e in termini percentuali, rispetto al 2011, l'aumento fa segnare un +67,3 per cento mentre, nelle case sparse, l'aumento percentuale è ancora più sostenuto e pari a +87,7 per cento, passando da circa 87 mila sezioni del 2011 a circa 228 mila del 2021. Valori che quantificano il lavoro di editing svolto e il maggiore dettaglio raggiunto nell'informazione cartografica. Nel paragrafo successivo si descriveranno più dettagliatamente alcune delle operazioni effettuate per la definizione delle BT2021. A livello di ripartizione geografica, le variazioni percentuali più elevate tra le sezioni definitive 2011 e 2021 sono quelle che ricadono territorialmente nel Nord-ovest, con valori superiori a 100 (105,5 per cento). I valori più contenuti delle cinque ripartizioni, invece, sono quelli che ricadono nella ripartizione Isole.

Da una analisi a livello regionale si nota un aumento delle variazioni percentuali 2011-2021, superiore al 100 per cento su nove delle 20 regioni italiane, mentre per le altre la differenza percentuale è sempre superiore al 50 per cento.

Le regioni che presentano le principali variazioni percentuali sono: la Valle d'Aosta/*Vallée d'Aoste* (135,8 per cento), il Piemonte (138,4 per cento) e il Molise (140,6 per cento). Viceversa i territori che mostrano delle differenze percentuali 2011-2021 più basse si trovano in Puglia (53,6 per cento), Emilia-Romagna (55,7 per cento) e Sicilia (55,0 per cento).

Considerando le località abitate, risulta un incremento di 60 centri abitati (0,3 per cento), 549 nuclei abitati (1,6 per cento) e 45 località produttive (1,6 per cento) rispetto al 2011.

Tavola 5.2 - Numero di comuni e di sezioni di Censimento per regione. Anni 2011 e 2021 (valori assoluti e variazione %)

Regioni Ripartizioni geografiche	Numero di comuni 2021	Sezioni 2011	Sezioni 2021 pre-validazione	Sezioni 2021 definitive	Variazione % 2011-2021 (a)
Piemonte	1.181	35.671	84.843	85.039	138,4
Valle d'Aosta/ <i>Vallée d'Aoste</i>	74	1.902	4.509	4.485	135,8
Liguria	234	11.054	19.798	19.793	79,1
Lombardia	1.506	53.173	96.093	99.886	87,9
Trentino-Alto Adige/ <i>Südtirol</i>	282	11.712	24.832	26.356	125,0
Bolzano/ <i>Bozen</i>	116	5.631	11.592	12.660	124,8
Trento	166	6.081	13.240	13.696	125,2
Veneto	563	33.883	68.265	68.807	103,1
Friuli-Venezia Giulia	215	8.278	17.202	18.960	129,0
Emilia-Romagna	328	38.603	57.702	60.112	55,7
Toscana	273	28.917	46.364	47.712	65,0
Umbria	92	7.480	12.560	13.318	78,0
Marche	227	11.862	20.970	21.113	78,0
Lazio	378	32.065	57.846	59.017	84,1
Abruzzo	305	9.529	18.585	19.342	103,0
Molise	136	2.821	6.645	6.787	140,6
Campania	550	24.323	48.562	49.716	104,4
Puglia	257	22.514	33.770	34.592	53,6
Basilicata	131	5.107	9.755	9.845	92,8
Calabria	404	13.121	28.914	29.875	127,7
Sicilia	390	36.681	56.568	56.869	55,0
Sardegna	377	13.981	24.437	24.635	76,2
Nord-ovest	2.995	101.800	205.243	209.203	105,5
Nord-est	1.388	92.476	168.001	174.235	88,4
Centro	970	80.324	137.740	141.160	75,7
Sud	1.783	77.415	146.231	150.157	94,0
Isole	767	50.662	81.005	81.504	60,9
ITALIA	7.903	402.677	738.220	756.259	87,8

Fonte: Elaborazioni su dati Istat

(a) Variazione % 2011-2021 delle sezioni di Censimento definitive.

5. Basi territoriali e microzonizzazione

Altri risultati sono visibili nella Tavola 5.3, dove è possibile apprezzare l'estensione delle località in km² e la loro variazione nel tempo. Complessivamente il territorio antropizzato delle località italiane è aumentato di +3,6 per cento. In termini assoluti, la variazione 2011/2021 è pari a +738,2 km². Se consideriamo che i primi due comuni per estensione territoriale sono Roma e Ravenna con superficie pari rispettivamente a 1.286,7 km² e 651,8 km², l'aumento delle superficie antropizzata nel corso del decennio corrisponde a quello della seconda città italiana in termini di estensione. A livello regionale si evidenziano variazioni della superficie delle località superiori al 5 per cento in Molise (13,1 per cento), Campania (7,4 per cento), Lazio (6,6 per cento), Valle d'Aosta/*Vallée d'Aoste* (6,1 per cento) e Piemonte (5,1 per cento).

Nonostante questo aumento, su tutto il territorio italiano vi sono alcuni casi particolari. Il caso più evidente è rappresentato dalla regione Trentino-Alto Adige/*Südtirol* dove il numero dei centri abitati passa da 839 del 2011 a 800 nel 2021, facendo registrare una variazione percentuale negativa pari a -4,6 per cento. Ciò è dovuto a una maggiore precisione nel disegno, all'effetto di conurbazione urbana con la conseguente fusione di località contigue e alla diminuzione delle piccole località montane.

Analizzando i dati per tipo di località (Tavola 5.3) si osserva, a livello nazionale, un incremento del 2,5 per cento della superficie (e in valore assoluto un +446,9 km²) per la tipologia centro abitato, dell'10,6 per cento (+183,8 km²) per i nuclei abitati e dell'11,4 per cento (+107,5 km²) per le località produttive rispetto alle BT del 2011.

A seguito della conferma post consolidamento delle località abitate, Umbria (-0,9 per cento) e Calabria (-4,6 per cento), in controtendenza con il dato nazionale, hanno subito una riduzione areale rispetto al 2011. Anche in questo caso si può ipotizzare che, tale dato derivi da un disegno più preciso dei contorni dei centri abitati stessi. Da una prima analisi delle variazioni di queste regioni, infatti, si evince come il lavoro apportato su ogni singola località abbia generato una contrazione sia nel numero delle località abitate sia nella superficie che occupano.

La geografia delle sezioni di Censimento 2021 è stata pubblicata sul sito Istat¹¹ e, nella loro versione definitiva, rappresenta un prezioso patrimonio informativo che descrive minuziosamente la situazione territoriale italiana. Insieme ai dati geografici l'Istat ha pubblicato anche le variabili definitive del Censimento permanente della Popolazione e delle abitazioni 2021 associabili alle partizioni del sistema delle Basi territoriali.

Gli oggetti contenuti nelle Basi territoriali, tra i quali le località abitate Istat, saranno così non soltanto funzionali alle operazioni censuarie, ma anche utilizzabili come elementi analitici per studi specifici (Ferrara *et al.* 2017; Lipizzi *et al.* 2017) nonché per un loro utilizzo istituzionale¹²: supporto per gli amministratori locali, per definire piani di emergenza della protezioni civile (Mugnoli *et al.* 2016) per fronteggiare eventi calamitosi, per affrontare nuove sfide sul cambiamento climatico (Abbatini *et al.* 2024) e, più in generale, per una conoscenza dettagliata sulle caratteristiche peculiari del territorio italiano.

11 Cfr. <https://www.istat.it/notizia/basi-territoriali-e-variabili-censuarie/>.

12 A questo proposito si consideri l'audizione del 30 gennaio 2019 "Esame delle proposte di legge in materia di consumo di suolo" consultabile al link www.istat.it/it/archivio/226984, o l'utilizzo per i Collegi elettorali, e le "Audizioni sulla sicurezza e stato di degrado delle città e delle loro periferie: follow-up" del 5 luglio 2017. www.istat.it/it/archivio/202052.

Tavola 5.3 - Superficie delle località abitate e produttive per regione. Anni 2011 e 2021 (valori in km² e variazione %) (a)

Regioni	Superficie 2011				Superficie 2021				Variazione % (c)
	CA (b)	NA (b)	LP (b)	Totale	CA (b)	NA (b)	LP (b)	Totale	
Piemonte	1.442,3	191,4	80,8	1.714,6	1.496,1	210,2	95,4	1.801,7	5,1
Valle d'Aosta/ <i>Vallée d'Aoste</i>	46,7	8,8	2,2	57,7	49,2	9,1	2,9	61,2	6,1
Liguria	480,7	39,3	5,5	525,4	485,2	41,9	5,7	532,9	1,4
Lombardia	2.805,3	157,3	85,8	3.048,4	2.910,6	164,7	92,0	3.167,3	3,9
Trentino-Alto Adige/ <i>Südtirol</i>	299,5	36,7	15,4	351,7	308,7	32,1	16,8	357,6	1,7
Bolzano/ <i>Bozen</i>	119,7	19,1	6,6	145,4	124,0	15,7	7,9	147,6	1,5
Trento	179,8	17,6	8,8	206,3	184,7	16,4	8,9	210,0	1,8
Veneto	2.059,5	234,6	78,5	2.372,6	2.140,6	262,4	84,8	2.487,7	4,9
Friuli-Venezia Giulia	616,8	31,2	38,6	686,6	637,1	34,7	44,6	716,4	4,3
Emilia-Romagna	1.470,5	161,1	80,8	1.712,5	1.491,5	162,7	86,6	1.740,8	1,7
Toscana	1.148,8	96,9	37,7	1.283,3	1.156,7	99,8	43,9	1.300,5	1,3
Umbria	321,1	37,1	9,9	368,2	314,5	39,8	10,6	364,8	-0,9
Marche	407,5	65,8	35,3	508,6	409,4	66,9	42,0	518,3	1,9
Lazio	1.541,3	159,0	71,6	1.771,9	1.624,2	191,2	72,7	1.888,0	6,6
Abruzzo	411,9	59,7	40,1	511,7	416,0	67,5	48,3	531,7	3,9
Molise	79,8	21,2	12,4	113,4	90,1	26,0	12,2	128,3	13,1
Campania	1.134,2	85,3	71,0	1.290,5	1.193,8	119,3	73,4	1.386,5	7,4
Puglia	826,5	53,3	96,9	976,7	817,4	67,2	106,1	990,7	1,4
Basilicata	126,0	24,0	27,7	177,8	131,7	25,3	28,1	185,2	4,2
Calabria	680,0	91,5	30,0	801,4	641,6	92,6	30,1	764,3	-4,6
Sicilia	1.176,0	109,8	64,6	1.350,4	1.198,3	124,6	83,8	1.406,7	4,2
Sardegna	507,8	62,1	59,3	629,1	516,7	72,0	71,6	660,3	5,0
ITALIA	17.582,3	1.726,2	944,2	20.252,6	18.029,3	1.909,8	1.051,7	20.990,8	3,6

Fonte: Elaborazioni su dati Istat

(a) I valori sono arrotondati alla prima cifra decimale. Di conseguenza i totali possono discostarsi leggermente dalla somma delle rispettive componenti.

(b) CA= Centri abitati, NA= Nuclei abitati, LP= Località produttive.

(c) Variazione % 2011-2021 sul totale delle località regionali.

6. INTEGRAZIONE DATI GEOGRAFICI E ANALISI SPAZIALE PER LA REALIZZAZIONE DELLE NUOVE BASI TERRITORIALI¹

6.1 Uno strumento di analisi: le sezioni di Censimento 2021

Come si è potuto constatare nel Capitolo precedente, le Sezioni di Censimento 2021 (SdC)², quindi, rappresentano una sintesi cartografica di strati informativi “condensati” in un unico *layer* che permette analisi spaziali statistiche relative a varie tematiche legate al territorio. Per non incorrere in ridondanze in queste pagine, ci soffermeremo principalmente su alcune interessanti possibilità analitiche che permettono di estrarre, a partire dalle SdC 2021, statistiche legate alla distribuzione spaziale della popolazione e relative alla copertura del suolo.

Infatti, le Basi territoriali prodotte dall'Istat e in particolare le SdC sono sempre state oggetto di integrazione cartografica e analisi spaziale territoriale, come lo dimostrano le numerose pubblicazioni a proposito³.

Inoltre, la fitta trama territoriale disegnata dalle SdC 2021, permette anche una più puntuale quantificazione di tipologie di copertura del suolo a partire anche da immagini telerilevate, come verrà descritto successivamente in questo Capitolo.

6.2 Integrazione SdC 2021 e altri GeoDB

6.2.1 Integrazione con dati di copertura/uso del suolo

Una delle maggiori novità delle SdC 2021 rispetto al passato è il Codice tipologia speciale (*Cod_Tipo_S*) (Mugnoli *et al.* 2023), un attributo legato a ogni poligono che costituisce il *layer*, il quale ha una stretta connessione con la copertura/uso del suolo. Tale relazione, resa possibile tecnicamente da algoritmi GIS di analisi spaziale, è basata soprattutto dalla sovrapposizione tra le SdC 2021 e lo strato denominato *Refresh esteso* e prodotto da AGEA (Agenzia per le erogazioni in agricoltura)⁴.

Ciò ha portato a un notevole aumento dell'informazione territoriale soprattutto in ambito extraurbano. Infatti, proprio per la collocazione spaziale delle sezioni di case sparse, queste sono delimitate a una scala che non ne permette un'analisi statistica dettagliata come le altre tipologie di sezione (centro abitato, nucleo abitato, località produttiva).

¹ Il Capitolo è stato redatto da Alberto Sabbi e Juri Corradi (paragrafi 6.1 e 6.3), Alberto Sabbi (paragrafo 6.2.1), Stefano Mugnoli (paragrafo 6.2.2).

² Le sezioni di Censimento 2021 sono consultabili al link: <https://www.istat.it/it/archivio/280254>.

³ Cfr. Chiocchini *et al.* 2014 e Mugnoli *et al.* 2011.

⁴ Per le principali caratteristiche dello strato *Refresh*: https://agrea.regione.emilia-romagna.it/settori-di-intervento/aiuti/pagamenti-diretti/altra-documentazione/specifiche_refresh_11.pdf.



Quindi, mediante delle semplici tabelle di transcodifica, si è potuto classificare ogni poligono di sezione di Censimento secondo la legenda *land cover* utilizzata nella indagine LUCAS (*Land cover and use area survey*)⁵ mediante la quale Eurostat diffonde i dati ufficiali di copertura del suolo per tutti i paesi europei coperti dall'indagine con dettaglio regionale (livello Nuts2).

La legenda LUCAS prevede due differenti elenchi per la copertura e l'uso del suolo ed è organizzata secondo otto voci principali:

A00 (*Artificial*) – Aree di natura antropica

B00 (*Cropland*) – Aree agricole

C00 (*Woodland*) – Boschi e foreste

D00 (*Shrubland*) – Cespuglieti

E00 (*Grassland*) – Praterie naturali

F00 (*Bare lands and lichens/moss*) – Aree senza copertura vegetale licheni/muschi

G00 (*Water areas*) – Aree coperte da acqua

H00 (*Wetlands*) – Zone umide

Nella Tavola 6.1 è riportata la distribuzione percentuale della superficie delle sezioni di Censimento 2021 articolata per regione e per classe di copertura LUCAS.

Tavola 6.1 - Superficie regionale per classe di copertura del suolo LUCAS. Anno 2021 (composizioni percentuali)

Regioni	A00	B00	C00	E00, D00	F00	G00, H00	Totale
Piemonte	9,0	40,7	37,6	4,0	7,7	1,0	100,0
Valle d'Aosta/Vallée d'Aoste	2,2	5,0	31,1	1,8	50,7	9,2	100,0
Liguria	10,6	8,6	74,6	4,6	1,5	0,1	100,0
Lombardia	14,5	42,5	28,2	1,8	9,2	3,8	100,0
Trentino-Alto Adige/Südtirol	3,1	18,8	54,2	6,8	15,3	1,8	100,0
Bolzano/Bozen	2,3	28,2	48,9	1,6	17,0	2,0	100,0
Trento	3,8	7,7	60,5	12,9	13,4	1,7	100,0
Veneto	14,4	48,7	24,0	4,3	3,0	5,6	100,0
Friuli-Venezia Giulia	9,9	33,1	40,5	5,6	6,3	4,6	100,0
Emilia-Romagna	8,4	58,5	25,9	3,9	0,6	2,7	100,0
Toscana	6,4	39,5	51,7	0,7	1,2	0,5	100,0
Umbria	4,9	47,8	41,3	4,0	0,1	1,9	100,0
Marche	5,9	54,9	33,7	4,8	0,4	0,3	100,0
Lazio	11,8	47,1	29,1	8,2	2,2	1,6	100,0
Abruzzo	5,4	42,0	31,1	17,7	3,5	0,3	100,0
Molise	3,4	50,2	37,0	7,8	1,2	0,4	100,0
Campania	10,8	51,4	28,3	8,0	1,0	0,5	100,0
Puglia	5,9	78,8	5,4	7,7	1,0	1,2	100,0
Basilicata	2,7	55,1	33,3	7,2	0,5	1,2	100,0
Calabria	5,3	35,3	39,4	17,3	1,3	1,3	100,0
Sicilia	6,4	62,5	8,3	19	3,2	0,6	100,0
Sardegna	3,5	40,1	23,6	30	1,7	1,1	100,0
ITALIA	7,2	43,0	33,9	8,3	5,6	2,0	100,0

Fonte: Elaborazioni Istat su dati AGEA, Istat ed Eurostat

La regione con la maggiore percentuale di copertura artificiale, in percentuale rispetto alla propria superficie, è la Lombardia (14,5 per cento) mentre la Puglia spicca per l'alto valore associato alla classe dell'agricolo (78,8 per cento). Il Trentino-Alto Adige/Südtirol presenta una estesa copertura di boschi (54,2 per cento) mentre le altre categorie che rappresentano la vegetazione (cespuglieti e praterie) sono più presenti in Abruzzo (17,7 per cento). Un aspetto particolare riguarda gli alti valori della categoria acque e rocce nude in Valle d'Aosta/Vallée d'Aoste (rispettivamente 50,7 per cento e 9,2 per cento), mentre per

5 Cfr. <https://ec.europa.eu/eurostat/web/products-statistical-working-papers/-/ks-tc-22-005>.

6. Integrazione dati geografici e analisi spaziale per la realizzazione delle nuove Basi territoriali

le aree non coperte da vegetazione il dato è assolutamente veritiero. Nel caso delle acque bisogna ricordare che la classe dei ghiacciai è inserita nelle *water areas* in LUCAS.

6.2.2 Analisi spaziali e indici radiometrici di vegetazione

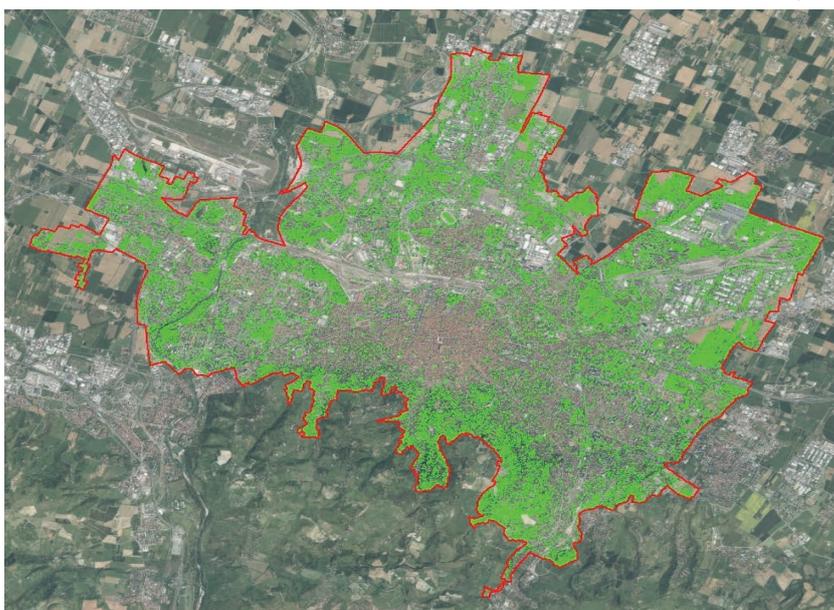
Conoscere anche approssimativamente la quantità di “verde” all’interno di un’area urbana è sicuramente un dato importantissimo per l’analisi e lo sviluppo di indicatori relativi a molti aspetti della vita nelle città. Fenomeni legati alla qualità della vita, in alcuni casi, sono direttamente proporzionali alla possibilità di fruizione delle aree verdi pubbliche e private (parchi, giardini, ville storiche, impianti sportivi, eccetera). Anche la qualità ambientale si basa sulla presenza e sullo stato di salute della copertura vegetale di un dato territorio.

Le SdC 2021 si prestano molto bene all’analisi integrata con prodotti di *image processing*, a partire soprattutto dalle foto aeree (ortofoto) di AGEA, che presentano una risoluzione spaziale molto più compatibile con le SdC 2021 rispetto alle immagini da satellite più utilizzate (ad esempio Sentinel e LANDSAT).

Una semplice elaborazione è l’estrazione dell’indice radiometrico di vegetazione NDVI (*Normalized difference vegetation Index*)⁶.

Una prima sperimentazione è stata sviluppata considerando i centri abitati principali delle 14 Città metropolitane. Nella Figura 6.1 viene presentato un esempio di estrazione del NDVI per il centro abitato principale del comune di Bologna. Disponendo del dato georiferito delle aree verdi all’interno di un centro urbano, oltre al calcolo di indicatori legati alla popolazione residente (ad esempio quantità di verde per abitante), si possono fornire anche indicazioni sulla fruibilità, accessibilità, qualità delle aree verdi stesse. Incrociando tali risultati con le indicazioni estratte da carte tematiche, si può prevedere anche una classificazione delle aree coperte da vegetazione (ville storiche, impianti sportivi, parchi urbani, eccetera).

Figura 6.1 - Risultato dell’estrazione del NDVI per il centro abitato principale del Comune di Bologna



Fonte: Elaborazioni Istat su dati AGEA e Istat

6 Cfr. https://it.wikipedia.org/wiki/Normalized_Difference_Vegetation_Index.

I dati così calcolati, oltre a complementare le statistiche legate alle località principali, possono essere di supporto per la validazione dei risultati legati ad alcune rilevazioni ufficiali effettuate dall'Istituto⁷.

Dai primi risultati ottenuti, assolutamente provvisori, emerge che le città metropolitane con più clorofilla, in percentuale rispetto alla propria estensione, sono Bologna e Firenze con oltre il 30 per cento. Le Città metropolitane siciliane si attestano tra l'11 per cento di Messina e il 21 per cento circa di Palermo. Roma ha circa il 21 per cento della sua superficie coperta da aree verdi; Milano circa il 23 per cento; Napoli tra il 15 e il 16 per cento; Torino circa il 25 per cento. Le altre principali realtà urbane italiane variano tra il 15 per cento di Bari e il 25 per cento di Reggio di Calabria.

6.3 Versatilità della cartografia dell'Istat

Da quanto descritto sopra, le nuove BT Istat 2021 (in particolare modo lo strato relativo alle SdC 2021), oltre a essere un *layer* geografico ottimale per la descrizione statistica dell'intero territorio nazionale, permettono l'integrazione con altri strati informativi georiferiti che possono ampliare l'informazione geografica prodotta dall'Istat.

La domanda di dati statistici georiferiti, anche a grandi scale con maggiore dettaglio geografico, sta crescendo esponenzialmente sia a livello nazionale sia continentale. La versatilità della cartografia istituzionale dell'Istat sicuramente può venire incontro alle esigenze dell'utenza a qualsiasi livello.

Ciò che è stato descritto in questo Capitolo è solo relativo a due esempi di possibile integrazione tra le BT Istat e altri strati georiferiti. Altri esempi sono riportati in bibliografia (Chiocchini *et al.* 2023; Lipizzi *et al.* 2021).

⁷ Cfr. IstatData, Ambiente Urbano: https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1.Z0920ENV.1.0/ENV_CITIES.

7. LA COMPONENTE INDIRIZZI IN RSBL¹

7.1 Struttura, composizione e aggiornamento del Registro

Il Registro statistico di base dei luoghi (RSBL), come menzionato in precedenza, è uno dei pilastri di tutte le attività che prevedono la georeferenziazione delle informazioni statistiche contenute negli altri registri dell'Istituto o raccolte attraverso le indagini statistiche.

La componente indirizzi di RSBL risponde alla necessità di strutturare un registro che contenga, a livello nazionale, tutti gli indirizzi esistenti poiché questi ultimi, essendo un *asset* informativo di importanza cruciale, sono presenti in una quantità enorme di dati pubblici e di indagini statistiche. Gli indirizzi sono utili per collocare nel territorio le unità statistiche socioeconomiche (popolazione e unità locali), senza tralasciare dati di natura ambientale. Essi costituiscono, inoltre, un elemento indispensabile per l'invio postale dei questionari e delle lettere informative da parte di Istat verso i rispondenti alle indagini e per la reperibilità da parte dei rilevatori in fase di intervista. Infine, in una ottica di Censimento permanente della Popolazione e delle abitazioni, in cui la popolazione è campionata annualmente, l'indirizzo può essere utilizzato come lista di campionamento nelle indagini di copertura.

Questa componente di RSBL viene alimentata acquisendo informazioni relative agli indirizzi presenti negli archivi amministrativi e alle coordinate geografiche provenienti da fonti commerciali e amministrative (come, ad esempio, gli archivi geografici disponibili presso le Regioni). Ogni indirizzo è abbinato a indicatori di qualità, a coordinate geografiche e alla relativa sezione di Censimento. Tale infrastruttura permette la georeferenziazione e la geocodifica alla sezione di Censimento e alla griglia regolare delle unità statistiche (individui, famiglie, unità locali, eccetera).

La componente indirizzi si caratterizza per non avere in sé soltanto il "contenitore" che custodisce tutti gli indirizzi delle fonti amministrative che tratta, ma anche al suo interno un processo di integrazione delle fonti.

La prima caratteristica è che ogni indirizzo² presente in RSBL dispone di un Codice univoco di indirizzo (CUI), il quale identifica, attraverso un apposito codice numerico e con assenza di ambiguità, la stringa indirizzo. Come detto poc'anzi, la componente indirizzo raccoglie differenti fonti amministrative al suo interno. Dunque, il CUI fa sì che identici indirizzi, anche di fonti diverse, abbiano sempre lo stesso codice identificativo. Il CUI svolge, quindi,

¹ Il Capitolo è stato redatto da Davide Fardelli.

² Pure non essendo questa la sede per specificare cosa rappresenti e come sia strutturato un indirizzo, di seguito qualche dettaglio. In generale possiamo affermare che ogni spazio del suolo pubblico, o aperto al pubblico, costituisce un'area di circolazione. L'amministrazione comunale provvede alla denominazione di tutte le aree di circolazione viaria. Ogni area di circolazione è distinta dal toponimo stradale, detto "odonimo". Quest'ultimo è costituito dalla Denominazione urbanistica generica (DUG) che identifica la specie o la tipologia di area di circolazione (ad esempio via, viale, piazza, eccetera) e dalla Denominazione urbanistica ufficiale (DUF) che specifica e indica il nome proprio assegnato a ogni area di circolazione. L'elenco degli odonimi utilizzati nel territorio comunale forma lo stradario. L'indirizzo, invece, identifica l'accesso che dall'area di circolazione immette, direttamente o indirettamente, alle unità ecografiche (abitazioni, esercizi, uffici, eccetera). Il numero civico è l'identificazione dell'accesso esterno e si presenta come il valore assegnato all'accesso stesso secondo la successione naturale dei numeri oppure secondo il sistema metrico. Esso può essere accompagnato da una parte letterale del numero civico, denominato esponente.

una funzione identificativa in modo univoco per ogni elemento acquisito all'interno del Registro. Questo codice è fondamentale per evitare ambiguità e garantire l'accuratezza delle informazioni. Tutti gli indirizzi nel momento in cui entrano all'interno del Registro hanno anche propri codici che fanno riferimento alla fonte di appartenenza. Ciò facilita il riconoscimento e l'integrazione nel Registro e tra i registri dell'Istituto, soprattutto nell'ambito del Sistema integrato dei registri (SIR).

Una seconda caratteristica che contraddistingue questa componente è la capacità di integrare indirizzi provenienti non solo da diversi archivi amministrativi, ma anche dalla stessa fonte in periodi temporali differenti. Questo processo di integrazione consente di unire dati eterogenei, garantendo una visione completa e coerente, sia nel caso in cui provengano da fonti diverse, sia quando si tratta di dati raccolti dalla stessa fonte, ma in annualità diverse.

La fase di integrazione è possibile grazie al riconoscimento dell'indirizzo attraverso un software di normalizzazione, procedimento che consente di standardizzare i dati relativi agli indirizzi, al fine di renderli coerenti dentro e fuori le fonti, e facilmente utilizzabili. Questo processo è fondamentale per garantire l'accuratezza e l'affidabilità dei dati. Infatti, in base alle casistiche analizzate, possono verificarsi casi anomali relativi al riconoscimento degli indirizzi. Due stringhe di testo, anche se semanticamente differenti, potrebbero identificare il medesimo luogo (Prospetto 7.1). Dunque, si tratta di indirizzi che, pur indicando lo stesso luogo fisico, sono scritti o formattati in modi differenti. Questa difformità semantica, derivante da variazioni linguistiche, abbreviazioni, errori di ortografia o formattazioni diverse, viene affrontata e corretta nella normalizzazione degli indirizzi. Infatti, quando gli indirizzi sono scritti in modi diversi, potrebbe risultare non automatico e non privo di errori, per database e sistemi informativi, riconoscere che si tratti dello stesso luogo. Questo può portare a duplicazioni all'interno del Registro e a difficoltà nell'integrazione dei dati. La normalizzazione degli indirizzi, quindi, trasforma gli indirizzi semanticamente diversi in un formato standardizzato e univoco. Sinteticamente si individuano le seguenti fasi:

1. standardizzazione: si convertono tutte le variazioni in un formato uniforme. Ad esempio, convertire tutte le abbreviazioni ("V." diventa "Via", "P.zza" diventa "Piazza", "V.le" diventa "Viale") e riportare a uno standard predefinito gli elementi;
2. correzione degli errori: si identificano e si correggono gli errori di ortografia e tipografici. Si utilizza il database di riferimento per verificare la correttezza degli indirizzi;
3. riconciliazione dei dati: si uniscono *record* duplicati che rappresentano lo stesso indirizzo, garantendo che ogni luogo fisico sia rappresentato una sola volta nel database.

Un indirizzo non normalizzato non è necessariamente un indirizzo errato, ma difficilmente può essere confrontabile e integrabile con altre fonti a causa della difformità negli standard di compilazione delle singole componenti o per assenza di alcuni di essi.

Diverse sono le tecniche di normalizzazione presenti sul mercato, tra le più conosciute ci sono gli algoritmi di Matching, utilizzati per confrontare e riconciliare indirizzi simili. Possono includere tecniche di similarità testuale e *fuzzy matching*. Altro strumento sono i database di riferimento, i quali sono utilizzati per verificare e standardizzare gli indirizzi, come ad esempio, elenchi ufficiali di vie e codici postali. Infine, sono disponibili software di normalizzazione, i quali automatizzano il processo di normalizzazione degli indirizzi, come *Google Address API*, *Melissa Data*, *BatchGeo*, *SmartyStreets* ed *Egon*.

L'attuale registro ha utilizzato una versione ibrida nella quale si utilizza un software di normalizzazione di natura commerciale che consente di riconoscere gli indirizzi attraverso metodi sia deterministici, sia di confronto tra stringhe rispetto a un proprio dizionario. In appoggio al software di normalizzazione si dispone dell'Archivio Nazionale dei Numeri Civici

7. La componente indirizzi in RSBL

delle Strade Urbane³ (ANNCUSU), il quale rappresenta la *spina dorsale* della componente indirizzi di RSBL. Attraverso l'ANNCUSU, infatti, vengono definite le forme ufficiali e standardizzate dell'indirizzo, in quanto seguono regole normative e circolari ufficiali, producendo uno standard unico dell'indirizzo.

Questo processo di integrazione e di riconoscimento delle differenti stringhe di indirizzo, comporta una deduplicazione delle stesse, sia interne sia esterne alle fonti. Ciò, come detto, è dovuto alle diverse modalità nelle quali è scritto un indirizzo. In questo modo all'interno del Registro degli indirizzi, vi è una sezione dove ci sono gli indirizzi che entrano per la prima volta all'interno del registro (sezione *Indirizzi*) e una seconda sezione (denominata *Thesaurus*), dove sono archiviate tutte le forme nelle quali si è presentato l'indirizzo nelle diverse fonti e all'interno della fonte stessa per entrare nel Registro. In questa sezione si vengono a creare dei veri e propri grappoli di indirizzi con lo stesso CUI (Prospetto 7.1).

Prospetto 7.1 - Esempio di rappresentazione della componente di RSBL con la sezione *Indirizzi* e *Thesaurus*

CUI	Indirizzo	Fonte	Anno	ID	Sezione RSBL
100	Viale Alessandro Manzoni, 5	ANNCUSU	2023	4522	Indirizzi
100	Via A Manzoni, 5	LAC	2023	7821	Thesaurus
100	V.le Manzoni, 5	LAC	2018	2547	Thesaurus
100	Viale Manzoni, 5	AT	2023	3654	Thesaurus
200	Via Venti Settembre	ANNCUSU	2023	5412	Indirizzi
200	Via XX Settembre	AT	2023	4501	Thesaurus
200	Via 20 Settembre	LAC	2023	9601	Thesaurus

Fonte: Elaborazioni su dati Istat

Una terza peculiarità che caratterizza la componente indirizzi di RSBL è l'associazione dell'indirizzo (CUI) a una coordinata geografica (latitudine e longitudine), la quale produce indirizzi georeferenziati. Come detto in precedenza, uno degli obiettivi di RSBL è quello della localizzazione nello spazio delle unità statistiche. In passato, l'Istat riusciva a produrre dati dettagliati a livello subcomunale (fino alle sezioni di Censimento) per le unità statistiche di tipo censuario. Oltre allo svantaggio di non avere un dettaglio geografico maggiore, si aveva anche il danno di avere comunque aree eterogenee, soprattutto nel tempo. La localizzazione alla sezione di Censimento (geocodifica) dell'unità statistica avveniva prevalentemente con una associazione tabellare dell'indirizzo alla sezione di Censimento, dunque, non sempre si utilizzavano operazioni di carattere spaziale per poterne ricavare il risultato. Questo tipo di approccio è definito come *area based*, nell'ambito del quale le unità statistiche erano assegnate a un'area, come ad esempio le sezioni di Censimento. Con la costruzione di RSBL, cambia l'approccio nella geocodifica delle unità statistiche, e si considera un metodo definito a *point-based geocoding infrastructure*. Questo avviene in maniera consistente

3 L'Archivio Nazionale dei Numeri Civici delle Strade Urbane (ANNCUSU), realizzato dall'Istat e dall'Agenzia delle Entrate, è un archivio informatizzato contenente gli stradari (elenco delle denominazioni delle aree di circolazione) e i numeri civici di tutti i comuni italiani, da questi ultimi certificati. Più in generale, l'introduzione dell'ANNCUSU comporterà un abbattimento dei costi legato alla gestione degli indirizzi, in quanto, sfruttando una economia di scala a livello nazionale, consentirà di disporre di un archivio informatizzato, standardizzato, codificato e dinamicamente certificato dai Comuni, da utilizzarsi come riferimento nella trattazione degli indirizzi da parte delle Pubbliche amministrazioni, dei cittadini e delle imprese, potendo assicurare nel tempo, grazie ai processi di aggiornamento e allineamento, un grado di integrazione sempre maggiore con altre banche dati nazionali. L'ANNCUSU costituisce, dalla data di attivazione, il riferimento nazionale per gli stradari e indirizzari comunali. L'aggiornamento è di competenza dei Comuni e il Comune utilizza nell'ambito delle attività di competenza esclusivamente i dati presenti nell'ANNCUSU. Eventuali disallineamenti tra gli indirizzi presenti negli atti rilasciati dal Comune successivamente alla data di istituzione dell'ANNCUSU, vanno segnalati direttamente al Comune. Cfr. <https://www.anncsu.gov.it>

considerando le coordinate geografiche dell'indirizzo, strumenti complementari per la localizzazione di fenomeni. Gli indirizzi offrono comprensibilità e usabilità, mentre le coordinate geografiche conferiscono accuratezza e universalità. Utilizzarli insieme può migliorare notevolmente l'efficacia delle operazioni di localizzazione e analisi. Gli indirizzi associati alle coordinate geografiche, offrono numerosi vantaggi rispetto agli indirizzi tradizionali basati solo su descrizioni testuali. Infatti, le coordinate geografiche forniscono una posizione precisa, evitando ambiguità ed errori comuni negli indirizzi.

Un secondo aspetto da citare è l'invariabilità. Le coordinate geografiche tendenzialmente, a meno di aggiornamenti qualitativi, non cambiano nel tempo, a differenza degli indirizzi che possono essere soggetti a modifiche dovute a nuove costruzioni, cambiamenti amministrativi, ridenominazioni stradali o rinumerazione civica della strada. L'assunzione delle coordinate geografiche come salto di qualità per collocare nello spazio le unità statistiche, comunque, non esula da problemi localizzativi. Le principali criticità sono: la copertura e la disponibilità di coordinate geografiche; la disponibilità di coordinate geografiche di qualità, rilevate a una scala ottimale e che presentino una accuratezza posizionale ottimale. Occorre, infatti, considerare la corretta associazione tra la stringa indirizzo e la coordinata geografica (X, Y) .

Un'ultima peculiarità della componente indirizzi di RSBL è la valutazione della qualità, nelle sue diverse dimensioni. Gli aspetti qualitativi sono una parte cruciale del Registro in quanto permettono il passaggio da archivio amministrativo a registro statistico attraverso la stesura di regole di eleggibilità per potere selezionare quali indirizzi, tra quelli raccolti nelle attività dell'Istituto, devono costituire la componente indirizzi e individuare un set di indicatori di qualità per valutare ciascun *record*.

La misurazione della qualità della componente indirizzi richiede la definizione di indicatori specifici come, ad esempio, il grado di completezza dell'indirizzo o la presenza delle altre variabili. La completezza dell'indirizzo concerne tutte le informazioni che sono necessarie a descrivere completamente l'indirizzo ovvero: DUG, DUF, civico e l'eventuale esponente. Le altre variabili sono rilevanti per le operazioni di localizzazione. Queste sono: i codici amministrativi, la sezione di Censimento, l'ID della griglia regolare europea e le coordinate geografiche.

Un terzo aspetto da tenere in considerazione è la coerenza tra la geocodifica alla sezione di Censimento e la georeferenziazione. Si veda il paragrafo successivo per maggiori approfondimenti.

Ultimo aspetto da valutare è il grado di abbinamento tra gli indirizzi provenienti da diverse fonti. Si tratta di misurare l'affidabilità dell'integrazione tra le diverse fonti valutando, ad esempio, il numero di fonti nelle quali l'indirizzo si presenta, opportunamente pesate, o il tipo di relazione (deterministica o probabilistica).

Una volta costruita la componente indirizzi, la valutazione della sua qualità complessiva richiede altri indicatori. Tra questi vi sono:

- accuratezza posizionale: nel caso in cui le coordinate X e Y o le sezioni di Censimento si discostino dalla posizione rilevata da una fonte più accurata;
- completezza: intesa come sovracopertura del Registro, ovvero l'inclusione di indirizzi non più esistenti sul territorio, e sottocopertura del Registro ossia l'assenza nella componente indirizzi di unità esistenti sul territorio;
- accuratezza sintattica dell'indirizzo: in questo caso si tratta di misurare i falsi non abbinamenti, ovvero quei *record* delle basi di dati che fanno riferimento alla stessa unità ma il *linkage* non è in grado di individuarli, in quanto almeno una variabile chiave

7. La componente indirizzi in RSBL

è affetta da qualche errore. In questo caso la normalizzazione non è andata a buon fine e gli indirizzi scartati faranno parte dell'insieme dei residui (è una delle possibili cause di sovracopertura). I falsi abbinamenti, ovvero alcuni *record* possono essere abbinati anche se in realtà fanno riferimento a unità diverse, rappresentano una delle possibili cause di sottocopertura⁴;

- accuratezza temporale: quando le variabili del registro hanno periodi di riferimento tra loro differenti.

Il primo impianto di RSBL, realizzato nel 2018 in forma prototipale e nel 2020 nella sua attuale struttura, era popolato da 3 fonti: l'Archivio Nazionale dei Numeri Civici delle Strade Urbane (ANNCSU); gli indirizzi estratti dalle Liste anagrafiche comunali (LAC); gli indirizzi dell'Anagrafe tributaria (AT). Successivamente sono stati caricati gli indirizzi degli immobili del catasto, gli indirizzi provenienti dal Censimento generale della Popolazione e delle abitazioni del 2011 e, infine, gli indirizzi provenienti da alcune indagini statistiche. Nella versione di giugno 2024 si contano in RSBL circa 30,2 milioni di CUI nella sezione *Indirizzi* e di 93 milioni di indirizzi nella sezione *Thesaurus*, per un totale di 123 milioni di stringhe indirizzo trattate, riconosciute ed entrate in RSBL.

La fonte ANNCSU è considerata la struttura portante del Registro in quanto rappresenta un archivio amministrativo a livello centrale che ha lo scopo di raccogliere tutti gli indirizzi presenti nel territorio di tutti i comuni. Esso è aggiornato dai singoli comuni, sotto il controllo di Istat e Agenzia delle Entrate, i quali verificano il rispetto delle regole e delle specifiche tecniche⁵. Pertanto, L'ANNCSU non solo dovrebbe garantire la massima copertura degli indirizzi nel territorio nazionale, ma dovrebbe garantire anche la coerenza al suo interno con le regole di standardizzazione, onde evitare duplicazioni intracomunali. Dal 2018 a oggi, oltre al primo caricamento, l'ANNCSU è stato aggiornato con cadenza pluriennale (2019, 2021, 2024). Nella componente indirizzi sono presenti circa 26,7 milioni di CUI provenienti dall'ANNCSU, i quali rappresentano l'88,5% dei CUI totali presenti nel registro. Nonostante si consideri l'ANNCSU un archivio controllato in entrata, presumibilmente per ragioni legate all'aggiornamento da parte dei comuni, il passaggio da archivio a registro comporta l'esclusione di circa 1,4 milioni di indirizzi (5 per cento), poiché questi non superano i controlli interni che rispettano le regole stabilite per potere attribuire un CUI all'indirizzo della fonte amministrativa.

La seconda fonte che alimenta la componente indirizzi di RSBL è la LAC (ora divenuta Anagrafe nazionale della popolazione residente - ANPR), dove gli indirizzi sono a cura dell'ufficio dell'Anagrafe di ogni singolo comune. Questa fonte è di una certa importanza poiché i CUI provenienti da essa sono il principale elemento di connessione con il Registro base degli individui e delle famiglie (RBI)⁶. In esso sono contenuti gli indirizzi delle LAC a partire dal 2014 fino al 2024. Vista l'importanza della fonte il suo aggiornamento avviene annualmente. Il totale degli indirizzi di fonte LAC sono circa 19,8 milioni e integrano gli indirizzi dell'ANNCSU nel registro indirizzi, come seconda fonte, per circa 1 milione di CUI, rappresentando il 3,6 per cento dei CUI totali. Per questa fonte è dedicato ampio spazio nelle procedure per il trattamento dell'indirizzo, affinché venga riconosciuto dal software

4 Per un ulteriore approfondimento sui problemi del *record linkage* cfr. Scanu M. 2013. *Metodi statistici per il record linkage*. Roma: Istat. https://www.istat.it/it/files/2013/12/met_norme_03_16_metodi_statistici_record_linkage.pdf.

5 Per approfondimenti si leggano le Istruzioni per l'Ordinamento Ecografico. https://www.agenziaentrate.gov.it/portale/documenti/20143/266134/02+Istruzioni+per+l%27ordinamento+ecografico_agg_20180801+per+stampa.pdf/17790a4d-3965-99d7-f3a1-75530cb7c736.

6 Attualmente la quota di popolazione che è associata a un CUI è pari al 97%. La restante parte di popolazione è priva di CUI in quanto gli indirizzi della fonte hanno un nome della strada non riconducibile agli standard qualitativi del Registro oppure sono primi di numerazione civica (SNC).

di normalizzazione. Le procedure di trattamento, pervenendo anche all'utilizzo di *linkage* probabilistici, cercano di riportare gli indirizzi a una forma standardizzata. Queste sono necessarie in quanto dalle fonti di provenienza non ci sono vincoli che ne impediscano l'ingresso. Gli indirizzi di fonte LAC, che non entrano nel Registro indirizzi, sono circa 200 mila.

La terza sorgente che alimenta RSBL è l'Anagrafe tributaria. In questo caso si tratta della fonte più complessa e con un contenuto ad alta entropia, in quanto gli indirizzi non sono compilati direttamente da un ente amministrativo (come il caso delle LAC), ma sono tutti gli indirizzi ai quali sono associati principalmente i codici fiscali e le dichiarazioni dei redditi delle persone fisiche. Ciò fa sì che per questo archivio si abbia una quota sproporzionata tra indirizzi in forma grezza della fonte e CUI entrati nel Registro. Come detto in precedenza, ciò è dovuto principalmente alle diverse forme, dunque duplicazioni, degli indirizzi che contribuiscono ad alimentare la sezione del *Thesaurus* del Registro. Gli indirizzi trattati ed entrati nella componente con un CUI sono circa 16,6 milioni, rispetto ai 50 milioni di indirizzi totali presenti nell'archivio amministrativo. Tale fonte porta a completare il registro indirizzi per una quota di circa 1,3 milioni di indirizzi (4,1 per cento). Gli indirizzi non riconosciuti e privi di CUI sono circa 3,5 milioni.

Il quarto archivio utilizzato e caricato nella componente è costituito da tutti gli indirizzi utilizzati durante l'ultimo Censimento generale della Popolazione e delle abitazioni del 2011. Questo archivio, in realtà, non può essere considerato al pari di una fonte di natura amministrativa, ed è stato utilizzato per recuperare le informazioni sulla geocodifica degli indirizzi alle sezioni di Censimento 2011. Si è trattato di un caricamento effettuato una sola volta in quanto non ci sono stati più aggiornamenti. Gli indirizzi trattati sono stati circa 10,8 milioni e la parte complementare che è andata a coprire i CUI ancora non presenti in RSBL è abbastanza marginale, pari a circa 79 mila CUI.

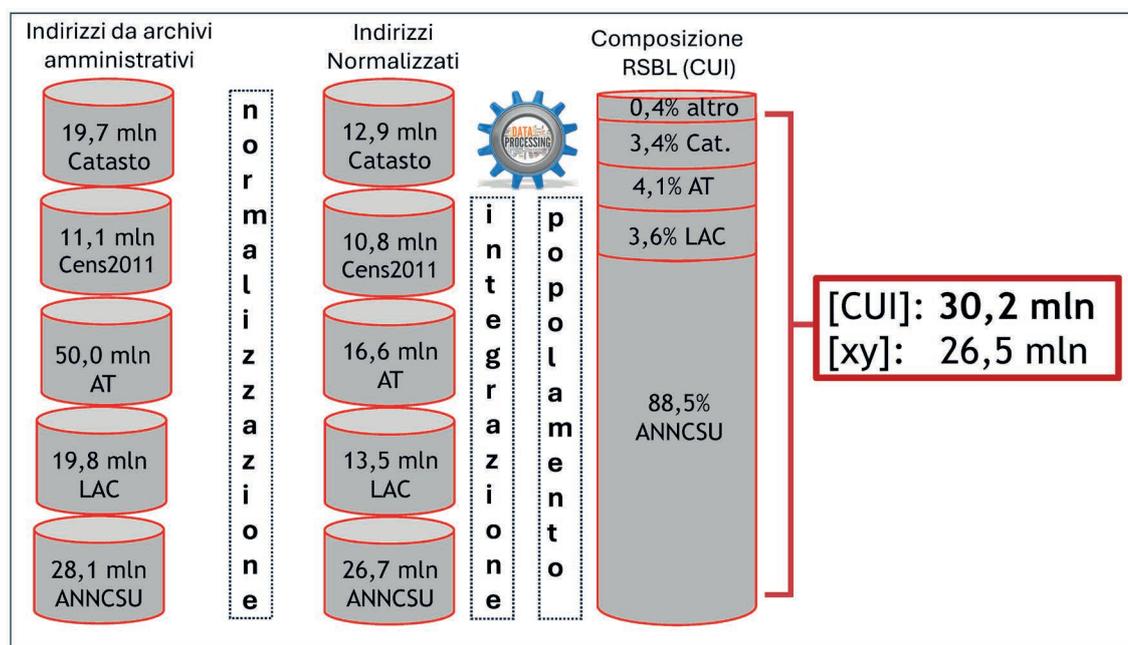
La quinta e ultima fonte amministrativa, che popola la componente indirizzi di RSBL, è il Catasto. Da questa fonte sono estratti tutti gli indirizzi ai quali è associato un immobile. Anche questa fonte riveste una certa importanza. Infatti, attraverso gli indirizzi di questa provenienza, può avvenire tanto l'integrazione interna tra la componente indirizzi e quella degli edifici e abitazioni, sia il completamento esterno con RBI. Tale integrazione permette anche la localizzazione dell'unità statistica (individuo/famiglia) all'interno del territorio comunale, anche attraverso l'utilizzo delle coordinate del centroide dell'edificio recuperato dalla cartografia catastale. Tali operazioni, denominate di geocodifica, hanno permesso di diffondere la distribuzione della popolazione legale relativa al Censimento 2021 sulla griglia regolare⁷, con celle di 1 km². Il dato è stato pubblicato e diffuso da Istat ed Eurostat, e hanno distribuito il dato per sezione di Censimento (Basi territoriali 2021). Gli indirizzi trattati dalla fonte catastale sono circa 19,7 milioni con un riconoscimento tramite CUI pari a circa 12,9 milioni di indirizzi e un conseguente non riconoscimento pari a circa 800 mila indirizzi (5,5 per cento). Tale fonte porta a completare il Registro indirizzi per una quota di circa 1 milione di indirizzi (3,4 per cento). Anche questa fonte ha un trattamento degli indirizzi dedicato ed esclusivo. La sua principale peculiarità è la compresenza di tre numeri civici per ogni stringa indirizzo (ad esempio Via Roma 1, 3, A oppure Via Milano 2, 4, A). Questa criticità, la quale tende a invalidare l'indirizzo, è stata sfruttata per fare sì che ogni civico superfluo diventasse un nuovo indirizzo. Questo trattamento, denominato verticalizzazione degli indirizzi, rispetta regole ben precise e stringenti che hanno dato luogo a circa 1 milione di indirizzi in più provenienti da fonte catastale, quasi del tutto inseriti in RSBL.

⁷ Cfr. <https://www.istat.it/notizia/statistiche-sulla-popolazione-per-griglia-regolare/>.

7. La componente indirizzi in RSBL

Infine, seppure in maniera molto marginale sono da annoverare anche gli indirizzi provenienti dalle indagini censuarie, come ad esempio l'Indagine areale e l'Indagine relativa alle popolazioni speciali (convivenze, campi attrezzati e senza fissa dimora). Queste ulteriori fonti hanno contribuito con circa 100 mila indirizzi. Per quanto concerne questi ultimi ingressi, non bisogna vedere l'aspetto marginale quantitativo su RSBL, ma l'aspetto qualitativo, poiché sono indirizzi che molto probabilmente non emergono da fonti amministrative, ma provengono direttamente da rilievi sul campo (rilevatori). Dunque, potrebbe trattarsi anche di nuovissimi indirizzi, appena nati che, dato il flusso di aggiornamento delle fonti, ancora non vengono recepiti dagli archivi stessi e, quindi, importanti perché in grado di soddisfare l'esigenza di una maggiore tempestività.

Figura 7.1 - La composizione della componente indirizzi di RSBL



Fonte: Elaborazioni su dati Istat

Il quadro rappresentato (Figura 7.1) raffigura la composizione attuale della componente indirizzi di RSBL con tutti gli archivi amministrativi che lo alimentano e le possibili integrazioni con gli altri registri dell'Istituto. È in fase di analisi lo studio di fattibilità per potere fare entrare anche gli indirizzi provenienti dal mondo economico. Tra gli archivi amministrativi più importanti si ricordano l'Anagrafe tributaria delle persone giuridiche e delle persone fisiche con partita Iva e Il Registro delle imprese di Unioncamere. Il riconoscimento degli indirizzi tramite CUI con questi archivi amministrativi, potrebbe permettere la connessione e l'integrazione di RSBL al Registro delle Unità Economiche del SIR.

Anche se in via sperimentale, sono state fatte le prime valutazioni quantitative e qualitative a livello di riconoscimento degli indirizzi e di geocodifica alle BT2011 delle unità locali. È emerso che circa il 95 per cento delle unità locali (sia attive sia non attive) hanno un indirizzo riconducibile a un CUI all'interno di *RSBL-Indirizzi*. Altra informazione interessante, che è emersa dalla sperimentazione, è che gli indirizzi che potenzialmente potrebbero entrare in RSBL ammontano a pochissime migliaia. Ciò sta a testimoniare la buona copertura di indirizzi presente all'interno del Registro a livello nazionale, che ingloba sia il tessuto demografico, sia quello economico.

7.2 L'indirizzo come elemento di geocodifica dell'unità statistica nel SIR

RSBL è in accordo col primo principio del GSGF, *Global Statistical Geospatial Framework*⁸ (cfr. Capitolo 10), il quale raccomanda l'adozione di un approccio comune e coerente per collocare ogni unità statistica di un *dataset* nel tempo e nello spazio, utilizzando infrastrutture geospaziali di riferimento. Secondo il GSGF, le unità statistiche devono essere geocodificate utilizzando identificativi geospaziali, così come particelle catastali, indirizzi o edifici. Successivamente il dato è riaggregato e rilasciato secondo il livello geografico/amministrativo di diffusione (comune, area subcomunale, sezione di Censimento, griglia).

L'infrastruttura RSBL si propone di seguire i principi GSGF adottando un approccio *point-based*, in modo tale da avere una produzione statistica flessibile, che può permettere di avere diversi livelli spaziali fino ad arrivare a livelli amministrativi subcomunali, come ad esempio, le sezioni di Censimento oppure le griglie regolari.

Il processo di geocodifica che porta all'associazione tra l'indirizzo (CUI) e la sezione di Censimento (BT2011) ha coinvolto tutte le fonti amministrative e geografiche che avevano informazioni spaziali, utilizzabili a concorrere e a costruire una base dati esauriente e affidabile.

Le principali fonti disponibili e adoperate per l'attribuzione della sezione di Censimento sono:

- 25,8 milioni di indirizzi con sezione di Censimento proveniente dall'ANNCSU;
- 23,5 milioni di indirizzi con sezione di Censimento proveniente dalle attività censuarie del Censimento 2011;
- 26,8 milioni di indirizzi con sezione di Censimento proveniente da coordinate geografiche con diversa qualità (rilevate/interpolate/approssimate).

Utilizzando le fonti sopra descritte si è proceduto a verificare il livello di uguaglianza/discordanza delle fonti per potere geocodificare l'indirizzo e creare un indicatore di qualità della geocodifica che misuri l'attendibilità della stessa. L'attendibilità della geocodifica è determinata sia dalla fonte utilizzata per geocodificare l'indirizzo, sia dal livello di concordanza/discordanza tra le diverse fonti.

Tutti i valori delle sezioni sono stati confrontati con la validità delle Basi territoriali del 2011. Se il valore non è riscontrato nelle BT del 2011, tale valore non è ammissibile e di conseguenza viene scartato. Dopo questa prima fase, è entrato nel processo di geocodifica anche la fonte LAC. Nella trasmissione delle LAC, i comuni hanno la possibilità di compilare e inviare l'informazione della geocodifica dell'indirizzo. Tale fonte è utilizzata solamente per gli indirizzi con una bassa attendibilità della geocodifica oppure totalmente privi. I valori della sezione proveniente dalla LAC servono per corroborare la sezione confermata per la geocodifica del CUI. Gli indirizzi la cui sezione di Censimento è stata utilizzata per il processo di geocodifica sono circa 300 mila.

Attraverso questa metodologia di geocodifica del CUI alla sezione di Censimento, si è arrivati a un risultato nel quale il 73 per cento dei CUI hanno una geocodifica ad alta attendibilità, il 18 per cento ha una geocodifica a media attendibilità, l'8 per cento dei CUI ha una geocodifica a bassa attendibilità e un 1 per cento è privo di geocodifica.

Si sono, quindi, considerate altre fonti e metodologie per poter recuperare la sezione di Censimento: la sezione indagine e la sezione imputata.

La sezione indagine proviene dall'indagine effettuata da Istat presso i Comuni italiani (dicembre 2020) per i CUI senza una qualsiasi informazione geografica e quelli con una geocodifica a bassa attendibilità. Dal ritorno dell'indagine, dopo uno stringente processo di

8 https://unstats.un.org/unsd/statcom/51st-session/documents/The_GSGF-E.pdf.

7. La componente indirizzi in RSBL

controllo e correzione del dato, si è valutata la qualità dell'informazione. Ciò è stato utile in quanto ha accresciuto la copertura degli indirizzi geocodificati di circa 100 mila unità.

L'ultimo passo per potere geocodificare l'indirizzo proviene attraverso metodologie di imputazione (sezione imputata) a carattere spaziale. Nello specifico, si è tentato di geocodificare l'indirizzo privo di sezione di Censimento cercando, nelle sue vicinanze (prossimità dei numeri civici), un indirizzo geocodificato. Una seconda metodologia utilizzata ha considerato l'imputazione della sezione di Censimento agli indirizzi privi di geocodifica la cui strada ricade in una o in due sezioni di Censimento. Con questa metodologia è stata attribuita la sezione di Censimento a circa 200 mila indirizzi. La quota di indirizzi rimasti privi di geocodifica ammonta a circa 150 mila.

Questa infrastruttura ha contribuito e ha permesso lo svolgimento, la produzione e la diffusione dei dati sia per sezione di Censimento⁹, sia per griglia regolare¹⁰, relativi al Censimento permanente della Popolazione e delle abitazioni.

Si tratta di un processo che richiede continui aggiornamenti, soprattutto con il rilascio delle nuove sezioni di Censimento (BT2021). Trattandosi di un procedimento complesso, sarà necessario revisionare il principio di assegnazione del CUI alla nuova sezione di Censimento in base alle nuove risultanze dei confronti tra le fonti, per le quali assume rilievo la sezione proveniente dal centroide dell'edificio (fonte catastale). In questa nuova *release* avremo per ogni CUI sia la coordinata geografica riferita all'accesso esterno sia al centroide dell'edificio associato all'indirizzo. Quest'ultima informazione potrà coprire quasi integralmente (97%) i CUI provenienti dalla fonte catastale. Dalle prime elaborazioni, ancora in fase di completamento, emerge che dall'integrazione dei dati geografici del catasto con *RSBL-Indirizzi* circa il 71 per cento della popolazione censuaria (2021) avrà anche una geocodifica alle BT21 direttamente connessa con il CUI. Successive analisi, sia statistiche sia spaziali, riusciranno sicuramente ad aumentare la copertura di georeferenziazione dei CUI.

9 Cfr. <https://www.istat.it/notizia/dati-per-sezioni-di-censimento/>.

10 Cfr. <https://www.istat.it/notizia/statistiche-sulla-popolazione-per-griglia-regolare/> e https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population_and_housing_census_2021_-_population_grids.

8. REGISTRO EDIFICI, ABITAZIONI E ALTRE UNITÀ IMMOBILIARI: IL PROCESSO DI INTEGRAZIONE E LOCALIZZAZIONE DEI DATI AMMINISTRATIVI¹

8.1 Il nuovo Registro edifici e alloggi: discontinuità con il passato e nuove possibilità

Gli edifici e le unità abitative (EUA) sono la parte del Registro statistico di base dei luoghi (RSBL) che fornisce dati essenziali per il Censimento della Popolazione e delle abitazioni e altre finalità istituzionali, come richiesto dal regolamento Eurostat 763/2008. Le informazioni sugli edifici includono caratteristiche strutturali e costruttive, mentre le unità abitative forniscono dettagli giuridici e strutturali. RSBL/EUA utilizza principalmente dati di fonte catastale e del Censimento 2011, integrati con altre fonti istituzionali e dati cartografici *open*, per fornire un quadro dettagliato dello *stock* edilizio, comprendendo sia edifici residenziali, sia non residenziali.

Il nuovo registro EUA rappresenta un cambiamento significativo rispetto alla rilevazione dedicata a edifici e abitazioni svolta nell'ambito del tradizionale Censimento generale della Popolazione e delle abitazioni. Al posto di una rilevazione puntuale tramite rilevatori, il nuovo registro si basa principalmente sull'utilizzo di archivi amministrativi, in particolare l'Archivio catastale. Questa scelta strategica comporta una serie di vantaggi e svantaggi che è importante considerare.

Tra i vantaggi del nuovo registro ci sono, ovviamente, gli stessi che hanno orientato l'intero Istituto a un maggiore utilizzo degli archivi amministrativi come ad esempio: la riduzione dei costi che, grazie all'utilizzo di archivi esistenti, elimina la necessità di costose operazioni di rilevazione sul campo; una maggiore efficienza grazie al processo di raccolta dei dati più rapido e automatizzato; una maggiore accuratezza dei dati provenienti da archivi ufficiali, generalmente considerati più affidabili di quelli raccolti tramite rilevazione diretta. Il registro è, infine, costantemente aggiornato con nuove informazioni, garantendo una visione più completa e attuale del patrimonio immobiliare.

Di contro, il nuovo registro sconta un minore dettaglio informativo. In particolare, tra le informazioni precedentemente rilevate sul campo e non disponibili negli archivi, si segnalano l'epoca di costruzione degli edifici, il loro stato di conservazione e la dotazione degli impianti di condizionamento delle abitazioni. Inoltre, a fronte di una generale maggiore completezza, i dati archiviati potrebbero, localmente, non essere completi o aggiornati.

Attualmente il registro conta, con riferimento all'anno 2021, oltre 35 milioni di abitazioni e 14,3 milioni di edifici residenziali (dati 2021). È in corso il controllo e la validazione del dato sugli edifici non residenziali tramite conteggio e integrazione con il Registro delle unità locali.

I due principali vantaggi della nuova architettura del registro sono:

- la disponibilità di codici identificativi specifici per le differenti unità, ovvero gli estremi catastali di immobili e fabbricati che permettono una precisa identificazione, riconoscibilità e tracciabilità nel tempo di ciascuna unità.

¹ Il Capitolo è stato redatto da Damiano Abbatini (paragrafo 8.1), Gianluigi Salvucci (paragrafo 8.2), Stefania Lucchetti (paragrafo 8.3).

- la localizzazione puntuale sul territorio (georeferenziazione), ottenuta grazie al lavoro di trasformazione dei file della cartografia catastale da semplici file di testo a *layer* vettoriali (in formato *shapefile*). Questa informazione è fondamentale sia per la geocodifica alle sezioni di Censimento 2021, sia per qualsiasi altro tipo di geocodifica.

Nei due paragrafi che seguono viene descritto con maggiore dettaglio l'uso che si fa di questi dati nel processo di integrazione di RSBL con gli altri registri, in particolare con RBI.

8.2 La cartografia catastale e l'integrazione delle altre fonti cartografiche

8.2.1 Il processo di acquisizione dei dati geografici

Il catasto, principale fonte di dati cartografici per il Registro degli edifici², è stato progettato su un sistema di riferimento geografico multiorigine, con una attenzione particolare al mantenimento delle distanze all'interno di fogli, considerati unità minime di coerenza geografica. Tuttavia, consuetudini e pratiche amministrative hanno compromesso, nel tempo, la possibilità di un riposizionamento automatico delle geometrie catastali, richiedendo correzioni puntuali che comportano un notevole dispendio di tempo e risorse (Carlucci 2024). Queste correzioni variano da semplici traslazioni a più complesse rototraslazioni delle geometrie. Nell'ambito della costruzione del Registro degli edifici dell'Istat, sono state adottate diverse metodologie per migliorare il georiferimento degli elementi catastali. Per rendere replicabile l'operazione in periodi successivi, è stata creata una banca dati dove, per ogni foglio catastale dell'anno t_0 , il file del sistema geografico dello *shapefile* (.prj) è stato corretto con uno spostamento uniforme che garantisce, per future forniture catastali al tempo t_1 , il riposizionamento automatico nella posizione desiderata. Tuttavia, questa procedura si è rivelata estremamente onerosa. Il catasto comprende oltre 300.000 fogli di mappa, e la verifica puntuale del loro posizionamento richiederebbe un controllo a videoterminale da parte di un operatore, con un enorme dispendio di tempo. D'altra parte, la qualità del lavoro di georiferimento è essenziale per garantire una corretta geocodifica delle informazioni alle sezioni di Censimento.

8.2.2 Il processo di georiferimento automatico: ASIR

Per migliorare il processo di riposizionamento è stato sviluppato il modello geografico ASIR (*Aggregate Shape Similarity Index - Relocation*) che combina due procedure: Duračiová (2023), con l'indicatore ASI per il riconoscimento di edifici simili, e Liu *et al.* (2020), con l'introduzione delle matrici di confronto e distanza, per migliorare il posizionamento dei dati catastali. Questo approccio si basa su studi che evidenziano l'importanza di una corretta georeferenziazione per la gestione dei dati spaziali.

La procedura mette a confronto due geodatabase di edifici: il primo, quello catastale, di cui si vuole controllare e misurare il corretto posizionamento, e il secondo, il Database di sintesi nazionale (DBSN)³ prodotto e rilasciato dall'Istituto geografico militare (IGM), usato

2 Si tratta della cartografia del catasto terreni, visualizzabile con servizio *wms on line* dal geoportale cartografico dell'Agenzia delle Entrate. https://geoportale.cartografia.agenziaentrate.gov.it/age-inspire/srv/ita/catalog_search#/home.

3 DataBase di Sintesi Nazionale (DBSN). <https://www.igmi.org/it/dbsn-database-di-sintesi-nazionale>.

8. Registro edifici, abitazioni e altre unità immobiliari: il processo di integrazione e localizzazione dei dati amministrativi

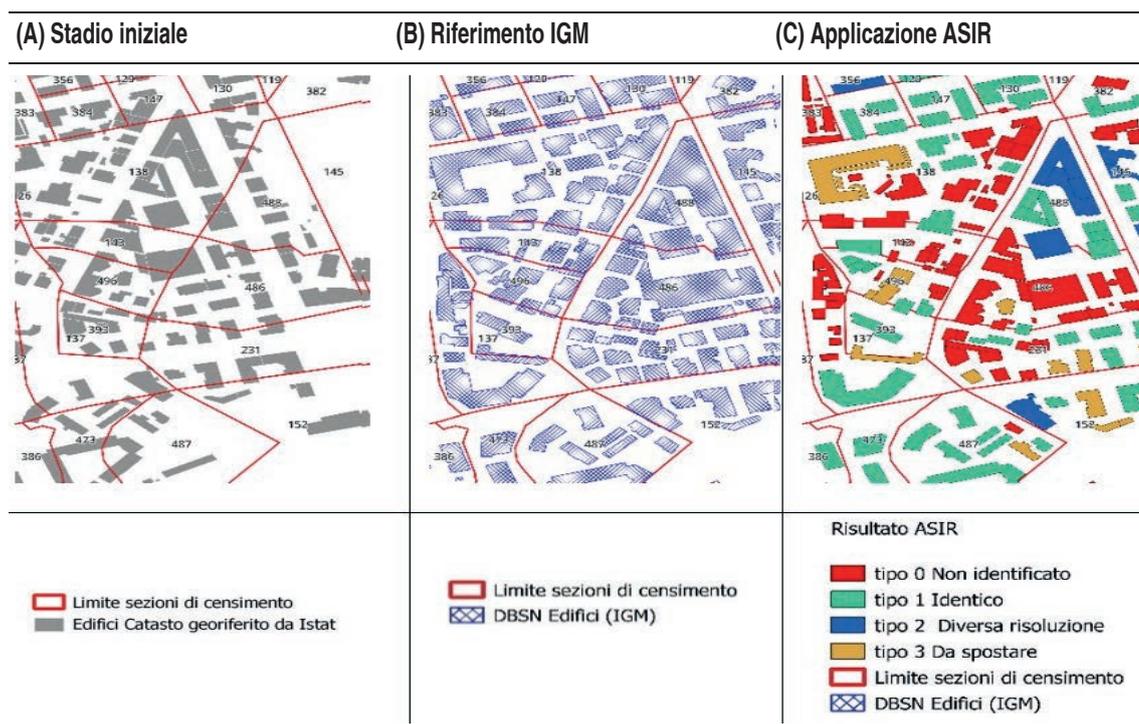
come *benchmark*. A partire da questi due geodatabase si costruisce la matrice degli edifici elegibili per coppie di edifici simili, entro una distanza prestabilita di 500 metri. Per migliorare il posizionamento del foglio catastale si sposta l'insieme degli edifici catastali sulla base dei vettori delle coppie di coordinate dei centroidi degli edifici riconosciuti come simili. Reiterando il processo, la valutazione dell'ASI per ogni foglio di mappa spostato consente di individuare il miglior posizionamento di quest'ultimo.

La procedura di riposizionamento ASIR, attraverso il calcolo del centroide e l'applicazione di criteri di similarità geometrica consente, quindi, un miglioramento significativo nella qualità del posizionamento dei dati catastali. La metodologia adottata, combinata con la valutazione ASI e l'uso di misure sintetiche di qualità, garantisce un processo maggiormente accurato e affidabile. Solo con ASIR è possibile georeferenziare al meglio il catasto, migliorando la precisione e l'affidabilità dei dati spaziali essenziali per la gestione e la lettura dei fenomeni sul territorio.

8.2.3 L'integrazione con altre fonti per la geocodifica alla sezione di Censimento e l'arricchimento della base informativa

L'integrazione dei dati geografici si fonda sul corretto posizionamento degli stessi: se il georiferimento degli elementi non è coerente nei database geografici messi a confronto, l'integrazione può risultare errata. Ad esempio in Figura 8.1 è riportato un caso in cui allo stato iniziale (A) i poligoni del catasto georiferiti da Istat non rispettano il tracciato delle sezioni di Censimento 2021, a differenza di quanto accade con i poligoni degli edifici del DBSN (B). Il modello geografico ASIR riesce a riconoscere e spostare gli edifici catastali in una posizione corretta (C). Questo garantisce una più esatta geocodifica degli elementi geografici.

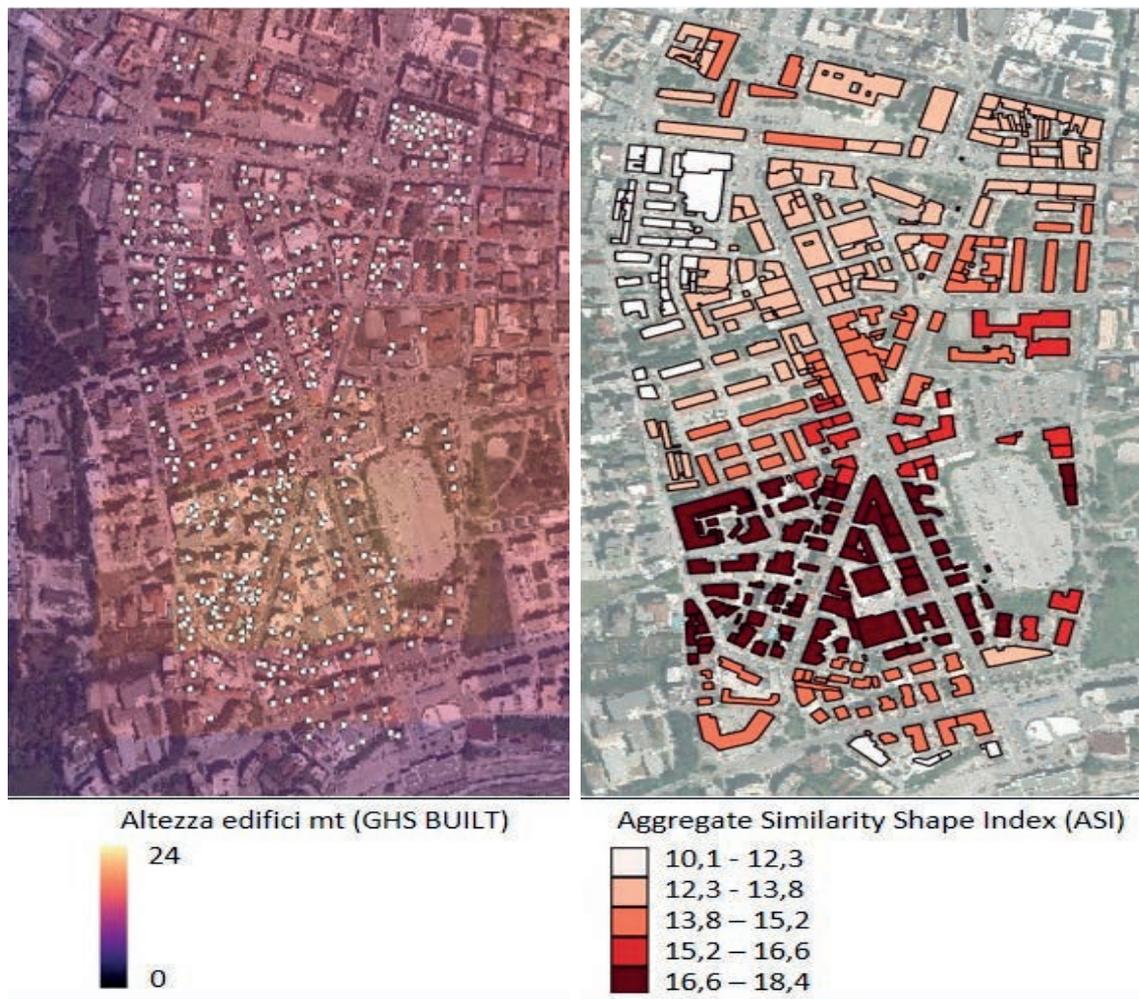
Figura 8.1 – Individuazione della sezione di Censimento e livello di qualità dell'informazione rilevata per il Comune di Cassino



Fonte: Elaborazioni su dati Istat e Istituto geografico militare (IGM)

Una volta migliorato il posizionamento degli edifici sul territorio, la possibilità di integrare altre informazioni presenti in altre fonti geografiche diventa vastissima. Ad esempio, è possibile stimare l'altezza dell'edificio e, quindi, il numero di piani fuori terra a partire dai dati stimati dal *Joint Research Centre* (JRC) della Commissione europea grazie alle informazioni ricavabili dal centroide dell'edificio riposizionato (Figura 8.2).

Figura 8.2 - Estrazione dell'altezza dell'edificio in base al centroide riposizionato con ASIR e assegnazione al poligono



Fonte: Elaborazioni su dati Istat e Joint Research Centre (JRC)

8.3 Abitazioni e altre unità immobiliari: elementi di integrazione con il Registro di popolazione e il Registro delle unità economiche

Come si è detto, all'interno del Sistema integrato dei registri, il Registro statistico dei luoghi (RSBL) fornisce dati statistici geografici a complemento delle informazioni statistiche provenienti da altri registri (sociodemografici o economici).

In questa Sezione si descrivono i principali elementi utilizzati, i problemi incontrati e le soluzioni adottate nel collegamento famiglie-abitazioni tra la popolazione residente nel Censimento della Popolazione e delle abitazioni 2021 e il Registro degli edifici e delle abitazioni (RSBL). Lo stesso approccio è in via di sperimentazione per considerare il registro ASIA Unità locali.

8. Registro edifici, abitazioni e altre unità immobiliari: il processo di integrazione e localizzazione dei dati amministrativi

Nel Censimento tradizionale, in cui le informazioni sulle famiglie venivano raccolte tramite specifici modelli di indagine consegnati alle famiglie e poi recuperati una volta compilati, il legame famiglia-abitazione era contestuale poiché le informazioni di entrambi erano contenute nello stesso questionario. Nel nuovo sistema integrato di registri, le unità statistiche di interesse provengono da archivi amministrativi diversi e il collegamento tra le varie unità deve essere stabilito attraverso codici di identificazione, quando disponibili e sempre in piena conformità con le normative sulla protezione dei dati identificativi personali e informazioni aggiuntive.

Da ottobre 2018, l'Istat ha avviato il Censimento permanente della Popolazione e delle abitazioni basato sulla integrazione delle informazioni disponibili dalle fonti amministrative con quelle acquisite da indagini campionarie rotative condotte in tutti i comuni italiani (Istat 2022). Per il nuovo Censimento permanente, l'archivio amministrativo di riferimento principale è l'Archivio nazionale della popolazione residente (ANPR), che raccoglie e centralizza i contenuti delle liste anagrafiche comunali; arricchito e corretto dai risultati delle indagini campionarie annuali menzionate sul territorio (*Master Sample*), l'Archivio integrato degli abitanti abituali in Italia (AIDA) e le informazioni rese disponibili dal Registro statistico di base delle persone, delle famiglie e dei conviventi (RBI).

L'elemento geografico che caratterizza gli individui e le famiglie censite è l'indirizzo di residenza⁴. Ogni indirizzo è codificato tramite un Codice univoco di indirizzo (CUI) (cfr. Capitolo 7). Per molti indirizzi sono disponibili anche le coordinate geografiche, corredate da un indicatore di qualità.

Come detto in precedenza, le unità immobiliari in RSBL provengono dal Catasto edilizio urbano, integrato con le principali informazioni del 15° Censimento generale della Popolazione e delle abitazioni del 2011 e con gli archivi dei contratti di locazione immobiliare. Nello specifico le abitazioni sono individuate tramite la categoria catastale da A01 a A11 a esclusione della categoria A10 (uffici)⁵. In questo caso gli elementi geografici sono il codice catastale del comune e i riferimenti di foglio e particella, ma si può arrivare anche al subalterno, associati alla singola proprietà. Nell'archivio catastale, inoltre, gli immobili sono associati a uno o più indirizzi. Anche questi indirizzi sono acquisiti in RSBL e a ciascuno di essi è associato un CUI. Per ogni immobile sono, infine, disponibili informazioni sulle persone fisiche e giuridiche titolari di diritti, il tipo di diritto e le rispettive quote di proprietà e di titolarità di altri diritti reali, nonché informazioni sulla locazione degli stessi e sulle persone fisiche o giuridiche locatarie.

Per la maggior parte delle abitazioni è possibile avere una georeferenziazione puntuale attraverso la coordinata dell'edificio in cui ricadono; questo è possibile se gli estremi catastali dell'immobile nel Catasto urbano coincidono con quelli dell'edificio cartografato nel Catasto terreni (cfr. paragrafo precedente). Attualmente, nel registro il 72,45 per cento delle abitazioni si collega direttamente a un poligono della cartografia tramite estremi catastali. Questa percentuale sale all' 89,43 per cento dopo avere corretto la chiave di link, quando possibile. In alternativa la georeferenziazione avviene tramite le coordinate degli indirizzi associati all'unità immobiliare.

4 L'indirizzo è una stringa che comprende il tipo di identificatore dell'area di circolazione (Via, Viale, Vicolo, Piazza, eccetera), il nome ufficiale assegnato all'area di circolazione (ad esempio Marco Polo, Giulio Cesare), il numero civico e l'eventuale esponente.

5 Ovviamente il registro accoglie anche tutti gli altri immobili diversi da abitazione; in questa accezione più allargata parliamo di Registro degli edifici e unità immobiliari.

Alla luce di tutto quanto detto, le principali informazioni presenti in RSBL o negli archivi a essi collegati, utili per l'integrazione con altri registri e altre unità statistiche sono:

- le coordinate degli edifici posizionati sulla mappa;
- gli estremi catastali degli immobili e degli edifici;
- le coordinate degli indirizzi degli immobili;
- il codice univoco di indirizzo (CUI) e le coordinate di tutti gli indirizzi considerati;
- le persone e le famiglie residenti agli indirizzi residenziali (nel caso di link con il Registro di popolazione);
- le unità economiche, comprensive di informazioni sugli indirizzi e sulle persone, fisiche e/o giuridiche, che ne hanno la titolarità (nel caso di link con il Registro delle unità economiche);
- le persone (fisiche e giuridiche) titolari di diritti reali sugli immobili;
- le persone (fisiche e giuridiche) e gli estremi catastali degli immobili interessati da contratti di locazione;
- la geocodifica alle sezioni di Censimento (2021, ma anche 2011)⁶ di tutte le unità considerate (persone, abitazioni, indirizzi, edifici, unità economiche).

Naturalmente, non sempre si dispone di tutte le informazioni sopra riportate in modo completo.

Per l'integrazione della popolazione residente nelle abitazioni e la successiva geocodifica delle unità alla sezione di Censimento, integrazione RBI-RSBL, è stata adottata una logica di massima valorizzazione delle informazioni disponibili e un principio di sussidiarietà tra queste. Quando una informazione è mancante, ad esempio non si dispone delle coordinate di un edificio, si utilizzano le coordinate degli indirizzi a esso collegate e viceversa. In pratica, si è cercato di ricostruire per ciascuna unità, una informazione congruente per tutti i livelli dei registri interessati. Tutte le persone della stessa famiglia devono avere lo stesso indirizzo di residenza, essere allocati nella stessa abitazione, all'interno di un unico edificio, geocodificato a una unica sezione di Censimento.

Per gli scopi censuari, che comprendono anche il calcolo e la diffusione di indicatori abitativi, le famiglie residenti in abitazione devono essere univocamente collocate in una abitazione.

Dal punto di vista teorico si tratta di un problema di collegamento deterministico dei *record* basato su precise variabili di corrispondenza (Scanu 2003; D'Orazio *et al.* 2006). Le due principali variabili di link sono l'indirizzo (CUI) della famiglia e dell'abitazione, e l'individuo (*CODICE_INDIVIDUO*) residente e titolare di diritti reali.

La procedura di link collega le famiglie alle abitazioni contemporaneamente tramite l'indirizzo e le informazioni sulla titolarità degli immobili. Per ogni CUI di residenza, vengono considerate tutte le abitazioni associate, formando così gruppi di n famiglie e m abitazioni caratterizzati da diverse qualità del collegamento dell'indirizzo, che vanno dal migliore livello (identità fino al numero civico) al peggiore livello (abitazione e indirizzo all'interno dello stesso comune). All'interno di ogni gruppo, vengono identificati, se presenti, i casi nei quali la famiglia possiede o affitta anche una delle abitazioni a quell'indirizzo.

⁶ La geocodifica alle sezioni di Censimento, intesa come passaggio finale della procedura finalizzata alla diffusione dei dati, è ovviamente la geocodifica alle sezioni 2021. Ciononostante, per le attività di individuazione e selezione di gruppi di abitazioni e famiglie, più circoscritti rispetto agli stessi insiemi calcolati per l'intero comune, è stata usata anche l'informazione relativa alla geocodifica delle unità alle sezioni di Censimento 2011.

8. Registro edifici, abitazioni e altre unità immobiliari: il processo di integrazione e localizzazione dei dati amministrativi

67

Tavola 8.1 - Famiglie in abitazione per qualità del link tramite indirizzo e titolarità sugli immobili. Anno 2021 (valori percentuali sul totale delle famiglie in abitazione) (a)

Qualità link indirizzo (b)	Proprietari	Affittuari	Altro	Totale
Nello stesso civico	44,00	3,78	6,41	54,19
In un civico adiacente (c)	16,43	1,97	10,55	28,95
Nella stessa strada	1,81	0,18	-	1,99
Nella stessa sezione	6,99	1,00	0,90	8,89
Nel comune	2,33	0,55	3,10	5,98
TOTALE	71,56	7,48	20,96	100,00

Fonte: Istat, Censimento permanente della Popolazione e delle abitazioni

(a) Numero di famiglie in abitazione: 25.364.168.

(b) I TIPO_LINK considerati nella procedura sono più articolati rispetto ai cinque livelli riportati in tabella.

(c) In questo caso sono stati considerati civici adiacenti anche tutti gli indirizzi che conducono allo stesso edificio.

Tutti i casi non univoci all'interno di questa tipologia, ovvero tutti i casi nei quali gruppi di n famiglie residenti a un indirizzo non potevano essere allocate in maniera univoca ciascuna in una delle m abitazioni presenti allo stesso indirizzo, *linkage* n a m , sono stati trattati grazie a una seconda metodologia deterministica che risolve il problema di assegnazione tramite un algoritmo noto in letteratura come algoritmo ungherese in tempo polinomiale.

In estrema sintesi (per una descrizione accurata cfr. il Capitolo 9), per ciascuno dei gruppi di n famiglie *linkate* a m abitazioni, con $n \leq m$, l'algoritmo crea una matrice $n \times m$ nella quale, per ciascuna cella c_{nm} viene calcolato un punteggio per l'associazione tra la famiglia n -esima e l'abitazione m -esima. Il punteggio viene calcolato sulla base di variabili ausiliarie, tra le quali il numero di componenti della famiglia, il numero di vani dell'abitazione, il tipo e la quota di diritti sul bene. Successivamente lo stesso algoritmo stabilisce la combinazione univoca tra le famiglie e le abitazioni in modo tale che la somma di tutti i punteggi per quella combinazione sia massima rispetto a tutte le altre possibili.

L'applicazione di queste metodologie ha permesso di collocare l'intera popolazione residente in abitazione in modo univoco e nel migliore modo possibile sulla base delle informazioni disponibili.

Una volta terminato il link famiglie-abitazioni, è stato possibile procedere con le operazioni di geocodifica delle sezioni 2021, sulla base dell'indirizzo di residenza e dell'edificio nel quale le abitazioni risultano collocate.

Il punto di partenza di questa operazione è stata la geocodifica alle sezioni 2021, tramite *spatial join*, delle coordinate degli indirizzi residenziali e delle coordinate dei centroidi degli edifici, per verificare la congruenza del reciproco posizionamento. Il risultato dell'operazione geometrica è riportato nella Tavola 8.2:

Tavola 8.2 - Famiglie in abitazione per qualità della geocodifica alle sezioni di Censimento. Anno 2021 (valori percentuali sul totale delle famiglie in abitazione) (a)

Qualità geocodifica	Totale
Stessa sezione per indirizzo ed edificio	74,6
Presenza della sola sezione associata all'edificio	3,8
Presenza della sola sezione associata all'indirizzo	10,6
Sezione differente per indirizzo ed edificio	10,6
Assenza di sezione di Censimento sia per indirizzo sia per edificio.	0,4
TOTALE	100,0

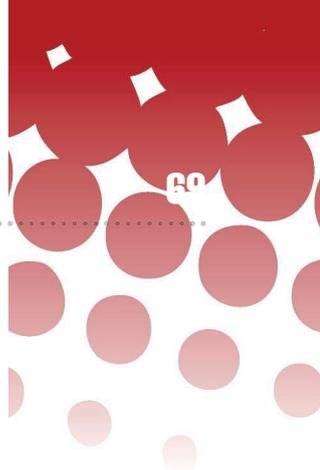
Fonte: Istat, Censimento permanente della Popolazione e delle abitazioni

(a) Numero di famiglie in abitazione: 25.364.168.



A partire da queste fattispecie si sono poi analizzati i casi che presentavano una incongruenza nella geocodifica o una assenza della stessa per riportare a coerenza e completezza l'intera tabella di lavoro. Si è proceduto con l'aiuto delle informazioni ausiliarie disponibili (ad esempio la distanza tra l'edificio e l'indirizzo, la geocodifica alle sezioni 2011, la qualità delle coordinate, eccetera), seguendo il criterio della massima valorizzazione delle informazioni e il principio di sussidiarietà.

Il risultato ottenuto è stato poi sottoposto alle procedure statistiche di controllo e correzione per la definitiva validazione e preparazione dei dati per la diffusione.



9. RSBL NEL SISTEMA INTEGRATO DEI REGISTRI¹

In questo Capitolo si descrive come il Registro statistico di base dei luoghi (RSBL) si integra con gli altri registri e, in particolare, come il suo ruolo sia centrale nell'ambito del Sistema integrato dei registri (SIR).

A prescindere dalla specifica unità statistica di un dato registro base, RSBL è fondamentale per collocare sul territorio tutte le unità statistiche del SIR e, quindi, funge da ponte con gli altri registri per realizzare statistiche sul territorio.

La georeferenziazione è realizzata attraverso delle metodologie generalizzate che effettuano l'integrazione, su base geospaziale, con diversi livelli di approssimazione ma anche a livello di codifica amministrativa.

L'intrinseca peculiarità di ogni unità statistica non altera i livelli di approssimazione geospaziale, ma può influire sulle caratteristiche degli oggetti di RSBL da integrare. Ad esempio, una famiglia risiede in una abitazione e una unità locale commerciale è presente in un negozio.

9.1 Metodologia di integrazione tra i registri RSBL e RBI

Il seguente paragrafo descrive il processo di integrazione (*linkage*) tra RSBL e il Registro di base degli individui (RBI) nelle sue unità statistiche individuo e famiglia. L'obiettivo della integrazione è di associare in modo univoco le famiglie residenti al territorio amministrativo, all'indirizzo di residenza e all'interno delle unità abitative georiferite e geocodificate.

Gli strumenti a disposizione per l'integrazione sono le unità immobiliari e gli edifici con i loro indirizzi, gli individui e le famiglie residenti con il loro indirizzi e, infine, gli individui titolari di proprietà e locazione di una unità immobiliare.

Il processo di *linkage* per la costruzione dell'associazione univoca di famiglie a indirizzo di residenza viene dettagliato nel paragrafo 9.2. Tale processo è basato sull'applicazione di diversi livelli di *linkage* (livelli di approssimazione geografica) tra diverse tipologie di famiglie ed è caratterizzato da due metodologie deterministiche.

In questo caso si considera una metodologia di associazione che per costruzione è univocamente deterministica, in quanto colloca i titolari e i non titolari nell'unico immobile disponibile nel migliore livello di link possibile e, successivamente, una seconda metodologia deterministica che elabora tutte le famiglie e gli immobili che non sono stati risolti dalla prima metodologia.

Questa seconda parte dalla costruzione di una matrice di associazioni tra n famiglie e m immobili, considerando tutti gli immobili disponibili alla residenza. Per tale ragione alla stessa residenza possiamo avere più famiglie e più immobili.

¹ Il Capitolo è stato redatto da Enrico Orsini (paragrafi 9.4 e 9.6), Roberta Radini (paragrafi 9.1 e 9.2), Andrea Pagano (paragrafo 9.3), Armando D'Aniello (paragrafo 9.5).

9.2 Integrazione tra i registri RSBL e RBI

RSBL, come detto nei paragrafi precedenti, contiene tutte le componenti descrittive delle unità: amministrative (esempio: comuni, regioni); funzionali (esempio: sistemi locali del lavoro); partizioni subcomunali (sezioni di Censimento, griglie regolari, eccetera); immobiliari (edifici e abitazioni); indirizzi. Il registro tratta la dinamica nel tempo delle singole unità territoriali e la relazione tra di esse.

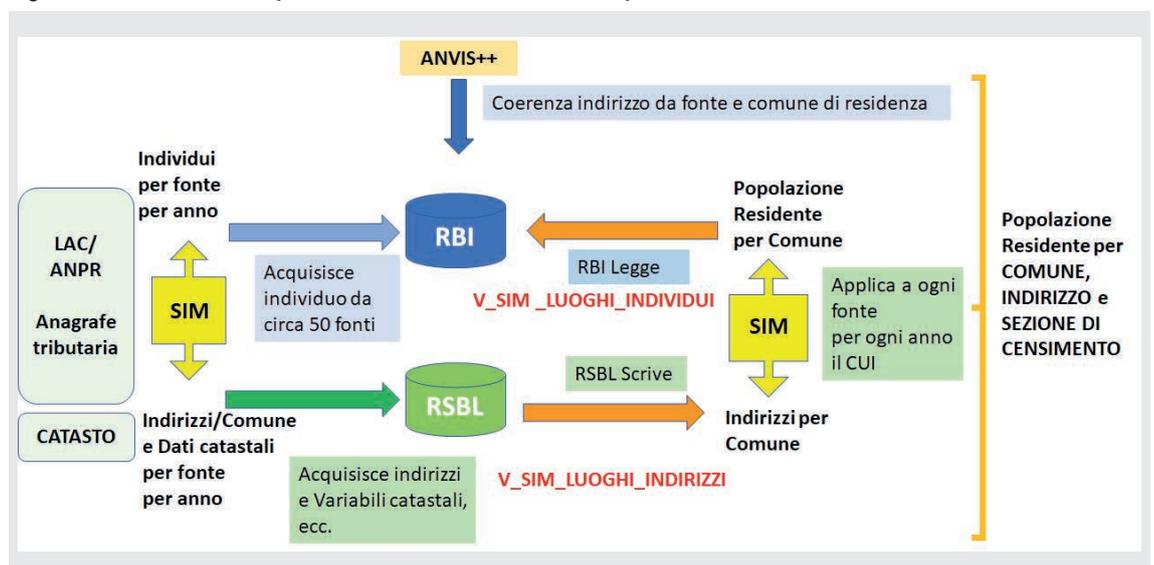
Per la particolarità delle unità trattate, il registro svolge anche la funzione di identificazione delle sue unità statistiche. Quindi, a differenza del Registro degli individui e del Registro delle unità economiche (ASIA), dove l'identificazione dell'individuo e della unità giuridica viene svolta dal processo del Sistema integrato dei microdati (SIM), in questo caso il registro svolge anche la funzione di identificazione delle unità territoriali. In particolare, nella fase di acquisizione, il processo SIM attribuisce agli indirizzi individuati nelle fonti amministrative un identificativo interno alla fonte.

La lista degli indirizzi, così individuata dal processo Sistema integrato luoghi e indirizzi (SILI), viene inviata al sistema di gestione del registro RSBL, componente indirizzi. Gli indirizzi di fonte amministrativa che vengono riconosciuti e identificati nel Registro dei luoghi, vengono restituiti al SIM con l'identificativo unico validato (CUI) e un indicatore di qualità che rappresenta il livello di riconoscimento dell'indirizzo (Cfr. Capitolo 7). Per quelli non riconosciuti, e, quindi, non identificati (definiti *no linked*), vengono restituiti al SIM senza attribuzione di un identificativo.

Tramite questi archivi il SIM appone il CUI a tutte le unità ivi presenti e le distribuisce a tutte le fonti da cui è stata estratta la stringa indirizzo. In questo ultimo caso l'indice di qualità definisce il livello di affidabilità del dato e, quindi, lascia la facoltà di decidere al processo di produzione il tipo di uso in funzione del livello di affidabilità dell'informazione. Soltanto tramite questo meccanismo, ad esempio, è possibile attribuire un identificativo di indirizzo di residenza in RBI.

Il processo di acquisizione del CUI da parte di RBI si può rappresentare nel flusso come in Figura 9.1.

Figura 9.1 - Processo di acquisizione dell'identificativo CUI nel processo di definizione di RBI



9. RSBL nel sistema integrato dei Registri

- Tale processo è suddiviso nei macropassi da svolgersi annualmente di seguito elencati.
- Per le fonti interessate dal processo SILI del SIM, vengono analizzate le stringhe indirizzo grezzo ed eseguite in alternativa le seguenti attività:
 - riapplicazione del CUI alla stringa se già individuato negli anni precedenti;
 - inserimento di una nuova riga relativa alla fonte e anno con il CUI non valorizzato e queste informazioni vengono passate al processo di identificazione di RSBL.
 - Per le fonti interessate dal processo di identificazione e caricamento di RSBL:
 - viene trattata la lista degli indirizzi privi di CUI fornita da SILI;
 - viene eseguito il processo di identificazione o di acquisizione di un nuovo indirizzo;
 - viene restituita al sistema SILI la lista degli indirizzi privi di CUI, elaborati e con il CUI valorizzato (se identificato) e il relativo indicatore di qualità;
 - viene inviata la lista degli indirizzi già identificati con un CUI, ma per i quali è aumentata la qualità.
 - Per le fonti interessate il processo SILI del SIM:
 - si applica il CUI alla stringa indirizzo nella vista che associa il CUI all'indirizzo.
 - Per le fonti interessate dal processo di attribuzione del CUI di residenza utilizzate da RBI:
 - viene estratto dalle fonti LAC (ANPR) e Anagrafe tributaria l'associazione CUI – *CODICE_INDIVIDUO* e qualità del CUI;
 - viene attribuito l'indirizzo di residenza alla popolazione del registro in funzione del tipo di fonte e della qualità del CUI. Questa attività consente di definire l'associazione CUI - individuo;
 - viene effettuata l'analisi della coerenza del territorio di riferimento del CUI e la sua geocodifica con l'attribuzione del comune di residenza dell'individuo, determinando la conferma dell'attribuzione dell'indirizzo di residenza (CUI) all'individuo.

Questa strategia a doppia fase di identificazione e di *linkage* delle unità statistiche dei registri con il territorio, consente agli altri registri di procedere nei rispettivi processi di alimentazione ed elaborazione, di effettuare un allineamento finale degli identificativi territoriali, che è possibile soltanto quando il registro RSBL ha terminato tutte le operazioni di identificazione e di *linkage* delle sue unità statistiche. I tempi di lavorazione e di consolidamento dell'identificazione dell'indirizzo e del territorio devono essere congrui con i tempi di produzione di tutti i registri. Questo aspetto risulta critico rispetto ai tempi di disponibilità del registro RSBL e dei diversi prodotti statistici dei registri. Per questo motivo l'allineamento avviene tramite un processo che si compone di più *step*. Tipicamente l'adeguamento per territorio amministrativo e indirizzo è molto veloce e, quindi, può essere fatto in prima istanza e, a seguire e se necessario, si possono integrare anche le informazioni di immobili ed edifici o in generale di informazioni catastali.

In generale, tutti i registri che si integrano con RSBL conservano le chiavi di integrazione con le unità statistiche di RSBL, come: il Codice unico di indirizzo (CUI); il codice di comune; il codice di unità immobiliare (UIU). Il dettaglio di queste unità è archiviato esclusivamente in RSBL.

9.3 Definizione dei livelli di link e dei gruppi di famiglie residenti

Le famiglie residenti vengono distinte in tre diverse tipologie identificate dalla presenza o meno di uno o più titolari di abitazioni all'interno del nucleo familiare (Prospetto 9.1).



Prospetto 9.1 - Classificazione in gruppi delle famiglie

Gruppo	Descrizione
1	Famiglia con almeno un componente proprietario di unità abitativa
2	Famiglia con almeno un componente affittuario di unità abitativa
3	Famiglia con tutti i componenti non titolari di alcuna unità abitativa

Fonte: Istat, Sistema integrato dei registri

La qualità del *linkage* famiglia e unità immobiliare è definita secondo una scala di dodici livelli, che varia a seconda della prossimità tra l'indirizzo di residenza e l'indirizzo dell'immobile (Prospetto 9.2).

Nel caso dei titolari (proprietari o locatari di un immobile), il vincolo tra la famiglia e l'immobile è molto forte per cui si possono utilizzare tutti i livelli di *linkage*.

Per il gruppo dei non titolari (né proprietari né locatari di alcun immobile), le uniche componenti considerate attendibili sono l'indirizzo e la strada, per questo vengono applicati solo i primi quattro livelli di *linkage*.

Prospetto 9.2 - Definizione dei diversi livelli di link in ordine di qualità

Tipo link	Descrizione
1	Famiglia con indirizzo di residenza coincidente all'indirizzo dell'abitazione
7	Famiglia con indirizzo di residenza coincidente a uno degli indirizzi dell'edificio
9	Famiglia con indirizzo di residenza contiguo all'indirizzo dell'abitazione (distanza massima 8 civici)
11	Famiglia con indirizzo di residenza contiguo a uno degli indirizzi dell'edificio (distanza massima 8 civici)
2	Famiglia con indirizzo di residenza non coincidente all'indirizzo dell'abitazione ma coincidente alla strada all'interno della stessa sezione di Censimento
8	Famiglia con indirizzo di residenza non coincidente all'indirizzo dell'abitazione ma coincidente a una delle strade dell'edificio all'interno della stessa sezione di Censimento
3	Famiglia con strada di residenza non coincidente alla strada dell'abitazione ma all'interno della stessa sezione di Censimento
4	Famiglia senza indirizzo di residenza o abitazioni senza indirizzo
5	Famiglia senza indirizzo di residenza e abitazioni senza indirizzo
10	Famiglia con indirizzo di residenza non coincidente all'indirizzo dell'abitazione ma coincidente alla strada in diverse sezioni di Censimento
6	Famiglia con strada di residenza non coincidente alla strada dell'abitazione in diverse sezioni di Censimento e con distanza massima di 200 metri
12	Famiglia con indirizzo di residenza nel comune dell'abitazione

Fonte: Istat, Sistema integrato dei registri

9.4 Prima metodologia deterministica

La prima metodologia deterministica è il processo di associazione più forte tra famiglia e immobile, ed è basato su due vincoli molto stringenti.

Il primo vincolo consiste nel considerare dapprima le famiglie titolari (gruppo 1 e 2) di unità abitative e poi le famiglie non titolari.

Il secondo vincolo consiste nella univocità delle associazioni, ovvero le famiglie titolari possiedono un unico immobile e per le famiglie non titolari esiste un unico immobile libero non occupato dai titolari.

L'associazione è costruita attraverso un algoritmo che verifica l'esistenza della condizione dei due vincoli, applicata alla migliore qualità del link possibile. Per i titolari, quindi, sono elaborate tutte le diverse tipologie di *linkage*, mentre per i non titolari solo le migliori quattro (cfr. paragrafo 9.3). Le famiglie associate tramite una determinata tipologia di *linkage* sono escluse dalle associazioni con le successive tipologie di link.

Ogni abbinamento prodotto può avere l'effetto di creare nuove possibili associazioni: tutte le tipologie di *linkage* sono reiterate fino a quando il risultato degli abbinamenti è esaurito.

9. RSBL nel sistema integrato dei Registri

Il risultato complessivo di questa prima metodologia di associazione è un primo output di famiglie associate a un'unica abitazione.

Questo output è poi ulteriormente sottoclassificato nell'insieme di una famiglia e un'abitazione (insieme 1 a 1) e nell'insieme di più famiglie e un'abitazione (insieme n a 1), dove quest'ultimo definisce la quota di famiglie coabitanti.

Le restanti famiglie non abbinate, sono le famiglie per le quali non sono verificati i vincoli sopra descritti, ovvero non esistono delle condizioni univoche di link ma all'opposto, da un punto di vista generale, costituiscono tanti insiemi di n famiglie associate a m immobili e si definiscono insieme k . La risoluzione di questo sottoinsieme è realizzata attraverso la seconda metodologia deterministica.

9.5 Calcolo della qualità delle associazioni famiglia-abitazione

Per la seconda metodologia deterministica è stato elaborato uno schema di punteggi da assegnare alle coppie famiglia-immobile, prevedendo come punteggio minimo zero e come punteggio massimo due per indicare la massima qualità di associazione.

Lo schema è basato su quattro variabili alle quali viene attribuito un diverso peso:

- combinazione di tipo link, qualità della coordinata dell'indirizzo di residenza e, nel caso questa risulti presente, la distanza tra la detta coordinata e quella del centroide dell'edificio dell'immobile. In particolare, per la distanza, si considera se questa sia o meno la minima tra tutte le coppie famiglia-immobile con stesso CUI di residenza (Tavola 9.1);
- la quota di possesso dell'immobile (Tavola 9.2);
- il rapporto tra il numero dei vani dell'immobile e il numero dei componenti del nucleo familiare (Tavola 9.3);
- il rapporto tra la superficie catastale dell'immobile (in m^2) e il numero dei componenti del nucleo familiare (Tavola 9.4).

Tavola 9.1 - Punteggi per tipo link, qualità della coordinata dell'indirizzo e distanza

Tipo link	Qualità della coordinata	Distanza	Punteggio
1	1 – 2	Min	0,75
1	3	Min	0,65
< > 1	1 – 2	Min	0,55
< > 1	3	Min	0,45
1	1 – 2	< > Min	0,65
1	3	< > Min	0,55
< > 1	1 – 2	< > Min	0,45
< > 1	3	< > Min	0,35
1	Non presente	-	0,45
< > 1	Non presente	-	0,25

Fonte: Elaborazioni su dati Istat, Sistema integrato dei registri

Tavola 9.2 - Punteggi per quota di proprietà dell'immobile

Suddivisione in classi della quota di proprietà dell'immobile	Punteggio
[0,75 – 1]	0,25
[0,50 – 0,75]	0,20
[0,25 – 0,50]	0,15
[0 – 0,25]	0,10

Fonte: Elaborazioni su dati Istat, Sistema integrato dei registri



Tavola 9.3 - Punteggi per rapporto tra i vani e i componenti del nucleo familiare

Suddivisione in classi del rapporto tra numero di vani e componenti del nucleo familiare	Punteggio
[1,3 – 1,7]	0,125
[1 – 1,3] - [1,7 – 2]	0,100
[2 – 3]	0,080
[3 – 5]	0,070
[5 – 7]	0,060
[0.5 – 1] - [7 – 10]	0,050

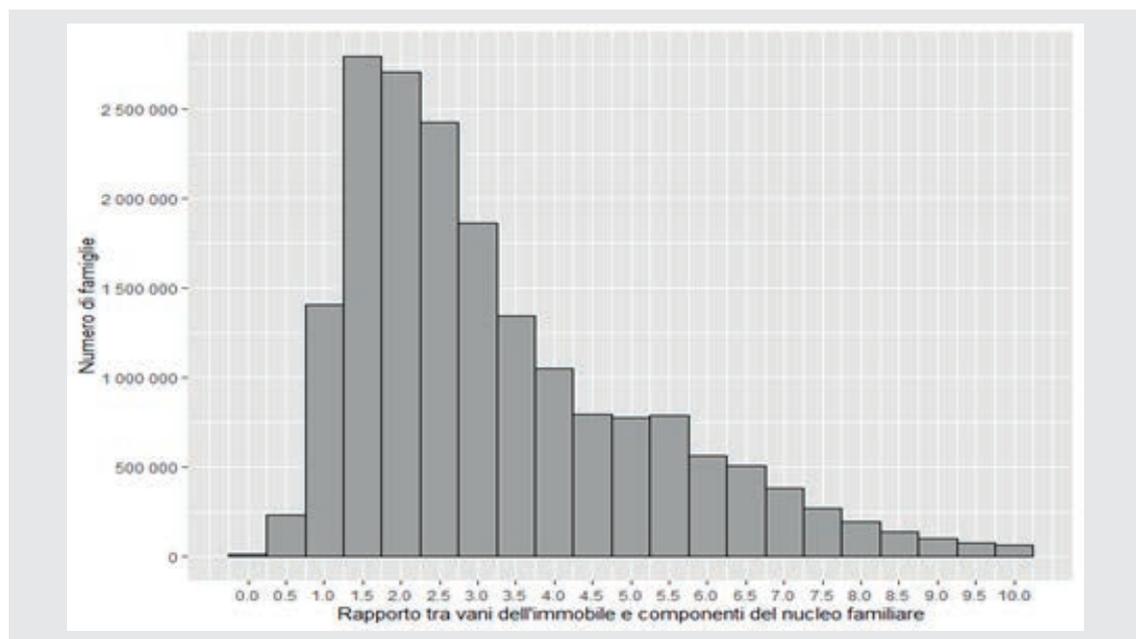
Fonte: Elaborazioni su dati Istat, Sistema integrato dei registri

Tavola 9.4 – Punteggi per rapporto tra la superficie dell'immobile e i componenti del nucleo familiare

Suddivisione in classi del rapporto tra la superficie dell'immobile e i componenti del nucleo familiare	Punteggio
[30 – 40]	0,125
[40 – 60]	0,100
[60 – 90]	0,080
[20 – 30] - [90 – 110]	0,075
[10 – 20] - [110 – 140]	0,050
[140 – 170]	0,025

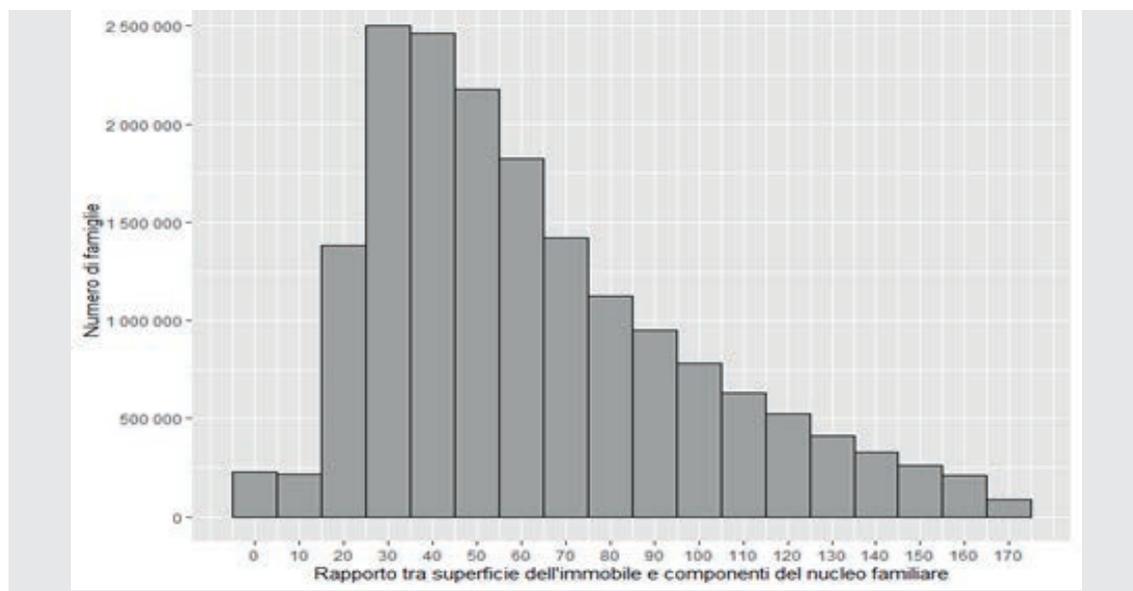
Fonte: Elaborazioni su dati Istat, Sistema integrato dei registri

Figura 9.2 - Rapporto tra i vani dell'immobile e i componenti del nucleo familiare per numero di famiglie in abitazione



Fonte: Elaborazioni su dati Istat, Sistema integrato dei registri

Figura 9.3 - Rapporto tra la superficie dell'immobile e i componenti del nucleo familiare per numero di famiglie in abitazione (a)



Fonte: Elaborazioni su dati Istat, Sistema integrato dei registri
(a) La superficie è quella catastale, in m².

Per la determinazione degli intervalli di punteggio per vani e superficie, è stata svolta un'analisi della distribuzione dei rapporti vani-componenti e superficie-componenti, nel caso delle coppie famiglia-immobile già validate nelle fasi precedenti del processo (Figura 9.2, Figura 9.3), assegnando punteggi maggiori alle casistiche più frequenti.

9.6 Seconda metodologia deterministica

La seconda metodologia deterministica è stata ideata per risolvere l'insieme k , applicando il criterio di ottenere le migliori associazioni univoche possibili.

Attribuiti i punteggi (cfr. paragrafo 9.5), è stato possibile applicare l'algoritmo ungherese, metodo di ottimizzazione combinatoria che risolve in tempo polinomiale il problema dell'assegnazione. Il metodo è stato sviluppato da Harold Kuhn nel 1955 (Kuhn 1955) e oggi è implementato da un pacchetto in R .

A partire da una matrice $n \times m$, determina le associazioni univoche, massimizzando la somma dei punteggi totali.

L'insieme k contiene tanti grafi bipartiti, ovvero tutte le possibili associazioni tra famiglie e immobili per una determinata tipologia di link (Figura 9.4).

L'applicazione dell'algoritmo ungherese, nel nostro specifico caso, è vincolata a due condizioni: la prima è che n deve essere minore uguale di m , ovvero devono esserci sufficienti abitazioni per le famiglie; la seconda è che i grafi bipartiti siano indipendenti tra loro, ovvero una famiglia e un immobile non possono essere in più grafi, altrimenti non si otterrebbe l'univocità.

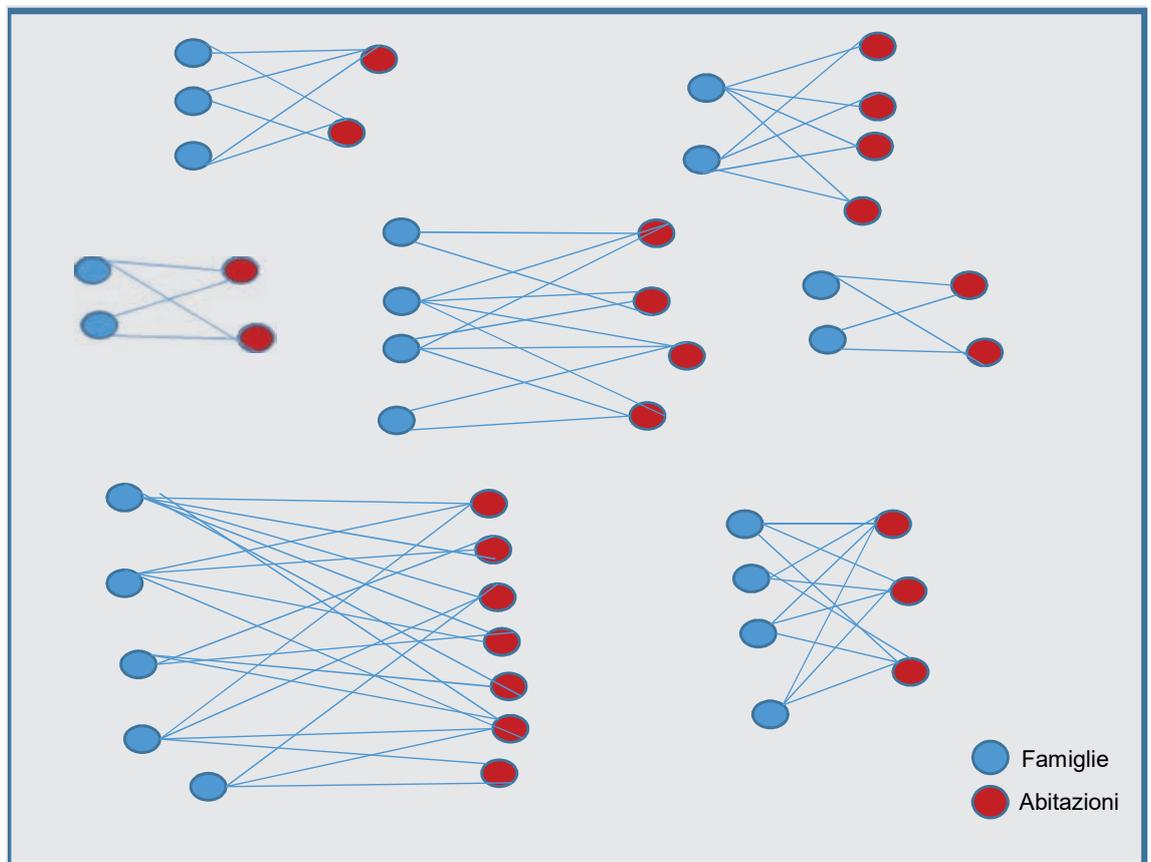
Per soddisfare tali condizioni sono stati aggiunti immobili fittizi con punteggio zero dove $m > n$ e sono stati costruiti dei nuovi macro grafi per accorpare tutti i grafi in correlazione per famiglia e immobile.

Grazie a questi accorgimenti si è potuto eseguire l'algoritmo su tutto l'insieme k e, successivamente, gli abbinamenti delle famiglie in immobili fittizi sono stati eliminati e hanno costituito un nuovo residuo.

Per esso è stato costruito un nuovo insieme $k1$ di grafi, ottenuto associando tutte le famiglie residenti in una determinata sezione di Censimento di residenza in tutte le abitazioni rimaste libere in quella sezione. Ricalcolato il punteggio secondo i criteri descritti, è stato riapplicato l'algoritmo ungherese.

Complessivamente questo secondo output ha permesso di completare l'associazione univoca tra famiglia e abitazione per il totale della popolazione residente italiana.

Figura 9.4 - Grafi bipartiti Famiglie - Abitazioni nell'insieme k



Fonte: Elaborazioni su dati Istat, Sistema integrato dei registri

10. RENDERE INTEROPERABILI I DATI GEOSPAZIALI DI RSBL¹

10.1 Il WebGIS e i GeoPortali

Gli strumenti GIS e WebGIS sono da tempo diventati fondamentali per la condivisione e l'analisi dei dati statistici georiferiti. Essi sono ormai indispensabili e imprescindibili per gli Istituti di statistica (INS) in tutte le fasi di trattamento dei dati statistici stessi a riferimento geografico. Il valore aggiunto derivante dall'integrazione dell'informazione geospaziale nella varie fasi del trattamento del dato statistico ha anche condotto alla definizione di *framework* e modelli di business sia a livello nazionale, sia internazionale. In particolare il *Global Statistical Geospatial Framework* (GSGF) (United Nations 2019) è una struttura di lavoro che, grazie ai suoi principi fondamentali, consente la produzione armonizzata e standardizzata di dati statistici georiferiti non ottenibili senza l'integrazione delle due tipologie di informazioni. In particolare, il principio 1 “*Use of fundamental geospatial infrastructure and geocoding*”, riguarda l'adozione di una infrastruttura geospaziale di base, comune e consistente, per georiferire ogni singola unità statistica. Potendo disporre di localizzazioni di riferimento di qualità (quali indirizzi, sezioni di Censimento, codici di edifici, eccetera) è possibile assegnare riferimenti geografici accurati (coordinate, codici di sezione, e/o edifici, eccetera) a ogni unità statistica a livello micro sul territorio. RSBL è perfettamente allineato a questo principio, come si evince anche dai Capitoli 5, 6, 7 e 8.

Sempre a dimostrazione dell'importanza di tale integrazione, in ambito Commissione economica per l'Europa (UNECE), si studia e si consolida l'estensione del GSBPM (Choi 2021) per integrare le funzionalità geospaziali e migliorare la produzione statistica.

In linea con i principi dell'apertura, del riuso e della condivisione dei dati esistono diverse direttive nazionali, internazionali ed europee per regolamentare e guidare la costituzione di *Spatial Data Infrastructure* (SDI) con l'obiettivo di consentire l'integrazione dell'informazione geospaziale, rendendola reperibile attraverso i GeoCataloghi e interoperabile attraverso l'adozione degli standard nazionali RNDT (Repertorio Nazionale dei Dati Territoriali) ed europei INSPIRE (*Infrastructure for Spatial Information in Europe*).

In Italia l'AGID (Agenzia per l'Italia Digitale) già da anni ha realizzato il GeoPortale RNDT (<https://geodati.gov.it>), un catalogo di geometadati che costituisce un punto di accesso unico all'informazione geografica, al quale le amministrazioni referenti devono contribuire per rendere reperibili i propri dati geospaziali e consentirne il riuso e l'interoperabilità.

RSBL, per la natura dei suoi output e delle indagini statistiche delle quali è alla base, è un perfetto esempio della integrazione tra informazione statistica e geospaziale. Esso ha, infatti, utilizzato e integrato, oltre i dati prodotti direttamente dall'Istituto (quali i confini amministrativi, le sezioni di Censimento, le località, eccetera), anche dati geospaziali di varie fonti, tra le quali cartografie disponibili online (sia di riferimento, sia tematiche, quali i database geotopografici pubblicati dalle regioni, la cartografia catastale dell'Agenzia delle entrate, eccetera), dati da EO (*Earth Observation*), immagini aeree, immagini di varia natura.

¹ Il Capitolo è stato redatto da Pina Grazia Ticca.



Le potenzialità dell'integrazione non riguardano la sola visualizzazione delle statistiche sulle mappe ma l'utilizzo di strumenti di analisi spaziale e calcolo di nuovi indicatori, proprio in virtù delle relazioni spaziali tra i fenomeni statistici georiferibili, basati sull'individuazione di relazioni tra i dati altrimenti non individuabili quali ad esempio, prossimità, distanza, inclusione, sovrapposizione, eccetera.

A oggi grazie alla tecnologia WebGIS è possibile interagire con i dati geospaziali attraverso i servizi di rete, realizzati secondo gli standard dell'*Open Geospatial Consortium* (OGC), che garantiscono l'interoperabilità dei dati geospaziali. Lo strumento di accesso ai dati è il GeoPortale, il quale rappresenta un punto di accesso unificato all'informazione geospaziale che ogni amministrazione referente dovrebbe offrire all'utenza. I GeoPortali vengono realizzati dalle differenti entità amministrative secondo le regole e direttive già citate (RNDT e INSPIRE) e offrono dei cataloghi di metadati geografici attraverso i quali consultare l'informazione richiesta. Molto spesso vengono anche offerti dei visualizzatori geografici con i quali consultare i dati attraverso diverse modalità, tra le quali mappe interattive e applicazioni specializzate secondo il tema di riferimento, oltre all'indicizzazione dei servizi di rete che l'utenza può utilizzare con strumenti GIS propri e integrarli con altri dati.

10.2 La piattaforma WebGIS dell'Istat

La piattaforma WebGIS (<https://gisportal.istat.it>) ha il compito di supportare l'Istat sia in risposta alle crescenti richieste di dati geografici e statistici georiferiti, con un livello di dettaglio informativo territoriale comunale e subcomunale (sezioni di Censimento e loro aggregazioni), sia alla crescente offerta di informazione statistica di qualità, la cui strategia di diffusione si basa sull'utilizzo di strumenti innovativi e soluzioni tecnologiche adeguate alla specificità dei contenuti, quali sono i dati geografici e statistici georiferiti.

In particolare, la piattaforma si adegua costantemente ai moderni canoni di condivisione, riutilizzo dei dati e non duplicazione, come sancito da norme e regolamenti nazionali ed europei (RNDT, INSPIRE, Open Data-HVD)

L'architettura viene aggiornata periodicamente per utilizzare al meglio le potenzialità offerte dalla tecnologia WebGIS a supporto, con l'obiettivo di realizzare il GeoPortale istituzionale, che deve costituire il punto di accesso unificato all'informazione geografica dell'Istat.

La piattaforma è dotata di una componente *portal* che consente dunque di pubblicare, in modo semplice e facilmente fruibile dall'utenza, dati statistici a base geografica. Costituisce una opportunità per incrementare la condivisione dei dati geografici e statistici georiferiti, nel quale le mappe sono l'elemento cardine dalle quali partire nell'analisi dei dati.

La piattaforma WebGIS dell'Istat consente di condividere su web sia le mappe, sia gli strumenti di analisi geografica attraverso la tecnologia *Web Map Service* e *Web Application GIS*.

La modalità di condivisione delle mappe e la tipologia dei servizi di rete utilizzati (i diversi tipi di *Web Map Service*) influenzano sia le funzionalità offerte, sia le prestazioni del sistema, consentendo di rispondere alla diversità degli utenti del web.

Attraverso il motore GIS è possibile configurare e incrementare sia i livelli geografici sia le funzionalità per l'interazione; spaziare dalla navigazione geografica, alle *query*, a complesse analisi spaziali, eccetera.

Le mappe condivise possono essere complessi servizi di rete interoperabili (*Web Map Service*) o semplici geografie, che un'applicazione *client* può elaborare modificando, per esempio, la classificazione dei valori e/o i colori degli indicatori statistici.

10. Rendere interoperabili i dati geospaziali di RSBL

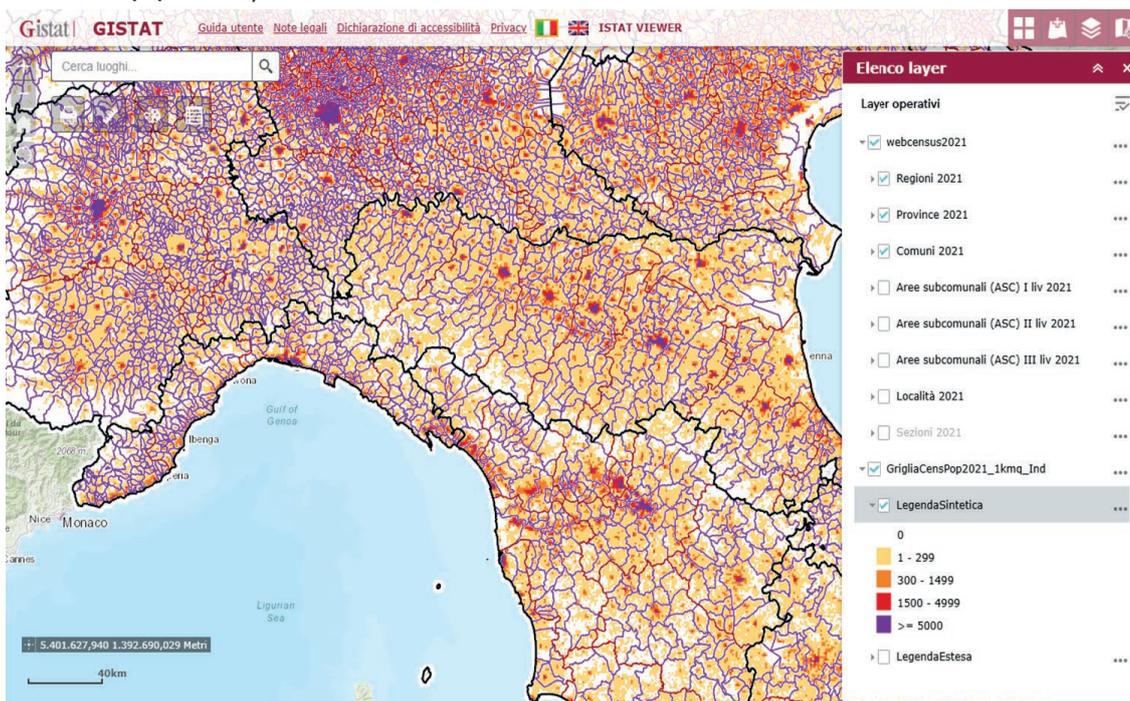
La piattaforma WebGIS dell'Istat offre, oltre a numerosi servizi di rete relativi a dati geografici e statistici georiferiti, documentati nel Repertorio Nazionale dei Dati Territoriali (<https://geodati.gov.it>) dell'AGID, anche alcune applicazioni che guidano l'utente nella consultazione e navigazione geografica di tali dati attraverso l'utilizzo di un *web browser*.

In particolare è disponibile *IstatViewer* (<https://gisportal.istat.it/IstatViewer>), un applicativo WebGIS dedicato alla visualizzazione e consultazione di livelli informativi di RSBL e, nello specifico, dei dati geografici delle basi territoriali e delle variabili censuarie per i Censimenti 2021, 2011, 2001 e 1991.

Per consentire la sovrapposizione di livelli informativi, in *IstatViewer* è possibile visualizzare insieme, per esempio, i dati delle basi territoriali e dei confini amministrativi e, tra gli altri, anche i dati di popolazione legale 2021, diffusi sulla griglia regolare, con celle di 1 km² (griglia europea rilasciata da Eurostat).

È possibile la navigazione geografica integrata di livelli geografici garantita dalla piattaforma WebGIS dell'Istat (Gisportal), ottenuta sovrapponendo i servizi di rete realizzati con i dati di RSBL (sezioni di Censimento e griglia di popolazione) (Figura 10.1).

Figura 10.1 - Esempio di sovrapposizione su web di livelli informativi di RSBL (sezioni di Censimento, griglia di popolazione) condivisi come servizi di rete



Fonte: Elaborazioni su dati Istat

11. IL PROCESSO TECNOLOGICO¹

11.1 L'industrializzazione

La costruzione dei registri è un tema complesso e sfidante per i vari settori dell'Istat coinvolti. In particolare RSBL è un registro molto strutturato (cfr. Capitolo 2), le sue componenti sono state sviluppate in modo incrementale, in tempi differenti, talvolta prendendo origine da esperienze pregresse che sono state arricchite e ampliate, in altri casi attraverso una progettazione ex novo di processi, database e interfacce a supporto.

Data la complessità dell'obiettivo che si voleva perseguire, la realizzazione è passata attraverso lo sviluppo iniziale di prototipi in via sperimentale sempre più consolidati che, in modo incrementale e iterativo, stanno passando a una gestione industrializzata, ossia controllata, monitorata, documentata, supportata da strumenti e interfacce, quanto più possibile automatizzata.

I primi passi sono una ingegnerizzazione, una strutturazione organica e una ottimizzazione del codice, inclusi i controlli e le verifiche effettuati manualmente nelle varie fasi di caricamento e produzione degli output. La scrittura di un codice robusto, con un'attenzione alle prestazioni e alle parametrizzazioni, consente di avere dei mattoncini che possono essere utilizzati nell'ordine necessario alla realizzazione del processo.

Fondamentale è poi l'analisi dei processi per individuare e descriverne fasi e sottofasi, associate alle relative procedure che le realizzano, propedeuticità e relazioni tra esse. Per la gestione dei processi sono state fatte delle sperimentazioni e delle analisi comparative per individuare gli strumenti adeguati (Casale *et al.* 2022; Vaste *et al.* 2021) per fornire un'automatizzazione che superasse il lancio manuale delle procedure e consentisse un monitoraggio semplice dello stato di esecuzione dei processi. Le principali funzionalità richieste a tali strumenti sono: organizzare l'esecuzione delle procedure esistenti attraverso l'utilizzo di interfacce web che consentano di visualizzare e monitorare in modo intuitivo e immediato l'andamento dei processi; automatizzare i processi attraverso la gestione del flusso, ossia prevedendo attività in sequenza e in parallelo, punti di decisione manuali o automatici attraverso regole di controllo predefinite, gestione degli errori, inserimento di parametri da interfaccia, eccetera; gestire la profilazione degli utenti e la relativa autorizzazione ai processi e funzionalità di competenza; inviare email per le notifiche in modo automatico.

All'interno del processo sono, inoltre, necessari momenti di analisi dei dati attraverso la visualizzazione di reportistica di output finali e parziali, di processo e di validazione, relativi a indicatori di qualità previsti nel processo, per valutare le varie fasi. Anche questa reportistica, in un'ottica industrializzata, deve essere fornita con strumenti idonei, semplici e al contempo *web-based*, condivisi e personalizzabili.

A seguito delle sperimentazioni effettuate, gli strumenti che si è deciso di adottare sono:

- *Flows for APEX* (<https://flowsforapex.org>) per il disegno, l'esecuzione e il monitoraggio dei processi;

¹ Il Capitolo è stato redatto da Paola Giorgetti (paragrafi 11.1, 11.2 e 11.3) e Giulia Vaste (paragrafo 11.1).

- *Oracle APEX* (<https://apex.oracle.com>) per la reportistica e, in futuro, la definizione di nuovi flussi o nuove parti per quelli già esistenti.

Flows for APEX si basa su BPMN (*Business Process Model and Notation* - <https://www.bpmn.org>) che è uno standard *de facto* per la modellazione dei processi. La potenza del BPMN risiede nell'essere una notazione grafica semplice e intuitiva, adatta anche a utenti non tecnici. Inoltre, si presta bene a una traduzione delle attività in software funzionante. Attraverso *Flows for APEX* è possibile:

- modellare in BPMN il proprio flusso, dettagliandolo in *task* automatici o manuali, disposti in sequenza o in parallelo, regolati da condizioni;
- associare agevolmente ai *task* automatici le procedure PL/SQL che li implementano;
- lanciare il processo e monitorarne l'avanzamento in modo intuitivo attraverso una interfaccia *user-friendly* che modifica i colori dei *task* via via che vengono eseguiti.

APEX è uno strumento *low code* per la realizzazione di applicazioni web e consente sviluppi rapidi attraverso la configurazione di funzionalità standard quali, ad esempio, la predisposizione di menù, la gestione utenti, la creazione e visualizzazione di reportistica, l'inserimento dati, la produzione di alcune tipologie di grafici, eccetera.

Entrambi gli strumenti sono attualmente utilizzati per la componente indirizzi di RSBL, come dettagliato nei paragrafi seguenti, e permettono all'utente di avere uno strumento completo per l'esecuzione, il monitoraggio e il controllo dei processi di caricamento, ma anche dei contenuti del registro stesso attraverso *report* dedicati. Per la componente edifici è stata svolta l'analisi del flusso di caricamento e si prevede l'adozione dei medesimi strumenti per la realizzazione dell'automatizzazione del flusso e per la reportistica. *APEX* è stato anche alla base della raccolta dei requisiti per la reportistica in SITUAS ed è stato utilizzato per l'implementazione di SITUAS-Servizi (cfr. Capitolo 4).

11.2 Il controllo di processo per la componente indirizzi

La componente indirizzi di RSBL è stata realizzata attraverso un primo popolamento massivo, e successivi caricamenti relativi ai soli nuovi indirizzi (delta), forniti dalle diverse fonti non ancora presenti nel registro (cfr. Capitolo 7). I processi di caricamento a delta sono stati raffinati negli anni e sono ormai consolidati: ciò ne ha permesso l'industrializzazione.

Il processo inizia con l'estrazione del delta dalle fonti originali. Gli indirizzi vengono, quindi, sottoposti a un processo di normalizzazione attraverso l'utilizzo di un software applicativo in grado di automatizzare il riconoscimento e la validazione degli indirizzi, basandosi su una importante base dati contenente tutti gli indirizzi esistenti in Italia, periodicamente aggiornata. Il software di normalizzazione può o meno riconoscere un indirizzo e, se lo riconosce, questo entrerà a fare parte del registro definitivo, altrimenti verrà inserito in una tabella di scarti. Questi ultimi verranno riproposti per la normalizzazione dopo una pulizia delle stringhe indirizzi, che a volte possono contenere caratteri sporchi come, ad esempio, il nome del comune, oppure a seguito di aggiornamenti della base dati del normalizzatore.

È stato necessario prevedere diversi modelli perché ciascuna fonte ha le sue peculiarità. Ad esempio, per il catasto è prevista una fase di verticalizzazione della riga di indirizzo, che può contenere fino a tre civici in modo che vengano generate, per una riga, un numero di righe pari al numero di civici presenti. Ciascun modello di caricamento è descritto nella base dati attraverso dei metadati che definiscono: le fonti di origine da leggere; le strutture dati di *working*; il flusso che descrive nel minimo dettaglio tutti i passaggi necessari per

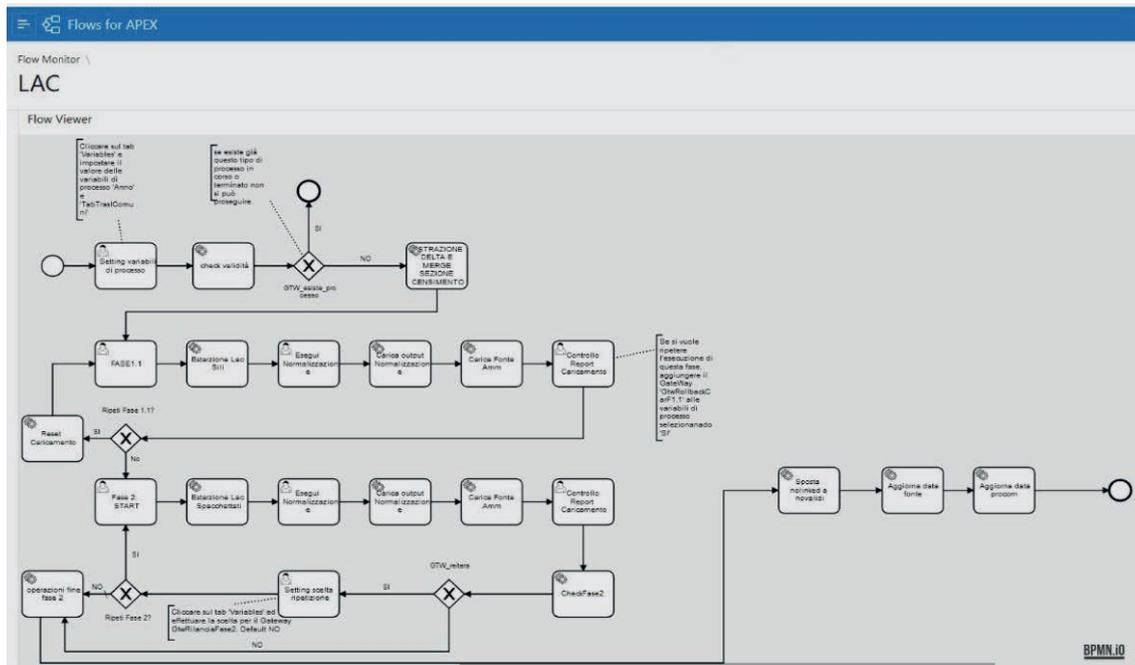
11. Il processo tecnologico

il caricamento completo degli indirizzi. Inoltre, nella base dati sono presenti tutte le procedure e le funzioni che devono essere eseguite nell'ordine descritto dal flusso e tutte le tabelle di *log* che permetteranno all'utente di tenere sotto controllo il processo attraverso la reportistica dedicata.

Per ogni fonte sono previsti più passaggi, o fasi, che costituiscono il flusso del processo di caricamento. In ogni fase viene raffinata la pulizia della stringa indirizzo, che potrebbe contenere caratteri non riconosciuti o errori sistematici, e ritentare la normalizzazione in modo da ottenere un riconoscimento.

La Figura 11.1 mostra un esempio di caricamento della fonte LAC, in cui è rappresentata la fase di estrazione del delta, la creazione del file con le stringhe indirizzo da sottoporre alla normalizzazione, il caricamento nella base dati della fonte amministrativa, uno *step* di pausa per la verifica che rimanda alla reportistica di controllo, descritta successivamente, e la reiterazione delle fasi di estrazione delle stringhe indirizzo e caricamento nella base dati, fino al termine del flusso che prevede aggiustamenti sulle date di riferimento della fonte e del territorio.

Figura 11.1 - Caricamento delle Liste Anagrafiche Comunali (LAC)



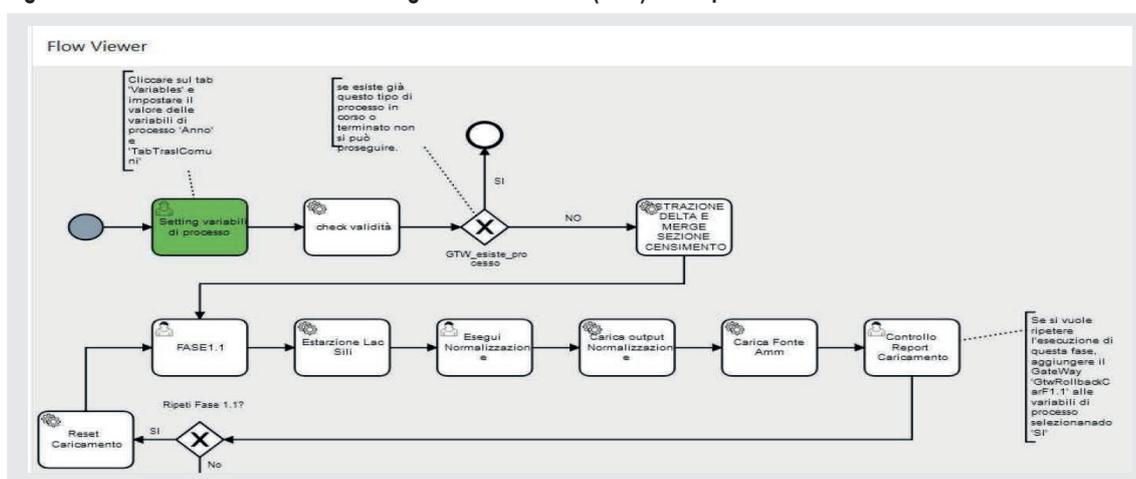
Fonte: Istat, Registro statistico di base dei luoghi - Indirizzi

Nella Figura 11.2, che mostra nel dettaglio una porzione del processo di caricamento della fonte LAC, i *task* con l'icona degli ingranaggi sono automatici e corrispondono a specifiche procedure del processo, mentre i *task* con l'icona dell'uomo sono manuali, ossia richiedono una interazione dell'utente. A questo punto si ha una pausa nell'esecuzione del flusso per consentire all'utente di eseguire le seguenti attività: eventuali verifiche e controlli attraverso la reportistica; valorizzare, laddove necessario e richiesto, le variabili di processo; mandare avanti il processo dopo avere effettuato una operazione manuale. Ad esempio, il *task* "esegui normalizzazione" prevede l'esecuzione, offline al momento, del software per normalizzare gli indirizzi, mentre i *task* "carica output normalizzatore" e "caricamento fonte amm" sono automatici e prevedono le chiamate alle procedure, presenti nel database, che

effettuano il caricamento del file prodotto dal normalizzatore, in una tabella di appoggio. L'elaborazione di tale tabella avviene in modo tale che vengano caricati gli indirizzi nelle opportune tabelle del registro.

Nella Figura 11.2 si mostra l'inizio del flusso rappresentato dal pallino (il colore grigio indica un'operazione già eseguita) e si attende l'avvio della *Fase 0* (il colore verde indica il *task* in esecuzione). La *Fase 0* è a carico dell'utente che inserirà le variabili di processo, tra le quali l'anno di elaborazione. I *task* di colore bianco devono ancora essere eseguiti e lo saranno secondo l'ordine sequenziale nel quale sono modellati. In particolare, i *task* automatici verranno eseguiti uno dopo l'altro senza alcun intervento manuale, mentre la *Fase 1* dovrà essere avviata dall'utente.

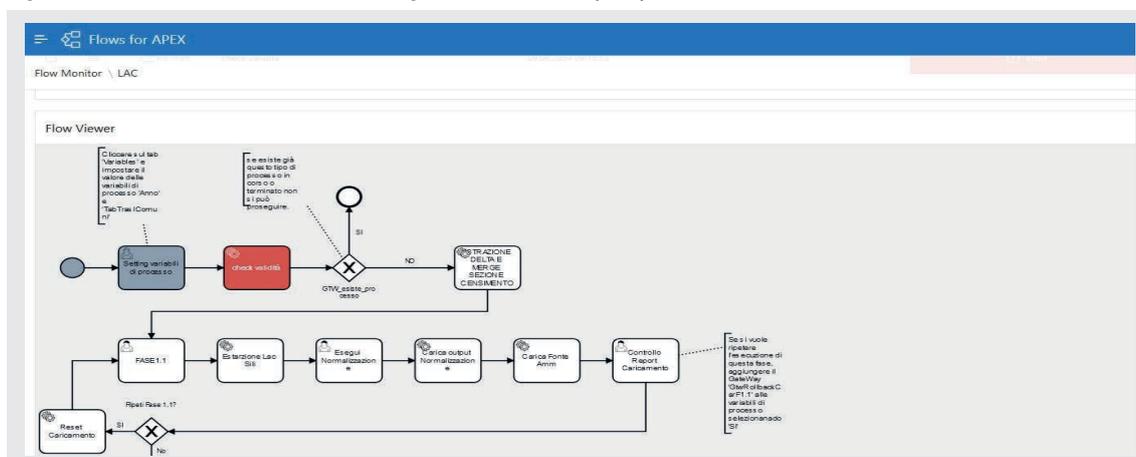
Figura 11.2 - Caricamento delle Liste Anagrafiche Comunali (LAC): vista parziale



Fonte: Istat, Registro statistico di base dei luoghi - Indirizzi

Se un *task* dovesse andare in errore diventa di colore rosso, evidenziando intuitivamente il punto dove si è verificato il problema e il processo si interrompe automaticamente (Figura 11.3).

Figura 11.3 - Caricamento delle Liste Anagrafiche Comunali (LAC): interruzione in presenza di un errore

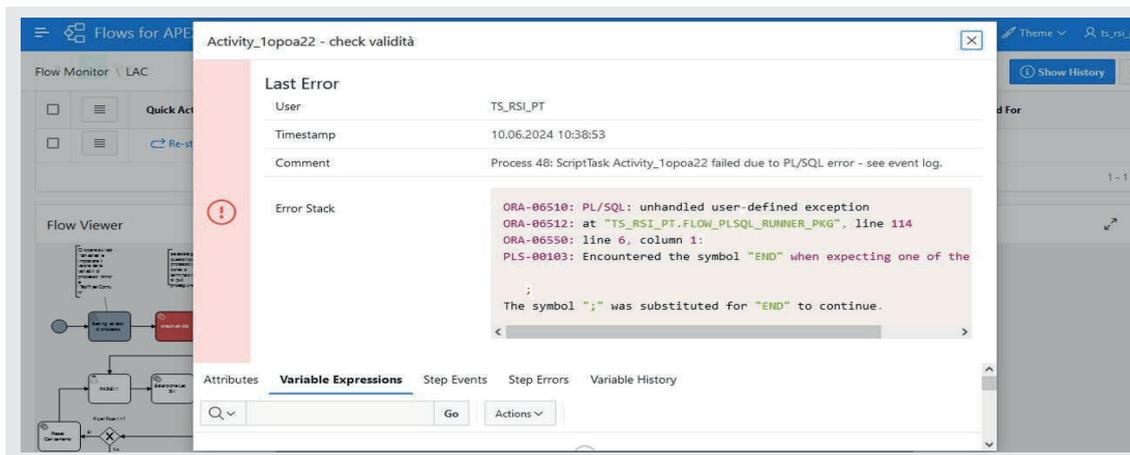


Fonte: Istat, Registro statistico di base dei luoghi - Indirizzi

11. Il processo tecnologico

È possibile visualizzare la descrizione dell'errore (Figura 11.4) per avere modo di risolverlo. Una volta risolto, si potrà riavviare il *task* e l'esecuzione del flusso proseguirà automaticamente fino al prossimo *task* manuale.

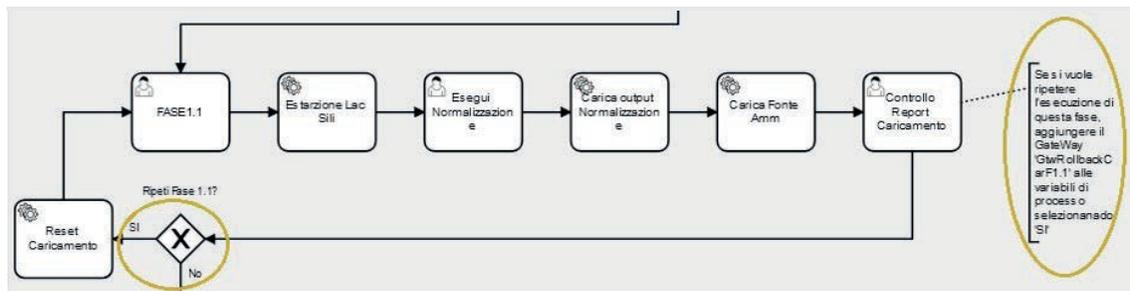
Figura 11.4 - Caricamento delle Liste Anagrafiche Comunali (LAC): visualizzazione della descrizione dell'errore



Fonte: Istat, Registro statistico di base dei luoghi - Indirizzi

In alcuni casi può essere necessario ripetere alcune sottofasi del processo successivamente ai controlli eseguiti sui dati appena caricati. Il flusso è stato pertanto modellato, ad esempio per la *Fase 1.1*, con un punto di decisione che consentisse all'utente di scegliere se annullare il caricamento appena fatto e ripetere la fase (Figura 11.5).

Figura 11.5 - Caricamento delle Liste Anagrafiche Comunali (LAC): punto di decisione relativo alla *Fase 1.1*



Fonte: Istat, Registro statistico di base dei luoghi - Indirizzi

11.3 La reportistica

Utilizzando *APEX* è stato realizzato un applicativo web che raccoglie la reportistica utile sia al monitoraggio delle varie fasi di caricamento sia alla interrogazione puntuale dei dati presenti nel registro. Ad esempio, è possibile visualizzare lo stato del flusso dei caricamenti interrogando il *report* che, per ogni fase, indica il numero dei *record* estratti, elaborati e caricati secondo l'operazione svolta. È possibile, inoltre, richiedere la storia di un particolare indirizzo per vedere quante volte e come è stato trattato e caricato o, ancora, visualizzare i *record* scartati da una fase del processo di caricamento per evidenziare eventuali errori sistemati su un certo comune (ad esempio, nella stringa indirizzo è sempre presente anche il nome del comune) e intervenire sugli indirizzi con una nuova fase di caricamento ad hoc.

Il *report* “distribuzione caricamenti” è utile per l'analisi della qualità degli indirizzi delle fonti: rappresenta per ogni comune la distribuzione per numero di indirizzi, per qualità CUI, per numero di forme (*Thesaurus*) e per scarti (Figura 11.6).

Figura 11.6 - Caricamento delle Liste Anagrafiche Comunali (LAC): *report* della distribuzione dei caricamenti

Anno Fonte	Fonte	Cod Rag	Cod Uts	Cod Pro Com	Fase	Qualità A	Qualità B	Indirizzi	Thesaurus	No Linked	No Validi	Totali	Perc No Link Su Tot
2023	40	01	001	001001	F1	0	0	0	0	0	0	0	0,00
2023	40	01	001	001002	F1	0	0	0	3	1	0	4	25,00
2023	40	01	001	001003	F1	0	0	0	7	0	0	7	0,00
2023	40	01	001	001004	F1	0	0	0	4	0	0	4	0,00
2023	40	01	001	001006	F1	0	0	0	23	0	0	23	0,00
2023	40	01	001	001007	F1	0	0	0	3	0	0	3	0,00
2023	40	01	001	001008	F1	0	0	0	111	0	4	115	0,00
2023	40	01	001	001009	F1	0	0	0	3	1	0	4	25,00
2023	40	01	001	001010	F1	0	0	0	4	0	0	4	0,00
2023	40	01	001	001011	F1	0	0	0	6	0	0	6	0,00

Fonte: Istat, Registro statistico di base dei luoghi - Indirizzi

Sui *report* può essere abilitata la possibilità di fare varie operazioni che permettono all'utente di creare in autonomia delle versioni differenti a partire da un *report* di base: filtrare il risultato su specifiche colonne; evidenziare risultati anomali a livello di riga o cella; creare nuove colonne risultato di quelle presenti; raggruppare il risultato presentato per una o più variabili, definendo le funzioni da applicare. Ad esempio, si può definire il raggruppamento per anno di elaborazione e fase e le funzioni di somma sulle variabili “Totali”, “Qualità A”, “Qualità B”, “Thesaurus”, “scarti” (*no linked*) e “No validi” (Figura 11.7).

Figura 11.7 - Raggruppamento per anno di elaborazione e fase di caricamento

Anno Fonte	Fase	Somma Totali	Somma Qualità A	Somma Qualità B	Somma Thesaurus	Somma No Linked	Somma No Validi
2023	F1	298.873	2.888	351	240.721	50.999	3.914
2023	F2_1	3.914	32	3	3.405	469	5
2023	F3_DUG	157.495	48	31	471	156.867	78
2023	F3_LOCALITA	12.927	13	5	320	12.587	2
2023	F3_INDILOC	38.223	196	72	20.909	17.035	11
2023	F3_SPACC	234.446	658	846	34.688	179.253	19.001
2023	F3_TRATT	33.084	54	22	1.855	30.919	234
2023	TOTALE	491.690	3.889	1.330	302.141	164.994	19.326
2023	TOTALE_ANNO	298.873	3.316	454	280.567	13.097	1.439

Fonte: Istat, Registro statistico di base dei luoghi - Indirizzi

Ogni *report* definito dall'utente può essere salvato, in modo che sia sempre disponibile per le future interrogazioni. Questo è uno strumento potente per l'utente che ha la possibilità di effettuare innumerevoli operazioni sui *report* forniti e così effettuare analisi sempre più approfondite, anche corredate da grafici generati automaticamente con il supporto del sistema.

I *report*, infine, possono essere scaricati o inviati per email nei formati Excel, csv e pdf. Ciò significa che l'utente può scaricare il *report* base con tutte le informazioni disponibili, oppure la versione che ha creato, come il *report* con il raggruppamento mostrato in precedenza (Figura 11.7).

La sinergia dell'impiego di *APEX* e *Flows for APEX* fornisce all'utente uno strumento semplice e completo che lo rende autonomo nel caricamento degli indirizzi e nella interrogazione del registro.

12. LA GRIGLIA DI POPOLAZIONE¹

12.1 Definizione e uso della griglia di popolazione

Il Regolamento di esecuzione (UE) 2018/1799 della Commissione europea prevede la diffusione di una selezione di variabili relative al Censimento della Popolazione e delle abitazioni del 2021, geocodificate in base a una griglia regolare con celle di 1 km². Come da Regolamento, entro il 31 dicembre 2022 gli stati membri hanno fornito a Eurostat il totale della popolazione per singola cella (in formato provvisorio) ed entro il 31 marzo 2024 hanno prodotto la versione definitiva con le informazioni statistiche relative alla popolazione.

Di seguito si descrive la nuova metodologia utilizzata da Istat per la produzione dello strato informativo per la rappresentazione della popolazione censuaria al 2021 per griglia regolare di 1 km². Tale metodologia si differenzia notevolmente dall'approccio utilizzato per la stima della popolazione relativa al 2011 (Istat 2011) sulla medesima griglia regolare. Tali elaborazioni sono, infatti, basate sul Registro statistico di base dei luoghi (RSBL).

La griglia a maglia regolare a fini statistici presenta diverse peculiarità, tra le quali le principali sono:

- tutte le celle della griglia hanno la stessa dimensione e possono essere facilmente confrontate tra loro;
- la griglia regolare, che copre tutto il territorio europeo, è stabile nel tempo, permettendo analisi sia spaziali sia longitudinali;
- i dati all'interno della griglia si integrano molto facilmente;
- un sistema basato sulla griglia può essere suddiviso e aggregato a prescindere dalle suddivisioni gerarchiche delle unità amministrative.

La griglia regolare Eurostat², come si può evincere dai punti elencati, semplificherebbe i confronti della distribuzione della popolazione tra i diversi paesi europei. Essa costituisce un potente strumento per l'analisi spaziale soprattutto per fini demografici e rappresenta la distribuzione della popolazione su una carta suddivisa in celle (o griglie regolari). In ogni cella è contenuto il numero di persone che vi abitano, basato su dati censuari o altre fonti demografiche. Le potenzialità di questo approccio sono numerose e si manifestano in vari ambiti.

La griglia di popolazione permette una maggiore precisione geografica rispetto all'utilizzo dei tradizionali confini delle unità amministrative, come comuni o province. Questi ultimi possono essere di dimensioni molto variabili, influenzando fortemente l'analisi dei dati. Una griglia uniforme, invece, consente di studiare la distribuzione della popolazione in modo più omogeneo, evidenziando meglio le aree ad alta o bassa densità, indipendentemente dalle suddivisioni amministrative.

La griglia di popolazione viene peraltro prodotta da diverse istituzioni internazionali. I principali produttori di griglie di popolazione includono:

¹ Il Capitolo è stato redatto da Davide Fardelli (paragrafi 12.1 e 12.3), Stefano Mugnoli (paragrafo 12.2).

² Per approfondimenti sulla griglia Eurostat cfr. https://gisco-services.ec.europa.eu/grid/GISCO_grid_metadata.pdf.

- Eurostat, come detto, regola la produzione della griglia di popolazione da parte di tutti i paesi membri dell'UE, basata su celle standard di 1 km². Eurostat utilizza dati forniti dai singoli stati membri, che a loro volta raccolgono informazioni attraverso censimenti e altre fonti amministrative;
- gli Istituti nazionali di statistica producono, quindi, la griglia di popolazione. Ad esempio, l'Istat elabora griglie basate sui dati del Censimento nazionale, adattando le informazioni ai confini geografici interni del proprio paese;
- Le Nazioni Unite attraverso il *Global Human Settlement Layer* (GHSL) prodotto dal *Joint Research Centre* della Commissione Europea, fornisce griglie di popolazione su scala globale. Il GHSL combina dati satellitari e censuari per mappare l'occupazione del suolo e la distribuzione della popolazione;
- la NASA e il *Center for International Earth Science Information Network* (CIESIN) della *Columbia University* producono il *Gridded Population of the World* (GPW). Questo *dataset* globale utilizza dati censuari e li converte in griglie uniformi per facilitare l'analisi della popolazione a livello mondiale.
- *WorldPop* è un progetto globale che fornisce griglie di popolazione ad alta risoluzione per vari paesi del mondo, utilizzando una combinazione di dati censuari, immagini satellitari e modelli statistici. È particolarmente utile per le analisi in paesi in via di sviluppo, dove i dati censuari possono essere scarsi o poco aggiornati.

12.2 La griglia di popolazione attraverso il metodo *top-down*

Il risultato finale pubblicato da Eurostat nel 2016 (su dati di popolazione 2011) (Istat 2011) ha rappresentato una proficua collaborazione tra l'Istat e Ispra (Istituto superiore per la prevenzione e la ricerca ambientale). Il *layer* prodotto è il risultato dell'integrazione di vari file georiferiti con la porzione italiana della *GEOSTAT Grid 2011*. Sono stati utilizzati a tale proposito molti strati geografici digitali e tra questi i più importanti sono stati sicuramente:

- lo strato delle sezioni di Censimento 2011;
- lo strato *raster HR Layer* relativo al grado di impermeabilizzazione del suolo, con pixel di 20 metri, prodotto del progetto *Copernicus*;
- la Carta sintetica statistica di copertura del suolo, un *layer* digitale prototipale che è stato la base per la produzione dello strato cartografico delle Sezioni di Censimento 2021;
- le carte tecniche e tematiche regionali.

Quanto esposto è stato valutato anche nel 2021 per la produzione della nuova griglia di popolazione al fine di considerare l'apporto complementare di questa metodologia utile principalmente nel territorio extraurbano, rispetto a quella *bottom-up* (cfr. paragrafo 12.3) sicuramente più performante in ambiente urbano.

La metodologia *top-down* utilizzata nel 2021 per la produzione dell'aggiornamento della griglia di popolazione prevede passaggi che si basano essenzialmente sull'elaborazione di strati georiferiti vettoriali e *raster*. Tra questi ultimi il principale è sicuramente lo strato *built-up* prodotto nell'ambito del progetto *Copernicus*.

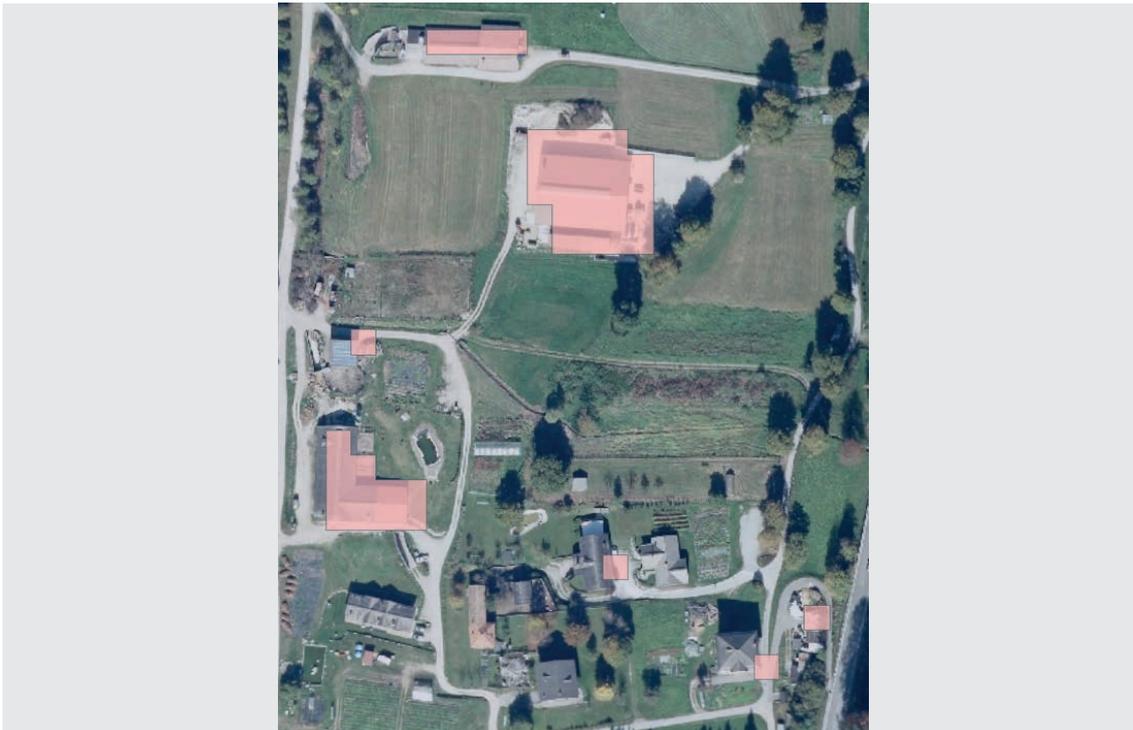
Il tutto utilizzando algoritmi GIS che permettono elaborazioni e analisi spaziali in maniera veloce e molto precisa.

Sono però necessari alcuni chiarimenti. Come si può osservare nella Figura 12.1, lo strato *built-up* (poligoni di colore rosa chiaro) in alcuni rari casi non abbracciano l'intera superficie coperta da edifici. In questo caso la popolazione della SdC 2011, laddove *intersect* con il *built-up* non fornisce alcun poligono, è stata riproporzionata associandola al

12. La griglia di popolazione

centroide della sezione (se questa era di Centro o Nucleo abitato), oppure, per le sezioni Case Sparse, al centroide della sezione di località abitata più vicina. Tale artefatto, anche in questo caso, non influenza il risultato della generalizzazione alla maglia di 1 km²; è utile comunque specificare che situazioni di questo tipo sono pochissime.

Figura 12.1 - Caso di mancata copertura dello strato 'Built-Up'



Fonte: Elaborazione Istat su dati Copernicus e Agea

12.3 Iter realizzativi per la produzione della griglia di popolazione finale: il metodo *bottom-up*

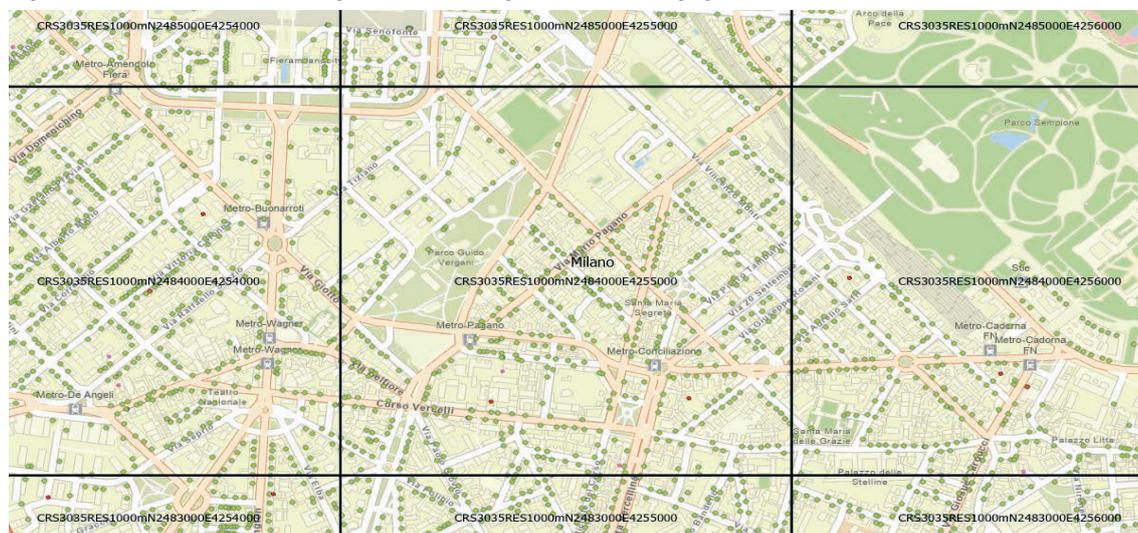
Di seguito si descriverà l'iter realizzativo per la produzione dello strato informativo che rappresenta la distribuzione della Popolazione censuaria al 2021 per maglia regolare di 1 km². L'attuale griglia della popolazione che si descriverà è quella rilasciata da Eurostat e pubblicata anche nel sito web dell'Istat (Istat 2011).

La metodologia denominata *bottom-up* prevede l'aggregazione di microdati georeferenziati. Questa metodologia, per potere essere attuata, richiede una ampia disponibilità di microdati geocodificati che vengono, a sua volta, aggregati nella cella per chilometro quadrato.

Questo metodo si attua con la sovrapposizione (*overlapping*) delle coordinate geografiche degli indirizzi e/o degli edifici, alle quali è associata la popolazione residente, con il reticolato elaborato da Eurostat. Tale sovrapposizione è effettuata mediante il tool di analisi spaziale di geoprocessing, denominato *Spatial Join*, presente in tutti i software GIS. Attraverso tale strumento si associa a ogni coordinata geografica il corrispondente codice di griglia³ in cui ricade la coordinata geografica (Figura 12.2).

³ Il codice della griglia è fornito direttamente da Eurostat.

Figura 12.2 - Esempio di indirizzi georeferenziati e geocodificati alla griglia



Fonte: Elaborazione Istat su dati Istat ed Eurostat

Tale associazione è univoca, ovvero per ogni coordinata si avrà esclusivamente e univocamente una cella di appartenenza. Come *step* successivo, si effettua il conteggio degli indirizzi o degli edifici all'interno di ogni cella e la somma degli individui presenti in ognuno di essi.

Questa tecnica è fortemente condizionata ed è sensibile principalmente a tre fattori:

- copertura e disponibilità di coordinate geografiche relative agli indirizzi e agli edifici;
- l'accuratezza posizionale delle coordinate geografiche dell'indirizzo, ovvero il grado di precisione nel posizionamento dei punti sulla carta rispetto alla realtà;
- qualità dei link tra indirizzo e famiglie, ed edificio e famiglie.

Sicuramente laddove i tre indicatori sopradescritti si presentino con valori elevati, la qualità dell'output è alta e il prodotto finale risulterà molto più accurato. È necessario precisare che, la generalizzazione del dato puntuale all'interno della cella di 1 km² porta comunque a un ottimo risultato, anche per quei punti la cui qualità non risulta eccellente per gli aspetti sopra esposti.

L'attuale griglia è stata prodotta sia con l'utilizzo delle coordinate geografiche degli indirizzi di residenza degli individui, sia con le coordinate del centroide dell'edificio al quale la famiglia è associata. A supporto delle precedenti informazioni geografiche si è utilizzata anche la geocodifica alla sezione di Censimento (BT2011) degli indirizzi di residenza. La popolazione residente coperta dalla coordinata geografica dell'indirizzo è pari a circa 56,6 milioni di individui (96,1 per cento). La popolazione residente coperta dalla coordinata geografica relativa al centroide dell'edificio, invece, è pari a 51,5 milioni di individui (87,5 per cento). Una quota marginale, pari a circa 500 mila individui (0,9 per cento), è localizzabile all'interno del comune, attraverso la geocodifica alla sola sezione di Censimento del 2011, senza l'utilizzo di coordinate geografiche.

Per potere geocodificare la popolazione alla griglia, si è optato per l'utilizzo di una metodologia che prende in considerazione il dato delle due coppie di coordinate (relative all'indirizzo e all'edificio). Le divergenze nei risultati tra le due fonti sono la conseguenza di una non ottimale accuratezza posizionale degli indirizzi e/o degli edifici alla quale la popolazione è associata. Il numero di individui per il quale si è riscontrata la divergenza, risultando in celle differenti della griglia, tra la coordinata dell'indirizzo e la coordinata dell'edificio,

12. La griglia di popolazione

ammonta a circa 4,9 milioni di individui. Invece, il livello della coerenza della geocodifica alla griglia tra le due fonti è pari circa a 46,6 milioni di individui. Per i restanti 7,3 milioni di individui, ove non vi erano informazioni geografiche relative alla copertura degli edifici, si è proceduto all'utilizzo esclusivo delle coordinate degli indirizzi (6,9 milioni).

In sintesi, la popolazione residente geocodificata alla griglia regolare è pari a 58.933.908, con una percentuale pari al 99,8 per cento. La popolazione inserita nella cella virtuale è pari a 96.225 unità. Tale popolazione è stata classificata come persone senza fissa dimora.

Le coordinate geografiche degli indirizzi sono di tipo rilevate, interpolate e, infine, approssimate. Le coordinate geografiche rilevate dovrebbero indicare il punto esatto dove si trova l'accesso dall'area di circolazione. Purtroppo, pur avendo coordinate rilevate, è stato riscontrato che non sempre l'accuratezza posizionale rispecchia la realtà. La popolazione geocodificata alla griglia con questa tipologia di coordinate è di circa 46,8 milioni di individui, per un totale di coordinate pari a circa 8,7 milioni. La seconda tipologia di coordinate utilizzate sono quelle interpolate dal grafo stradale. La popolazione coperta da questa tipologia è pari a 3,9 milioni di individui, per un totale di coordinate pari a circa 900 mila. Infine, l'ultima tipologia di coordinate utilizzate sono quelle approssimate al numero civico più vicino. L'approssimazione massima utilizzata è quella a 10 numeri civici, indipendentemente dal lato della strada in quest'ultima categoria sono presenti circa 5,9 milioni di individui, per un totale di coordinate pari a 2,1 milioni.

Analizzate le differenze tra le due metodologie (*top-down* e *bottom-up*) sono state rilevate differenze non rilevanti tra gli output dei due metodi. Poiché il metodo *top-down* parte da dati aggregati, si rischia di perdere la precisione nella distribuzione della popolazione nelle piccole aree, specialmente se la densità della popolazione non è uniforme. La difficoltà insita nel metodo *bottom-up*, è legata alla necessità di avere consistenti dati dettagliati, difficili e onerosi da ottenere e da processare. Ciò è stato superato grazie all'infrastruttura di RSBL. Il metodo *bottom-up* rende possibile la costruzione della griglia, mantenendo integro il microdato a livello di individuo/famiglia. La griglia di popolazione produce, non solo dati relativi al conteggio di popolazione ma, anche, altre variabili socio-economiche. Per la produzione su griglia di queste ulteriori informazioni statistiche, adottando il metodo *bottom-up*, ogni individuo/famiglia conserva in maniera integra le proprie caratteristiche, le quali sono proiettate all'interno della griglia. Tale processo non è possibile adottando il metodo *top-down* se non con l'implementazione di stime statistiche.

La griglia di popolazione trova applicazioni pratiche in settori come l'urbanistica, la gestione delle risorse e la pianificazione dei servizi pubblici. Ad esempio, per la costruzione di infrastrutture, come scuole o ospedali, la griglia offre una base dati affidabile per individuare le aree con maggiore necessità di determinati servizi. Anche in situazioni di emergenza, come disastri naturali, una mappatura precisa della popolazione permette di organizzare gli interventi in modo efficiente.

La griglia di popolazione è uno strumento fondamentale per la sostenibilità ambientale. Integrare i dati demografici con informazioni su uso del suolo e risorse naturali aiuta a sviluppare politiche di sviluppo sostenibile, bilanciando crescita urbana e tutela dell'ambiente.

La griglia di popolazione è, quindi, una risorsa versatile che, grazie alla sua precisione e alla sua capacità di adattarsi a diverse scale geografiche, offre importanti vantaggi per la pianificazione territoriale e lo sviluppo di politiche pubbliche.

13. IL CENSIMENTO PERMANENTE DELLA POPOLAZIONE E DELLE ABITAZIONI E IL REGISTRO RSBL¹

13.1 Il Censimento permanente della Popolazione e delle abitazioni

13.1.1 Cambiamenti metodologici rispetto ai Censimenti tradizionali

Il passaggio dai Censimenti tradizionali al Censimento permanente della Popolazione e delle abitazioni ha richiesto una profonda riorganizzazione dell'intera macchina censuaria e notevoli cambiamenti nella metodologia di produzione dei dati.

Fino al 2011 il Censimento generale della Popolazione e delle abitazioni aveva carattere universale, veniva realizzato ogni dieci anni attraverso una rilevazione porta a porta e vedeva coinvolte tutte le famiglie presenti sul territorio nazionale. Dal 2018, con l'avvio del Censimento permanente, si è invece puntato sull'impiego combinato di archivi amministrativi, registri statistici e indagini con cadenza annuale su un campione di famiglie (Istat 2022). Questo cambio di paradigma, grazie a innovative tecniche organizzative e statistiche, rende possibile l'aggiornamento annuale delle principali variabili censuarie della popolazione, quali sesso, età, cittadinanza, grado di istruzione e condizione professionale. Anche la produzione di statistiche su famiglie, abitazioni ed edifici, nonché la diffusione di informazioni dettagliate fino a un livello territoriale subcomunale, può essere garantita con maggiore frequenza e tempestività rispetto al passato.

Nel Censimento permanente un ruolo determinante è assunto dal Registro di base degli individui che, insieme agli altri registri tematici come quelli sull'occupazione e sull'istruzione, consente il conteggio annuale della popolazione abitualmente dimorante e la fornitura delle principali variabili socio-demografiche della popolazione. Il conteggio viene effettuato sulla base dei "segnali di vita amministrativi" degli individui (Borrelli *et al.* 2016), integrando le informazioni contenute negli archivi amministrativi e nei registri tematici con i risultati delle rilevazioni sul campo, anche attraverso il supporto di modelli statistici. Le informazioni socio-economiche quali, ad esempio, la condizione professionale, il grado di istruzione, le caratteristiche delle abitazioni occupate e gli spostamenti pendolari per studio e per lavoro, sono, invece, ottenute attraverso l'integrazione di registri statistici, archivi e i dati rilevati attraverso le indagini campionarie annuali.

Altro registro fondamentale per una parte importante della produzione statistica del Censimento permanente è il Registro statistico di base dei luoghi (RSBL). Esso consente, infatti, di ottenere sia le informazioni sulla consistenza numerica delle abitazioni e degli edifici, sia di ricondurre i dati di popolazione, famiglie, abitazioni ed edifici, a un livello territoriale micro (sezioni di Censimento).

La necessità di adottare una nuova modalità di allocazione della popolazione a livello subcomunale è un'altra conseguenza del passaggio dal Censimento tradizionale a quello

¹ Il Capitolo è stato redatto da Raffaele Ferrara e Alessandro Sasso.

permanente. Nel Censimento tradizionale la geocodifica alle sezioni di Censimento (unità minima territoriale di rilevazione) della popolazione avveniva contestualmente alla rilevazione sul campo (erano i comuni, tramite rilevatori, che enumeravano tutte le unità statistiche presenti nelle varie sezioni di Censimento); nel Censimento permanente la geocodifica della popolazione e delle abitazioni avviene, invece, in una fase successiva alle indagini sul campo e si basa sul collegamento tra i registri RBI e RSBL. Il collegamento tra i due registri consente di associare gli individui e le famiglie rilevati annualmente al Censimento alle rispettive abitazioni ed edifici, nonché di stabilire una geocodifica territoriale univoca e coerente per tutte le unità statistiche, con la conseguente possibilità di fornire, anche per livelli territoriali molto fini, variabili o incroci di variabili di notevole importanza per la diffusione censuaria.

13.1.2 Il processo di produzione dei dati sulle abitazioni

La nuova strategia censuaria definisce il numero totale di abitazioni attraverso il trattamento statistico² delle informazioni presenti in RSBL, in particolare della componente del Registro degli edifici e delle abitazioni, la cui fonte primaria è costituita dal catasto dei fabbricati. Il numero di abitazioni occupate e di abitazioni non occupate viene ottenuto associando, attraverso un complesso processo di *linkage* (cfr. Capitoli 8 e 9), le abitazioni del suddetto registro alle famiglie rilevate al Censimento.

Nella tornata censuaria del 2021, per garantire una rapida diffusione dei risultati comunali³, il processo di *linkage* tra i microdati è stato affiancato da una procedura di calcolo realizzata con dati aggregati a livello comunale. Questa procedura ha individuato il numero di abitazioni non occupate per differenza tra il numero di abitazioni totali (definito nel Registro degli edifici e delle abitazioni) e il numero di famiglie, scorporando da queste ultime le cosiddette popolazioni speciali, ossia tutte le famiglie che non vivono in una abitazione, e tenendo conto dell'elemento della coabitazione. Successivamente le abitazioni occupate sono state ottenute per differenza tra le abitazioni totali e le abitazioni non occupate.

Le informazioni sulle caratteristiche delle abitazioni sono, invece, il risultato di modelli di stima che integrano i dati presenti nel Registro degli edifici e delle abitazioni con quelli delle indagini campionarie annuali del Censimento permanente.

In considerazione delle novità metodologiche intervenute, va sottolineato che i confronti tra i dati attuali e quelli del passato potrebbero risentire non solo dell'effettiva evoluzione dei fenomeni in esame, ma anche del diverso sistema di produzione dei risultati definitivi. Il dato delle abitazioni occupate si ritiene in ogni caso piuttosto robusto e stabile, proprio perché ottenuto dall'associazione con le famiglie.

13.1.3 Allocazione delle unità statistiche a livello subcomunale

Sulla base del collegamento tra le famiglie di RBI e le abitazioni di RSBL, richiamato al punto precedente, ha origine la lavorazione per l'allocazione a livello subcomunale di tutte le unità statistiche, così come illustrato nel Capitolo 9. Tale lavorazione è riassumibile in

² Sono state adottate una serie di procedure di coerenza e verifica tra il dato delle abitazioni di fonte catastale e il numero di famiglie censite per ogni comune.

³ La validazione dei risultati a livello subcomunale della procedura di linkage tra i microdati di famiglie e abitazioni 2021 è terminata alla fine del 2023; i dati delle abitazioni occupate e non occupate 2021 a livello comunale sono stati diffusi prima dell'estate del 2023.

tre fasi principali: una prima fase nella quale, per ciascuna famiglia collegata a un alloggio, vengono calcolate, tramite operazioni di *spatial join*, le sezioni di Censimento 2021 dell'indirizzo di residenza e dell'edificio nel quale la famiglia è stata collocata; una seconda fase dove vengono confrontate le due sezioni ottenute; una terza fase nella quale vengono risolti i casi eventuali di incoerenza tra le due geocodifiche.

Per collocare tutta la popolazione censita sul territorio è necessario geocodificare anche le cosiddette popolazioni speciali, ovvero gli individui che vivono in convivenze anagrafiche, quelli che vivono nei campi attrezzati e i senza fissa dimora. Gli individui in convivenza e quelli che vivono nei campi attrezzati, vengono allocati nelle sezioni di Censimento utilizzando la coordinata degli indirizzi comunicati dai comuni durante le indagini ad hoc che si svolgono annualmente su tali *target* di popolazione. I senza fissa dimora vengono, invece, collocati in sezioni di Censimento fittizie appositamente previste per ciascun comune.

L'allocazione della popolazione sul territorio viene poi controllata e validata con una analisi su più livelli territoriali subcomunali. Per tutti i comuni vengono eseguiti controlli di coerenza tra il tipo di sezione di Censimento e la numerosità (e i *target*) della popolazione in esse collocate. Si valuta, ad esempio, che nelle sezioni di centro urbano vi sia generalmente allocata della popolazione. Nelle sezioni con strutture particolari (ad esempio, caserme, centri di accoglienza per migranti, centri di detenzione, ospedali, scuole, campi attrezzati, eccetera) si valuta che non vi sia una presenza anomala di popolazione residente, a meno che non si tratti di individui che vivono in convivenza anagrafiche o in alloggi non convenzionali. Per i grandi comuni viene anche effettuato un confronto con il passato tra i dati aggregati per aree subcomunali amministrative, al fine di valutarne robustezza e dinamiche evolutive. L'esito della validazione può richiedere, in alcuni casi, anche un riesame delle informazioni di base ed eventualmente una nuova esecuzione delle procedure, laddove i riscontri ottenuti si ritengono palesemente non coerenti o poco verosimili⁴.

Collocata la popolazione censuaria sul territorio è possibile associare a tutti gli individui le relative caratteristiche socio-demografiche e familiari, rendendo in tale modo possibile una diffusione subcomunale molto dettagliata e ricca di variabili classificatorie. A tale proposito, si segnala che il 26 luglio 2024 sono stati rilasciati nuovi dati subcomunali di popolazione e abitazioni, riferiti alla data del 31 dicembre 2021 per le sezioni di Censimento delle nuove Basi territoriali del 2021 (Istat 2024d). Utilizzando questi nuovi dati, nel paragrafo a seguire verrà analizzata, a titolo esemplificativo, la dislocazione territoriale della Popolazione e delle abitazioni occupate all'interno del Comune di Roma.

13.2 Distribuzione della popolazione e delle abitazioni occupate all'interno del Comune di Roma

Il Comune di Roma si estende su di una superficie territoriale di circa 1.288 km² e ha una densità di popolazione che nel 2021 è risultata pari a 2.134 abitanti per km² (Tavola 13.1). Nell'ultimo decennio intercensuario, la popolazione residente è aumentata da 2.617.175 individui nel 2011 a 2.749.031 individui nel 2021, con una variazione percentuale del +5 per cento rispetto all'anno base.

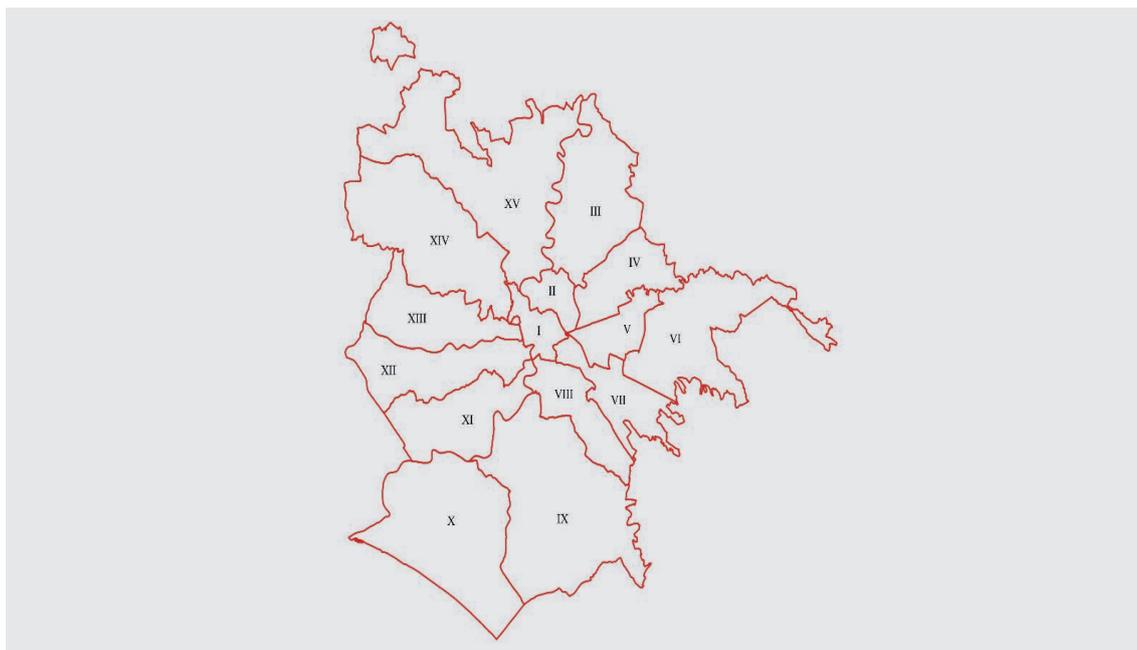
⁴ Nel caso del 2021 è stato necessario, ad esempio, intervenire sul Comune di Napoli al fine di integrare le informazioni di base provenienti dal Catasto che avevano inizialmente determinato una anomala distribuzione della popolazione tra le aree amministrative subcomunali.

La distribuzione della popolazione romana tra i vari municipi (Figura 13.1) non appare cambiata: il Municipio VII, con una superficie di circa 47 km² e una densità di popolazione pari a più di 6.500 abitanti per km², continua a ospitare la quota più ampia di popolazione capitolina con oltre 11 per cento del totale, seguito soltanto a distanza dal Municipio V (8,6 per cento), dal Municipio VI (8,5 per cento) e dal Municipio X (8,1 per cento).

Al contrario, i Municipi VIII e XIII, pure avendo una estensione territoriale maggiore del Municipio VII, ospitano le quote più esigue di popolazione, rispettivamente il 4,5 per cento e il 4,7 per cento.

I Municipi più estesi sono il IX e il XV (entrambi con oltre 180 km²) e ospitano rispettivamente il 6,5 per cento e il 5,6 per cento della popolazione del comune; quelli meno estesi sono invece i Municipi I e II, che si trovano al centro della Capitale (entrambi con circa 20 km²): hanno quote di popolazione pari al 5,9 per cento rispetto al totale, ma una densità di popolazione che è la più elevata di tutto il Comune di Roma (oltre 8 mila abitanti per km²).

Figura 13.1 - Municipi di Roma. Anno 2021



Fonte: Istat, Basi territoriali

Interessanti sono le dinamiche di popolazione intercorse tra il 2011 e il 2021 all'interno di ciascun municipio, in quanto si segnalano variazioni relative chiaramente eterogenee: i Municipi XV, IX e XIV hanno, infatti, incrementato in modo rilevante la loro popolazione, di oltre il 15 per cento il primo e di oltre il 9 per cento gli altri due; i Municipi I e IV hanno, al contrario, subito una leggera variazione negativa (rispettivamente -0,9 per cento e -0,7 per cento). In generale, si può dire che a guadagnare popolazione sono state soprattutto le aree più periferiche della capitale.

Le abitazioni del Comune di Roma sono 1.441.947. Di esse risultano occupate 1.279.874, ovvero l'88,8 per cento, una quota ben più elevata della media nazionale che è pari al 72,8 per cento (Istat 2024e).

13. Il Censimento permanente della Popolazione e delle abitazioni e il registro RSBL

La distribuzione delle abitazioni occupate tra i municipi del Comune è naturalmente molto simile a quella della popolazione: il Municipio VII accoglie la quota più elevata delle abitazioni comunali, ovvero 11,5 per cento (stessa quota anche nel 2011), mentre seguono a distanza i Municipi V (8,6 per cento), X (7,7 per cento) e VI (7,6 per cento); in coda si trovano i Municipi XIII (4,5 per cento) e VIII (4,8 per cento).

Anche nel caso delle abitazioni occupate le variazioni intercensuarie mostrano gli aspetti più interessanti. Tutti i municipi hanno incrementato il numero delle loro abitazioni occupate, ma con modalità molto differenti: il Municipio XV riporta la crescita più rilevante di abitazioni occupate, con un +24,4 per cento rispetto al 2011 (in linea con quella delle famiglie, +24,6 per cento), valore molto superiore alla media comunale che è pari al 12,5 per cento. Anche i Municipi IX e XIV riportano una crescita importante delle unità abitative occupate, quasi il 18 per cento in più e il 16,2 per cento in più rispetto al 2011. Al contrario, le variazioni più esigue appartengono ai Municipi VIII e IV con il 6,4 per cento e il 7 per cento rispetto al primo anno di riferimento.

Tavola 13.1 - Popolazione e abitazioni occupate nel Comune di Roma e nei suoi municipi. Anno 2021 (valori percentuali e valori assoluti)

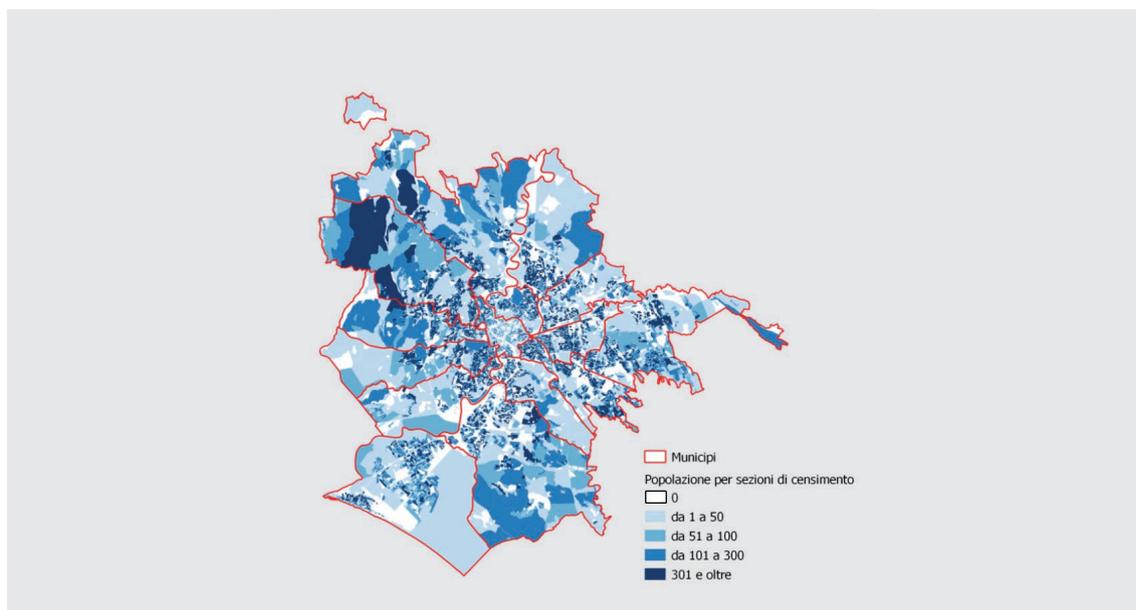
Municipi	Superficie (in km ²)	Densità abitativa.	Popolazione			Abitazioni occupate		
			%		Variazione % 2021-2011	%		Variazione % 2021-2011
			(sul totale comunale)			(sul totale comunale)		
2011	2021	2011	2021	2011	2021			
Municipio I	20,1	8.104,3	6,3	5,9	-0,9	6,3	6,3	12,5
Municipio II	19,7	8.167,0	5,9	5,9	4,4	6,4	6,4	11,3
Municipio III	98,1	2.051,1	7,3	7,3	4,9	7,6	7,5	10,7
Municipio IV	48,0	3.472,4	6,4	6,1	-0,7	6,3	6,0	7,0
Municipio V	30,7	7.710,6	8,6	8,6	4,7	8,8	8,6	10,7
Municipio VI	109,3	2.141,7	8,6	8,5	3,4	7,6	7,6	13,1
Municipio VII	47,0	6.505,9	11,1	11,1	5,4	11,5	11,5	12,5
Municipio VIII	47,4	2.630,1	4,8	4,5	0,3	5,1	4,8	6,4
Municipio IX	183,3	973,2	6,2	6,5	9,4	6,0	6,3	17,9
Municipio X	152,2	1.461,9	8,0	8,1	6,2	7,6	7,7	13,1
Municipio XI	71,5	2.075,5	5,4	5,4	5,6	5,4	5,3	11,3
Municipio XII	73,1	1.881,8	5,0	5,0	5,5	5,2	5,1	12,2
Municipio XIII	67,0	1.912,1	4,7	4,7	3,9	4,6	4,5	10,3
Municipio XIV	133,6	1.395,0	6,5	6,8	9,0	6,5	6,7	16,2
Municipio XV	187,1	826,3	5,1	5,6	15,2	5,0	5,6	24,4
Roma	1.288,2	2.134,0	100,0	100,0	5,0	100,0	100,0	12,5

Fonte: Istat, Censimento permanente della Popolazione e delle abitazioni

Aumentando il dettaglio territoriale, le Figure 13.2 e 13.3 consentono di apprezzare la distribuzione della popolazione e delle abitazioni occupate tra le sezioni di Censimento della Capitale.

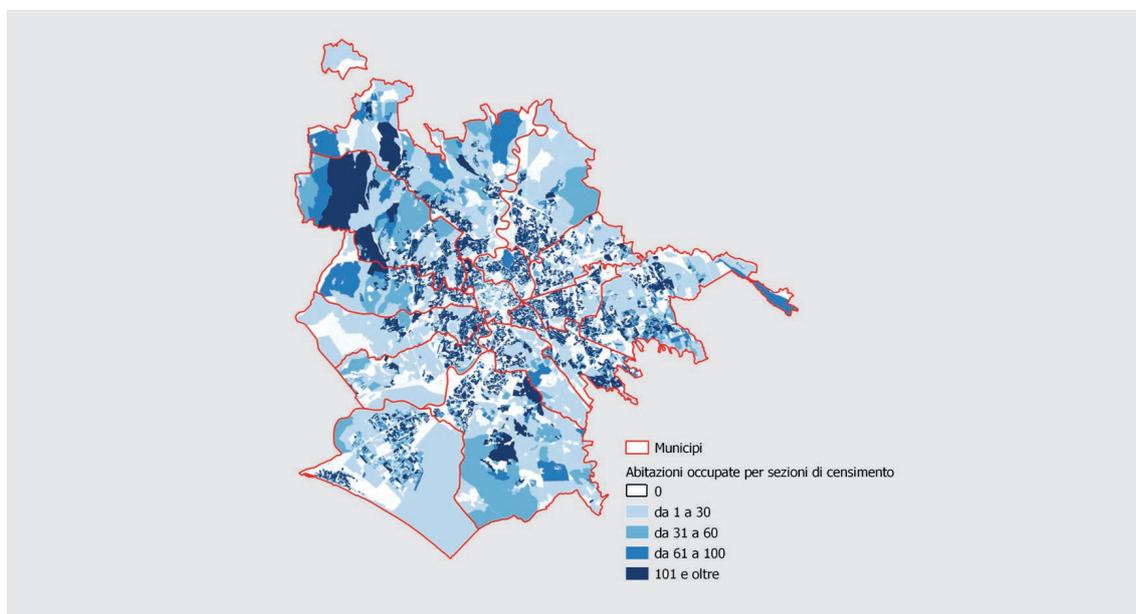


Figura 13.2 - Popolazione residente nelle sezioni di Censimento del Comune di Roma. Anno 2021 (valori assoluti)



Fonte: Istat, Censimento permanente della Popolazione e delle abitazioni

Figura 13.3 - Abitazioni occupate nelle sezioni di Censimento del Comune di Roma. Anno 2021 (valori assoluti)



Fonte: Istat, Censimento permanente della Popolazione e delle abitazioni

Una semplice analisi dei due cartogrammi è sufficiente a rivelare l'esistenza di situazioni molto eterogenee anche all'interno dei domini municipali, rendendo in tale modo manifesta l'utilità di disporre di dati a un livello territoriale sempre più piccolo, per potere approfondire le specifiche caratteristiche della popolazione nei territori locali. Analisi come quella che vedremo nel Capitolo seguente, avente a oggetto lo studio delle periferie urbane dei comuni capoluogo delle Città metropolitane.

14. STUDIO SULLE PERIFERIE URBANE DEI COMUNI CAPOLUOGO DELLE CITTÀ METROPOLITANE: PRIMI RISULTATI¹

14.1 Finalità e obiettivi

Il 26 giugno 2024, in occasione dell'audizione alla Commissione parlamentare di inchiesta sulle condizioni di sicurezza e sullo stato di degrado delle città e delle loro periferie², l'Istat ha presentato i risultati di un lavoro di analisi volto a fornire elementi conoscitivi utili a descrivere in profondità le realtà dei 14 comuni capoluogo delle Città metropolitane³ e a individuare le aree a maggiore criticità e disagio economico.

Il lavoro si è avvalso del potenziale informativo che è parte integrante del Registro statistico base dei luoghi.

Le analisi condotte a livello comunale celano spesso eterogeneità interne tipiche di realtà complesse e articolate quali sono, appunto, i grandi comuni. In essi, infatti, coesistono aree nelle quali la tipologia di insediamenti (produttivi, di servizi, residenziali) e le caratteristiche della popolazione che vi insiste sono associabili a un certo livello di benessere, con aree più periferiche nelle quali è maggiore la vulnerabilità sociale.

Al fine di leggere e interpretare queste differenze, l'analisi è stata condotta a un livello territoriale molto fine, con una granularità che consente di cogliere gli indispensabili dettagli informativi.

Le evidenze fornite hanno come base di partenza le sezioni di Censimento, unità territoriali minime che coprono l'intero territorio nazionale. Con l'edizione 2021 del Censimento permanente della Popolazione e delle abitazioni sono state aggiornate le Basi territoriali⁴ (BT), ovvero l'articolazione del territorio nazionale in sezioni di Censimento (cfr. Capitolo 5).

L'aggiornamento ha garantito un significativo miglioramento della qualità del disegno delle BT, coerenza territoriale e dettaglio: le sezioni di Censimento sono passate dalle circa 403 mila del 2011 alle oltre 756 mila del 2021, quindi, con una capacità di lettura dei fenomeni microterritoriali quasi doppia rispetto al passato.

I dati disponibili per sezione di Censimento sono stati aggregati in aree subcomunali più ampie (Zone statistiche, Zone urbanistiche, Aree elementari, Quartieri, Nuclei di identità locale, e così via), aree internamente omogenee per storia, cultura, stili di vita, opportunità e quindi diverse per ogni Comune considerato; ciò ha consentito di disporre di una lettura "funzionale" dei territori ritenuta più consona per una corretta interpretazione dei fenomeni analizzati.⁵

1 Il Capitolo è stato redatto da Marianna Mantuano, Simona Mastroluca, Fabio Lipizzi e Raffaele Ferrara.

2 Cfr. Camera dei Deputati. Delibera 23 marzo 2023. "Istituzione di una Commissione parlamentare di inchiesta sulle condizioni di sicurezza e sullo stato di degrado delle città e delle loro periferie. (23A01991)". *Gazzetta Ufficiale Serie Generale N. 75 del 29 marzo 2023*. <https://www.gazzettaufficiale.it/eli/gu/2023/03/29/75/sg/pdf>.

3 I comuni capoluogo delle Città metropolitane sono: Bari, Bologna, Cagliari, Catania, Firenze, Genova, Messina, Milano, Napoli, Palermo, Reggio di Calabria, Roma, Torino e Venezia.

4 Cfr. <https://www.istat.it/notizia/basi-territoriali-e-variabili-censuarie/>.

5 Le aree subcomunali identificate per l'analisi dei comuni capoluogo delle Città metropolitane hanno una connotazione storica e funzionale e sono state selezionate per la maggiore omogeneità morfologica, ambientale e demografica

Il lavoro si basa, infatti, sulla costruzione di indicatori demografici e socio-economici, territorialmente dettagliati, ricavati dai dati del Censimento permanente del 2021⁶ e dalle fonti amministrative disponibili, insieme ad alcuni indicatori preliminari sulla disponibilità di servizi sul territorio.

Come già detto, a partire dal 2018, il Censimento permanente della Popolazione e delle abitazioni consente la produzione e la diffusione di dati a un livello territoriale fine, più tempestivi rispetto al tradizionale intervallo decennale. In tale contesto i risultati censuari a livello comunale e subcomunale vengono prodotti massimizzando l'utilizzo delle informazioni incluse nelle fonti amministrative.

Gli indicatori selezionati sono organizzati in diverse aree tematiche: territorio, demografia, istruzione, economia/lavoro, vulnerabilità sociale e materiale, disagio socio-economico, disponibilità di servizi.

Per ciascun comune capoluogo delle Città metropolitane sono stati predisposti elaborati cartografici per inquadrare territorialmente i singoli fenomeni indagati e una breve scheda di approfondimento che ne riporta le evidenze principali. L'elenco dettagliato degli indicatori, le relative modalità di calcolo e i valori ottenuti per ciascuna area, sono consultabili nell'allegato statistico disponibile sul sito web dell'Istat⁷, insieme a tutta la documentazione prodotta per la Commissione parlamentare.

14.2 Periferie urbane di alcuni comuni capoluogo

Si riportano di seguito alcuni dei risultati emersi, concentrando l'attenzione su sei indicatori e tre comuni capoluogo. Gli indicatori considerati sono i seguenti:

- l'indice di vecchiaia;
- la quota di stranieri;
- l'indice di non completamento del ciclo di scuola secondaria di primo grado;
- la quota di giovani che non studia e non lavora;
- la quota di famiglie con figli la cui persona di riferimento ha fino a 64 anni e nelle quali nessun componente è occupato o percettore di una pensione da lavoro;
- la quota di famiglie che non vivono in una abitazione di proprietà.

I sei indicatori considerati permettono di cogliere la complessità sul territorio dei fenomeni di marginalità sociale ma non esauriscono naturalmente l'ampio quadro informativo prodotto e disponibile online.

I tre comuni selezionati sono: Milano, Roma e Napoli. Le partizioni territoriali di riferimento per l'analisi sono i Nuclei di identità locale (NIL) per Milano, le Zone urbanistiche (ZU) per Roma e i Quartieri per Napoli.

rispetto alle altre aree subcomunali esistenti. Alcune di queste, soprattutto nei comuni che presentano un elevato numero di suddivisioni, sono poco popolate e caratterizzate territorialmente dalla presenza di parchi, monasteri, ospedali, cimiteri, eccetera. In questi casi i valori esigui del numero di abitanti e di occupati generano valori degli indici statistici fuori scala e, quindi, non significativi; le aree in questione, relative ai comuni di Bologna, Firenze, Milano, Roma e Torino, sono state, dunque, escluse dall'analisi. Inoltre, in tutti i comuni analizzati vi sono sezioni di Censimento nelle quali sono presenti individui senza fissa dimora. Tali sezioni e la relativa popolazione sono state escluse dall'analisi.

6 Il 26 luglio 2024 sono stati aggiornati i dati subcomunali di popolazione e abitazioni riferiti alla data del 31/12/2021. Cfr. <https://www.istat.it/notizia/dati-per-sezioni-di-censimento/>.

7 Cfr. <https://www.istat.it/audizioni/sicurezza-e-stato-di-degrado-delle-citta-e-delle-loro-periferie/>.

14.2.1 Indice di vecchiaia e quota di stranieri residenti

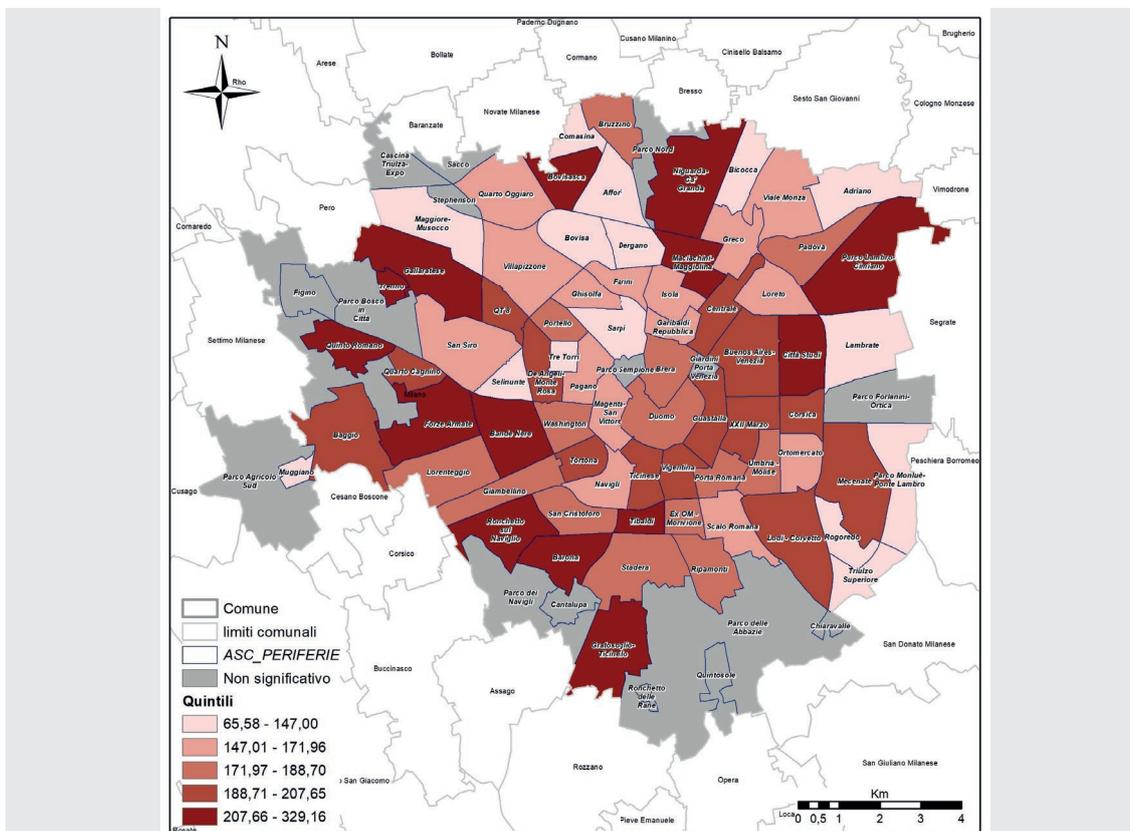
L'Italia è un paese che invecchia: nel 2011 l'indice di vecchiaia era pari a 148,7 persone con più di 64 anni ogni 100 ragazzi di età compresa tra 0 e 14 anni; nell'arco di dieci anni questo rapporto è notevolmente aumentato e nel 2021 è pari a 187,6. A tale risultato hanno contribuito fattori quali l'aumento della speranza di vita alla nascita e la riduzione della natalità che, nel tempo, hanno portato a un cambiamento della struttura demografica. Per contro, l'incremento della componente straniera della popolazione, per effetto di una struttura per età più giovane rispetto a quella degli italiani, ha rallentato il processo di invecchiamento complessivo della popolazione.

Nel Comune di Milano si contano 187,8 stranieri ogni mille abitanti. La popolazione in età avanzata è più del doppio di quella giovane in ben 18 degli 88 NIL. Tra questi, nei NIL Galleratese e Trenno, tra loro confinanti e collocati nell'area Nord-ovest della città, l'indice di vecchiaia rispettivamente sfiora e supera il rapporto di tre a uno tra chi ha almeno 65 anni e chi ne ha al più 14.

Nella quasi totalità dei 18 NIL, la componente straniera della popolazione si attesta su quote inferiori al valore comunale.

Il NIL Triulzo Superiore è il più giovane: gli *over* 64 anni sono i due terzi dei residenti con al massimo 14 anni di età e l'incidenza degli stranieri è più elevata (595,7 ogni mille abitanti). La componente giovane della popolazione prevale su quella anziana anche nei NIL di Rogoredo, Muggiano e Tre Torri.

Figura 14.1 - Indice di vecchiaia nel Comune di Milano. Anno 2021 (a)



Fonte: Istat, Censimento permanente della Popolazione e delle abitazioni
 (a) Rapporto tra la popolazione residente di 65 anni e oltre e la popolazione residente nella classe di età 0-14 anni, moltiplicato per cento.



Gli indicatori demografici denotano situazioni eterogenee nel Comune di Roma, che è al contempo la città più popolosa, con la maggiore estensione territoriale e con il più elevato numero di aree subcomunali. Poco più di metà delle ZU di Roma (82), site prevalentemente all'interno del Grande raccordo anulare (GRA), hanno un indice di vecchiaia superiore a quello rilevato a livello comunale. Le aree con più popolazione in età avanzata sono concentrate, in particolare, nel centro geografico della città (Foro italoico, Prati, XX settembre, Centro storico). Al contrario, l'indicatore propende per la componente più giovane della popolazione in 13 ZU, delle quali nove al di fuori del GRA. A queste si aggiungono Borghesiana, tra le aree subcomunali più popolate (oltre 50 mila residenti), nella quale l'indice di vecchiaia supera di poco la soglia del 100 e si attesta a 101,5, e Torre Angela (circa 85 mila residenti, con un indice di vecchiaia pari a 125,3), situate nella periferia Est della Capitale.

Dei circa 339 mila stranieri che risiedono a Roma nel 2021, oltre 15 mila dimorano abitualmente a Torre Angela (180,3 per mille) e circa 10 mila a Borghesiana (194,4 per mille). Livelli più alti dell'indicatore si rilevano nella parte Nord della città, in particolare a Grottarossa Ovest, dove ogni mille residenti 312,5 sono stranieri; nella stessa area geografica si collocano anche Tomba di Nerone, Cesano e La Storta che presentano una incidenza di stranieri superiore al 200 per mille. Elevata la proporzione anche nelle ZU centrali della Capitale: all'Esquilino l'indicatore è del 234,5 per mille, quota poco più alta di quella registrata nel Centro storico che supera comunque la soglia del 200 per mille anche grazie alla presenza di alcuni centri assistenziali.

Napoli si caratterizza per un indice di vecchiaia pari a 152,6 abitanti *over* 64 anni ogni 100 ragazzi tra 0 e 14 anni. Sebbene tale valore sia inferiore al dato medio nazionale, in nessuna area subcomunale l'indicatore scende al di sotto della soglia 100 (rapporto 1:1 tra le due fasce di età considerate). Ciò è probabilmente riconducibile anche a una ridotta incidenza di residenti stranieri, mediamente inferiore ad altre realtà comunali e pari solo a 58 individui ogni mille abitanti. È quanto si osserva, ad esempio, nei Quartieri di Miano, Secondigliano e San Pietro a Patierno, tutti collocati nell'area Est della città, nei quali l'indice di vecchiaia supera di poco la soglia di parità tra le due generazioni e gli stranieri che vi risiedono arrivano al più a 36,3 ogni mille abitanti.

Per contro, Quartieri quali Arenella, Vomero e San Giuseppe sono caratterizzati più di altri dall'invecchiamento della popolazione; in essi la popolazione in età più avanzata è più del doppio di quella *under* 15. Se consideriamo anche Fuorigrotta (197), Chiaia (191,6) e Posillipo (187,8), possiamo dire che l'area a Ovest, a partire dalla zona collinare della città, è quella nella quale risiede la popolazione più anziana.

I Quartieri più attrattivi per gli stranieri si trovano nelle aree centrali di Napoli: nella Zona industriale se ne contano circa 202 per mille: seguono, con valori più bassi, i Quartieri di Mercato, Stella, San Lorenzo, Pendino, Montecalvario e Avvocata.

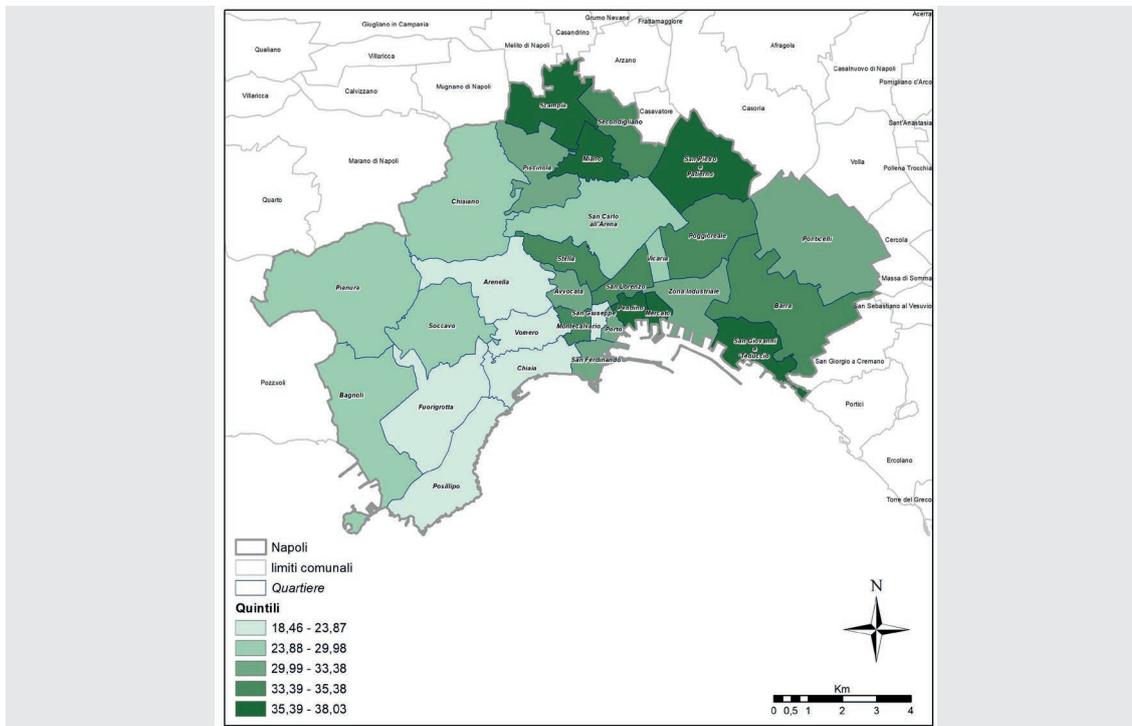
14.2.2 Bassa istruzione e quota di giovani che non studiano e non lavorano

Il Comune di Napoli sembra diviso sostanzialmente in due aree: quella centro-orientale, dove si riscontrano i livelli più elevati di persone meno istruite e di giovani che non studiano e non lavorano, quindi, più esposti a rischio marginalità, e la parte centro-occidentale, dove si colloca la componente di popolazione più istruita e attiva del comune. La popolazione meno istruita raggiunge comunque i livelli più elevati nei quartieri centrali della città, tra il 12,6 e l'8,7 per cento. Le soglie più critiche di giovani che non studiano e non lavorano sono localizzate negli stessi quartieri centrali di Pendino (38 per cento) e Mercato (37,6 per

cento), oltre che in quelli orientali come Scampia (36,7 per cento), San Pietro a Patierno (37,5 per cento) e San Giovanni a Teduccio (37,1 per cento).

A Roma la percentuale di popolazione senza licenza media è appena il 2,3 per cento e raggiunge il suo massimo nell'area subcomunale di Santa Palomba, una ZU con meno di 1.500 residenti situata al confine con i comuni di Albano, Ardea e Pomezia dove si trova un comprensorio residenziale con edifici popolari. L'indicatore si attesta su livelli elevati anche in zone ben più popolate della capitale come il Quadraro, l'Esquilino, la Magliana e Torpignattara. La quota di giovani che non seguono un ciclo di istruzione e non svolgono alcuna attività lavorativa è del 20,8 per cento a livello comunale: il dato cittadino viene abbondantemente superato a Grottarossa Ovest, Santa Palomba e alla Magliana dove supera il 30 per cento, quota di poco più elevata di quella registrata nelle seppure centrali Trastevere, Centro Storico, Aventino e San Lorenzo.

Figura 14.2 - Incidenza dei giovani che non studiano e non lavorano nel Comune di Napoli. Anno 2021 (a)



Fonte: Istat, integrazione dei dati del Censimento della Popolazione e delle abitazioni con informazioni provenienti da archivi amministrativi
 (a) Rapporto tra la popolazione residente nella classe di età 15-29 anni che non studia e non lavora, e la popolazione residente nella medesima classe di età, moltiplicato per cento.

A Milano la situazione generale rispetto ai livelli di povertà educativa e ai giovani che non studiano e non lavorano risulta decisamente meno preoccupante. Non mancano tuttavia situazioni locali più a rischio marginalità rispetto ad altre.

I NIL di Triulzo Superiore e Parco Monluè - Ponte Lambro sono caratterizzati da una più bassa istruzione (rispettivamente 8,4 per cento e 7,6 per cento dei 15- 52 enni, a fronte di una media comunale del 3 per cento) al pari di altre aree nella cintura più periferica della città. Sempre Parco Monluè - Ponte Lambro e Triulzo Superiore, ai quali si aggiungono Barona, Selinunte e i più centrali Duomo e Tre Torri, sono i NIL che mostrano le quote più elevate di giovani che non studiano e non lavorano.

14.2.3 Famiglie con potenziale disagio economico e famiglie che non vivono in casa di proprietà

La maggiore vulnerabilità di alcune aree subcomunali dei capoluoghi considerati si evince anche dal potenziale disagio economico delle famiglie con figli, nelle quali la persona di riferimento ha meno di 64 anni e nessun componente è occupato o percepisce una pensione da lavoro e di quelle, con o senza figli, che non vivono in una abitazione di proprietà.

A Milano l'incidenza delle famiglie senza un reddito da lavoro sul totale delle famiglie residenti si attesta all'1,4 per cento. Le quote più alte sono state rilevate nel NIL Tre Torri (3 per cento) e, a seguire, nell'area Triulzo Superiore (2,6 per cento). In questo ultimo NIL, caratterizzato anche dalla maggiore presenza di alloggi impropri (il 3 per cento degli alloggi occupati è rappresentato, infatti, da luoghi di riparo non identificabili come abitazioni), oltre la metà dei residenti ha cittadinanza straniera. Continuando sul versante del disagio socio-economico delle famiglie, nel NIL Parco Monluè – Ponte Lambro è massima la percentuale di famiglie che non vivono in una abitazione di proprietà (59,7 per cento). Anche a Triulzo Superiore, Selinunte e Quarto Oggiaro più di una famiglia su 2 risiede in un alloggio in affitto o occupato ad altro titolo.

Nella Capitale l'incidenza di famiglie con figli più esposte al disagio economico è più elevata rispetto a Milano ed è pari al 2,3 per cento. Tra le Zone urbanistiche più vulnerabili troviamo la già citata Santa Palomba (5,4 per cento), seguita dalla ZU della Magliana (5,3 per cento) che si contraddistingue, a sua volta, per una percentuale di alloggi impropri (roulotte, caravan, garage, eccetera) del 7,3 per cento, a fronte di una media comunale che non va oltre lo 0,2 per cento.

La rappresentazione cartografica dell'indicatore mostra che, in numerose aree al di fuori del GRA, queste famiglie superano la soglia del 3 per cento, tra le quali: Torre Angela, Borghesiana e Lunghezza a Est; Cesano, La Storta a Nord; Boccea e Massimina a Ovest; Acilia e Infernetto a Sud. Anche in città, in particolare all'Aventino e al Celio, il fenomeno si manifesta con quote maggiori del 3 per cento.

Sempre a Santa Palomba, circa il 70 per cento delle famiglie non ha una casa di proprietà, il 60,1 per cento a Tufello. A Torre Angela, Cesano e Centro Storico le quote si attestano tra il 44 per cento e il 45 per cento circa, al 47,5 per cento a San Basilio.

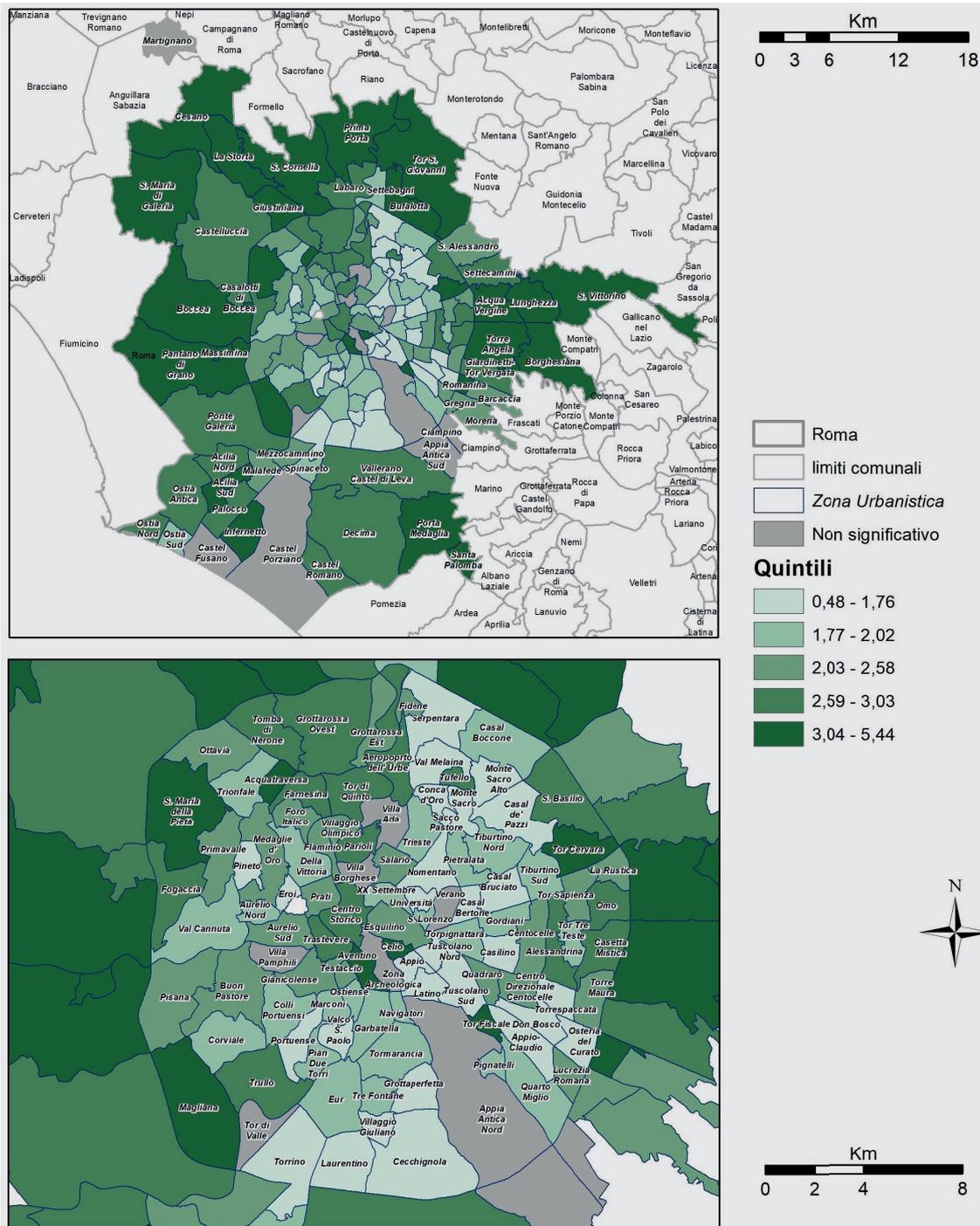
La distribuzione dei dati sul territorio è piuttosto eterogenea: rispetto al dato medio comunale pari al 32,3 per cento, si passa dal 16,7 per cento di Grottaperfetta al già citato 70 per cento rilevato a Sud-est di Roma.

A Napoli, in una ottica di confronto con gli altri comuni analizzati, l'incidenza delle famiglie più vulnerabili raggiunge livelli decisamente più elevati. Nel capoluogo partenopeo, infatti, sei famiglie su 100 devono provvedere al fabbisogno dei figli e di tutti gli altri componenti, senza potere contare su una occupazione o su una pensione da lavoro.

A San Pietro a Patierno la percentuale si colloca al di sopra del 9 per cento ed è oltre l'8 per cento nei Quartieri di Pendino, Secondigliano, San Giovanni a Teduccio e Miano. Decisamente diversa la situazione nei più agiati Arenella e Vomero, dove la quota si ferma, rispettivamente, al 2,7 per cento e 3,2 per cento.

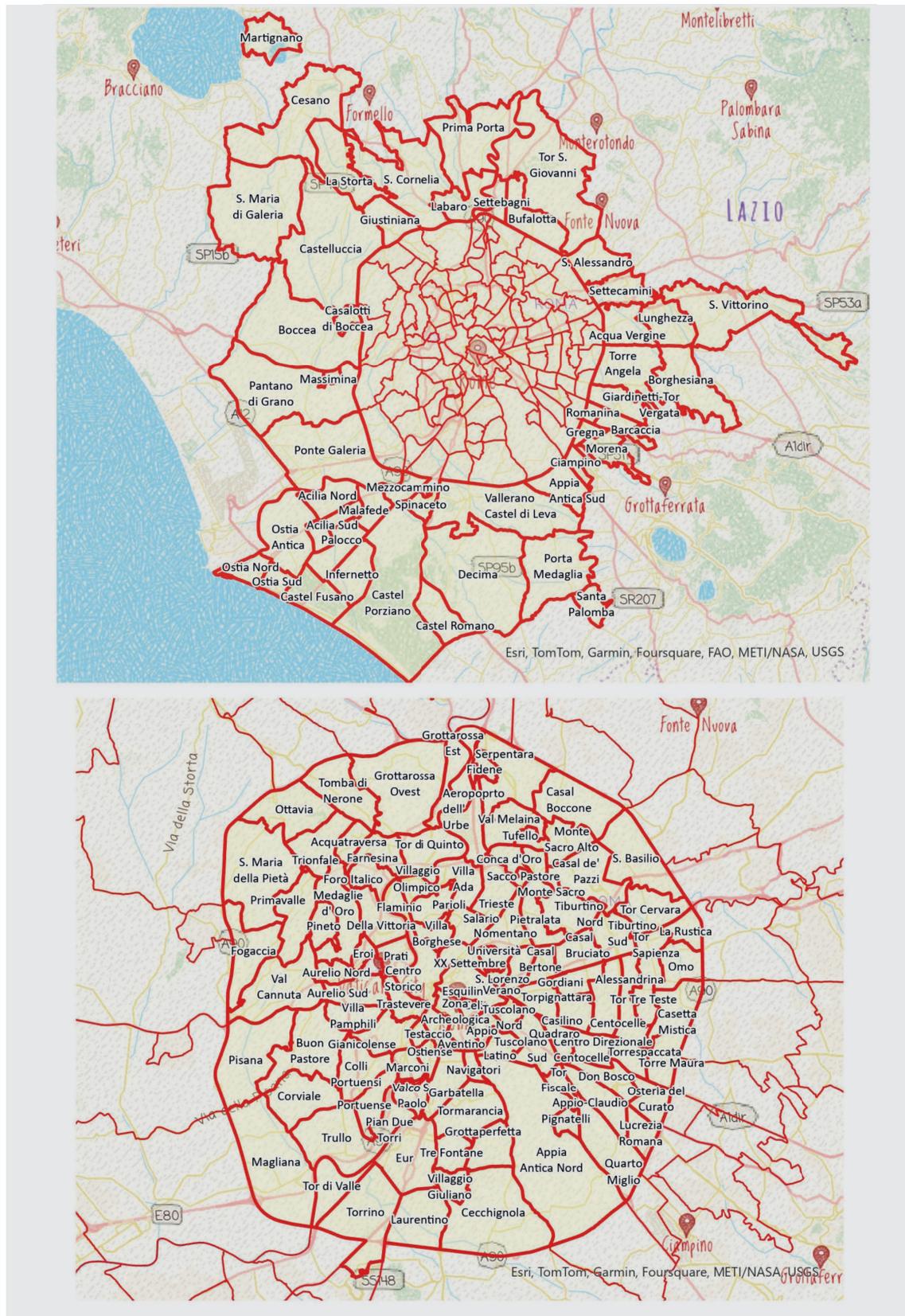
Quasi una famiglia su due tra quelle residenti a Napoli non vive in una casa di proprietà. I dati del Censimento del 2021 mostrano, anche in questo caso, una marcata eterogeneità sul territorio con San Pietro a Patierno, Secondigliano, Piscinola, Miano e Scampia che registrano più del 60 per cento di famiglie con una situazione abitativa meno rassicurante, al contrario di quanto avviene sempre ad Arenella e Vomero in cui l'incidenza approssima il 30 per cento del totale delle famiglie.

Figura 14.3 - Incidenza delle famiglie con potenziale disagio economico nel Comune di Roma. Anno 2021 (a)



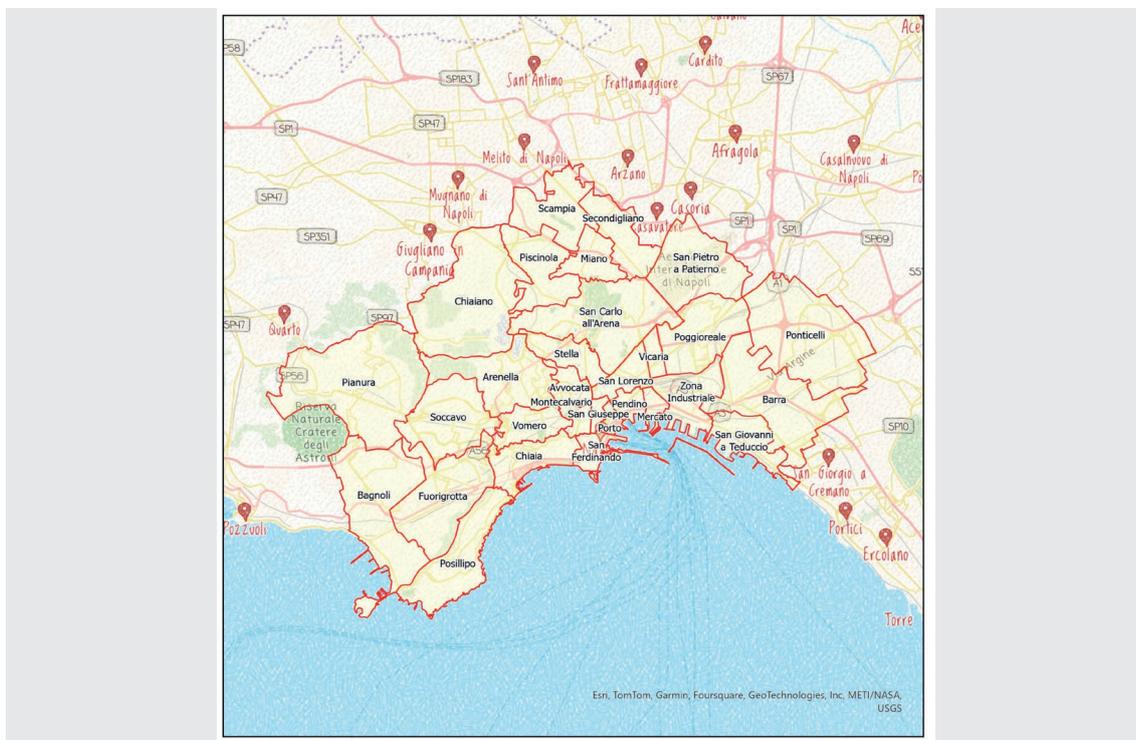
Fonte: Istat, integrazione dei dati del Censimento della Popolazione e delle abitazioni con informazioni provenienti da archivi amministrativi
 (a) Rapporto tra il numero di famiglie con figli la cui persona di riferimento ha fino a 64 anni e nelle quali nessun componente è occupato o percettore di pensione da lavoro, e il totale delle famiglie, moltiplicato per cento.

Figura 14.5 - Le 155 Zone urbanistiche (ZU) del Comune di Roma, all'esterno e all'interno del Grande Raccordo Anulare. Anno 2021



Fonte: Istat, Basi territoriali

Figura 14.6 - I 30 Quartieri del Comune di Napoli. Anno 2021



Fonte: Istat, Basi territoriali

GLOSSARIO

Circondario: ente territoriale di livello intermedio tra la provincia e il comune, previsto e disciplinato dall'ordinamento degli enti locali dal 1861 al 1927. Il capoluogo del circondario era sede di sottoprefettura, di tribunale, di catasto e di uffici finanziari.

Compartimento: ripartizione territoriale del Regno d'Italia elaborata dalla Divisione di statistica del Ministero di agricoltura, industria e commercio, fondata, anziché su criteri politici, su ragioni topografiche e di omogeneità statistica delle province.

Comune: ente pubblico territoriale rappresentativo di una collettività stanziata su un determinato territorio, dotato di propria autonomia statutaria, normativa e amministrativa nell'ambito dei propri confini e secondo i limiti fissati dalla Costituzione (art. 114) e dalla normativa nazionale (Testo unico degli enti locali). I confini amministrativi vengono elaborati dall'Istat in occasione dei periodici censimenti generali della popolazione e validati dai singoli comuni in accordo con quelli confinanti. L'elenco dei comuni viene costantemente aggiornato in funzione delle variazioni che occorrono nel tempo.

Distretto: ente territoriale di livello intermedio tra la provincia e il comune. Al pari dei circondari, tali divisioni contraddistinsero la ripartizione dei territori annessi dall'Impero asburgico fino alla loro soppressione e conseguente trasformazione in circondari nel 1915.

Edificio: Una costruzione generalmente di concezione ed esecuzione unitaria, che può avere uno o più piani, con le seguenti caratteristiche: dotata di una propria struttura indipendente; contenente spazi utilizzabili stabilmente da persone per uso residenziale (alloggi) e/o per la produzione di beni e servizi (uffici, studi, laboratori eccetera); delimitata da pareti, esterne o divisorie, e da coperture; dotata di almeno un accesso dall'esterno.

Georeferenziazione: il processo di associare dati, come immagini, mappe o elementi geografici (ad esempio edifici o indirizzi), a specifiche coordinate spaziali in un sistema di riferimento geografico. Questo permette di posizionare gli oggetti sulla superficie terrestre in modo accurato, rendendoli utilizzabili all'interno di un Sistema informativo geografico (GIS) per l'analisi spaziale e la visualizzazione su mappe digitali.

LUCAS (*Land use and cover area frame survey*): indagine campionaria standardizzata e organizzata a livello continentale, aggiornata ogni tre anni a partire dal 2006.

NDVI (*Normalised vegetation index*): indice calcolato secondo il rapporto tra bande di una immagine multispettrale come indicato di seguito:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

dove **NIR** è la Banda spettrale dell'infrarosso vicino e **RED** rappresenta la Banda spettrale del rosso.



Provincia: ente territoriale di livello intermedio riconosciuto dall'articolo 114 della Costituzione italiana, costituito da un aggregato di comuni limitrofi, il più importante dei quali ne costituisce il capoluogo. Dal 2015, per effetto della normativa contenuta nella legge 7 aprile 2014, n. 56, le province, al pari delle città metropolitane, sono riconosciute enti di area vasta di secondo livello.

Regione: ente pubblico a rilevanza costituzionale (art.114), rappresentativo di una collettività stanziata su un determinato territorio, dotato di propri poteri e funzioni e di un ordinamento autonomo nei limiti prefissati dalla Costituzione e dalle leggi costituzionali. Nel 1946, la Costituzione ha previsto l'esistenza di 19 regioni, divenute 20 nel 1963 a seguito dello scorporo degli Abruzzi e Molise nelle attuali regioni dell'Abruzzo e del Molise; cinque sono dotate di uno statuto speciale di autonomia (Costituzione, art. 116) e una di queste, il Trentino-Alto Adige, è costituita dalle uniche due province autonome: Trento e Bolzano/*Bozen*, dotate di poteri legislativi analoghi a quelli delle regioni a statuto speciale. Nel rispetto delle minoranze linguistiche, il Trentino-Alto Adige e la Valle d'Aosta sono riportati con le denominazioni bilingue di Trentino-Alto Adige/*Südtirol* e Valle d'Aosta/*Vallée d'Aoste*.

Sistema informativo geografico - GIS (*Geographic information system*): insieme ordinato di hardware e software, dati geografici e risorse umane destinate a rilevare, inserire, aggiornare, manipolare, analizzare e visualizzare dati georeferenziati, associando a ciascun elemento geografico una o più descrizioni alfanumeriche.

***Spatial join*:** operazione utilizzata nei Sistemi informativi geografici (GIS) per combinare due insiemi di dati spaziali basandosi sulla loro posizione geografica. Consente di associare attributi da un *layer* (ad esempio, edifici) a un altro *layer* (ad esempio, confini amministrativi), confrontando le loro geometrie e sovrapponendo i dati. In questo modo, è possibile arricchire le informazioni di un *dataset* con quelle provenienti da un altro, in base alla loro vicinanza o contenimento spaziale.

Stato: ente territoriale primario e originario, destinatario di norme di diritto internazionale, autonomo e indipendente da ogni altra entità politica la cui sovranità si estende alla popolazione che insiste su un determinato territorio.

Territorio estero: area geografica su cui si estende tutto o in parte la sovranità di uno Stato nazionale, verso cui sussistono legami storici di dipendenza.

Unità abitativa - Abitazione: singola entità residenziale all'interno di un edificio, progettata per ospitare uno o più individui. Una unità abitativa può essere una casa indipendente o un appartamento all'interno di un edificio con più abitazioni. Include gli spazi interni utilizzati per vivere, dormire, cucinare e svolgere altre attività domestiche. Le unità abitative sono generalmente identificate e registrate nel catasto e nei Censimenti, e possono essere associate a specifici indirizzi e proprietà.

Unità amministrativa territoriale (Ua): le unità territoriali di natura amministrativa previste e regolamentate dalla Costituzione, di qualsiasi livello gerarchico, a partire dalla unità amministrativa locale che corrisponde al comune, fino ad arrivare allo Stato. Con tale termine si identificano, inoltre, gli enti locali indicati dal Testo unico degli enti locali, nella fattispecie unioni di comuni e/o comunità montane.

Unità non amministrative: unità territoriali di livello intermedio tra il comune e la regione, valide ai soli fini statistici, corrispondenti alle ex province del Friuli-Venezia Giulia cessate secondo le modalità espresse con legge regionale 20 dicembre 2016, n. 20.

Unità statistica territoriale (Us): si comprendono nella dicitura le unità territoriali sia di natura amministrativa sia quelle costituite per finalità statistiche nazionali ed europee, nonché funzionali alla rappresentazione di fenomeni specifici.

Unità territoriale: si comprendono nella dicitura le unità territoriali sia di natura amministrativa sia quelle costituite per finalità statistiche nazionali ed europee, nonché funzionali alla rappresentazione di fenomeni specifici.

Unità territoriale sovracomunale (Uts): unità territoriale di livello intermedio tra il comune e la regione. Tali unità, elaborate a fini statistici, a partire dal 2015 ricomprendono, oltre le province, anche le Città metropolitane, i liberi consorzi e le unità non amministrative (ex province del Friuli-Venezia Giulia).

Variazioni amministrative e territoriali dei comuni: le prime riguardano l'istituzione, la soppressione, il cambio di appartenenza alla provincia e/o regione e il cambio di denominazione di un comune. Le variazioni territoriali consistono nello scambio di territori tra due o più comuni che innescano due relazioni reciproche: acquisizione di territorio e cessione di territorio. Tutte le variazioni sono avvalorate da un provvedimento amministrativo che le sancisce (per esempio, legge regionale), pubblicato su un documento istituzionale ufficiale (per esempio, gazzetta, bollettino regionale).

RIFERIMENTI BIBLIOGRAFICI

- Abbatini, D., T. Clary, R. Chiocchini, D. Fardelli, A. Ferruzza, L. Franconi, F. Lipizzi, S. Lucchetti, S. Mugnoli, E. Orsini, A. Pagano, A. Sabbi, G. Salvucci, A. Sera, and P. Ticca. 2024. "Statistical register of places: opportunities for sustainable and climate change related indicators". *Rivista italiana di Economia Demografia e Statistica/ Italian Journal of Economic, Demographic and Statistical Studies*, Volume LXXVIII, N. 1: 101-112. <https://www.rieds-journal.org/rieds/issue/view/61>.
- Agenzia per l'Italia Digitale - AGID. 2020. *Guida operativa per la compilazione dei metadati RNDT – v. 3.0*. Roma: AGID. https://geodati.gov.it/geoportale/images/struttura/documenti/Manuale-RNDT_2-guida-operativa-compilazione-metadati_v3.0.pdf.
- Aracri, R. M., R. Radini, M. Scannapieco, and L. Tosco. 2017. "Using Ontologies for Official Statistics: The ISTAT Experience". In Garrigós, I., and M. Wimmer (eds.) *Current Trends in Web Engineering*. ICWE 2017. Lecture Notes in Computer Science. Volume 10544: 166-172. Cham: Springer.
- Australian Bureau of Statistics - ABS. 2018. "Geocoding Unit Record Data Using Address and Location". *Statistical Spatial Framework Guidance Material*. Website. Canberra, Australia: ABS. <https://www.abs.gov.au/statistics/statistical-geography/statistical-spatial-framework-ssf/statistical-spatial-framework-guidance-material>.
- Borrelli, F., A. Chieppa, S. Di Domenico, G. Gallo, S. Rosati, e V. Tomeo. 2016. *Primi risultati della sperimentazione condotta su fonti amministrative capaci di valutare i segnali di dimora abituale in Italia e l'individuazione di sottopopolazioni critiche*. Istat working papers, N. 23/2016. <https://www.istat.it/it/archivio/210582>.
- Carlucci, R. 2004. "Digitalizzazione del Catasto, nuovi rilasci". Sito web *GEO for ALL*. Roma: mediaGEO. <https://rivistageomedia.it/rilievo-e-localizzazione/digitalizzazione-del-catasto-nuovi-rilasci>.
- Casale, D., M.P. Di Maio, e G. Vaste. 2022. "Tirocini informatici online in Istat: esperienze di learning-by-doing di coding e di utilizzo di strumenti low-code". In Marafioti, R.S., P. Ciancarini, P. Ravotto, e M. Gentile (a cura di). *Atti Convegno Nazionale Didamatica 2022*: 311-318. Milano: Associazione Italiana per l'Informatica e il Calcolo Automatico - AICA. <https://www.aicanet.it/documents/10776/4555506/ATTI+Didamatica+2022/469726b1-58d5-4f48-8a8f-0fa1d2fb0f34>.
- Chiocchini, R., S. Mugnoli, e A. Sabbi. 2023. "Indici per la descrizione morfometrica del territorio italiano". In *Atti della conferenza ESRI Italia 2023*. Roma: ESRI Italia. https://www.esriitalia.it/media/sync/Stefano_Mugnoli_ISTAT.pdf.
- Choi, I. 2021. "Geospatial view of Generic Statistical Business Process Model – GeoGSBPM". *UNECE Statswiki*. <https://statswiki.unece.org/display/GSBPM/GeoGSBPM>.
- Commissione parlamentare di inchiesta sulle condizioni di sicurezza e sullo stato di degrado delle città e delle loro periferie. 2024. *Sicurezza e stato di degrado delle città e delle loro periferie*. Audizione parlamentare del 9 gennaio 2024. Roma: Camera dei deputati. <https://www.istat.it/audizioni/sicurezza-e-stato-di-degrado-delle-citta-e-delle-loro-periferie/>.
- Console, M., D. Lembo, V. Santarelli, and D.F. Savo. 2014. *Graphol: Ontology Representation through Diagrams*. Roma: Sapienza Università di Roma. <https://cs.unibg.it/savo/papers/CLSS-DL-14.pdf>.
- D'Orazio, M., M. Di Zio, and M. Scanu. 2006. *Statistical Matching: Theory and Practice*. Chichester, UK: John Wiley & Sons.
- De Fausti, F., M. Di Zio, G. Lancioni, S. Mugnoli, A. Sabbi, and F. Sisti. 2024. "Data science at Istat for urban green". In Plaia, A., L. Egidi, and A. Abbruzzo (eds.). *Proceedings of the SDS 2024 Conference - New perspectives on Statistics and Data Science*. Palermo: Università degli Studi di Palermo. <https://iris.unipa.it/retrieve/7461fea8-4e45-4346-9520-2a480ae6ee5a/SDS%202024%20Giada%2C%20Giada%2C%20Marcello.pdf>.

- De Giacomo, G., D. Lembo, M. Lenzerini, A. Poggi, and R. Rosati. 2018. "Using Ontologies for Semantic Data Integration". In Flesca, S., S. Greco, E. Masciari, and D. Saccà (eds.). *A Comprehensive Guide Through the Italian Database Research Over the Last 25 Years*. Studies in Big Data, Volume 31: 187-202. Cham: Springer.
- Di Pede, F. 2018. *Istruzioni per l'ordinamento ecografico*. Roma: Istat.
- Dijkstra, L., H. Poelman, and P. Veneri. 2019. *The EU-OECD definition of a functional urban area*. OECD Regional Development Working Papers, N. 11/2019. Paris: OECD Publishing. <https://doi.org/10.1787/d58cb34d-en>.
- Đuračiová R. 2023. "An Aggregated Shape Similarity Index: A Case Study of Comparing the Footprints of OpenStreetMap and INSPIRE Buildings". *ISPRS International Journal of Geo-Information*, Volume 12, N. 495. <https://doi.org/10.3390/ijgi12120495>.
- Egenhofer, M.J., and R.D. Franzosa. 1991. "Point-Set Topological Spatial Relations". *International Journal for Geographic Information Systems*, Volume 5, N. 2: 161-174. <https://doi.org/10.1080/02693799108927841>.
- European Commission - EC, INSPIRE Maintenance and Implementation Group - MIG. 2024. D2.8.III.1 *Data Specification on Statistical Units – Technical Guidelines*. Bruxelles: EC - MIG. https://knowledge-base.inspire.ec.europa.eu/publications/inspire-data-specification-statistical-units-technical-guidelines_en.
- European Environment Agency - EEA. 2006. *Urban Sprawl in Europe - The ignored challenge*. EEA report, N. 10/2006. Copenhagen: EEA. https://www.eea.europa.eu/publications/eea_report_2006_10/eea_report_10_2006.pdf/view.
- Fardelli, D., e S. Mugnoli. 2024. *Nota Metodologica per l'elaborazione della distribuzione della popolazione censuaria per griglia regolare rilasciata ad EUROSTAT*. Roma: Istat. <https://www.istat.it/wp-content/uploads/2023/07/NotaMetodologicaGriglia2021-Ind.pdf>.
- Ferrara, A., F. Lipizzi, e S. Mugnoli. 2017. "Livelli, dinamica e forme del consumo del suolo nelle principali realtà urbane - anni 1991-2011". In Istituto Nazionale di Statistica - Istat. *Forme livelli e dinamiche dell'urbanizzazione in Italia*: 188-213. Roma: Istat. <https://www.istat.it/it/files/2017/05/Urbanizzazione.pdf>.
- Gallego, F. J. 2010. "A population density grid of the European Union". *Population and Environment*, Volume 31, N. 6: 460-479.
- Haldorson, M. 2019. "High demand for local area level statistics - How do National Statistical institutes respond?". *Regional Statistics*, Volume 9, N. 1: 168-186.
- Hornik, K. 2005. "A CLUE for CLUster Ensembles". *Journal of Statistical Software*, Volume 14, N. 12: 1-25. <https://www.jstatsoft.org/article/view/v014i12>.
- Kuhn, H. W. 1955. "The Hungarian method for the assignment problem". *Naval Research Logistics Quarterly*, Volume 2, N. 1-2: 83-97.
- Istituto Nazionale di Statistica - Istat. 2024a. *Statistiche sulla popolazione per griglia regolare*. Notizia, data aggiornamento 10 settembre 2024. Roma: Istat. <https://www.istat.it/notizia/statistiche-sulla-popolazione-per-griglia-regolare/>.
- Istituto Nazionale di Statistica - Istat. 2024b. *Basi territoriali e variabili censuarie*. Notizia, data aggiornamento 10 luglio 2024. Roma: Istat. <https://www.istat.it/notizia/basi-territoriali-e-variabili-censuarie/>.
- Istituto Nazionale di Statistica - Istat. 2024c. *Rapporto SDGs 2024. Informazioni statistiche per l'Agenda 2030 in Italia*. Roma: Istat. <https://www.istat.it/produzione-editoriale/rapporto-sdgs-2024/>.
- Istituto Nazionale di Statistica - Istat. 2024d. *Dati del Censimento Permanente al 31.12.2021 sulle nuove Basi Territoriali 2021*. Nota metodologica. Roma: Istat. <https://www.istat.it/wp-content/uploads/2023/06/Nota-Metodologica-sezioni-censimento-2021.pdf>.

- Istituto Nazionale di Statistica - Istat. 2024e. *Censimento permanente 2021: caratteristiche delle abitazioni*. Statistiche today. Roma: Istat. <https://www.istat.it/comunicato-stampa/censimento-permanente-2021-caratteristiche-delle-abitazioni/>.
- Istituto Nazionale di Statistica - Istat. 2023. *Rapporto SDGs 2023. Informazioni statistiche per l'Agenda 2030 in Italia*. Roma: Istat. <https://www.istat.it/produzione-editoriale/rapporto-sdgs-2023/>.
- Istituto Nazionale di Statistica - Istat. 2022. *Nota tecnica sulla produzione dei dati del Censimento Permanente: la popolazione residente per genere, età, cittadinanza e grado di istruzione al 31.12.2021*. Nota metodologica. Roma: Istat. <https://www.istat.it/it/files/2022/12/Nota-metodologica-censipop-2021.pdf>.
- Istituto Nazionale di Statistica - Istat. 2018. *Struttura e dinamica delle unità amministrative territoriali italiane. Dall'Unificazione del Regno al 2017*. Annali di statistica, Anno 147, Serie XIII, Volume 1. Roma: Istat. <https://www.istat.it/produzione-editoriale/struttura-e-dinamica-delle-unita-amministrative-territoriali-italiane/>.
- Istituto Nazionale di Statistica - Istat. 2016a. *Il programma di modernizzazione dell'Istat*. Roma: Istat. https://www.istat.it/it/files/2010/12/Programma_modernizzazione_Istat2016.pdf.
- Istituto Nazionale di Statistica - Istat. 2016b. *Il Sistema Integrato dei Registri nell'ambito del processo di modernizzazione. Versione 0.1 dell'8 febbraio 2016*. Documento tecnico interno. Roma: Istat.
- Istituto Nazionale di Statistica - Istat. 2015a. *La nuova geografia dei sistemi locali*. Letture statistiche - Territorio. Roma: Istat. <https://www.istat.it/produzione-editoriale/la-nuova-geografia-dei-sistemi-locali/>.
- Istituto Nazionale di Statistica - Istat. 2015b. *I distretti industriali 2011. 9° Censimento dell'industria e dei servizi e Censimento delle istituzioni non profit*. Roma: Istat. <https://www.istat.it/produzione-editoriale/i-distretti-industriali-2011/>.
- Istituto Nazionale di Statistica - Istat. 2011. *Nota tecnica per l'elaborazione della distribuzione della popolazione censuaria per griglia regolare*. Sito web Istat. Roma: Istat. https://www.istat.it/wp-content/uploads/2015/04/Nota_Tecnica.pdf.
- Istituto Nazionale di Statistica - Istat. 2001. *Variazioni territoriali e di nome dal 1861 al 2000. Popolazione legale per comune ai censimenti dal 1861 al 1991 ai confini dell'epoca*. Roma: Istat. https://ebiblio.istat.it/digibib/Unit%C3%A0%20amministrative%20e%20comuni/IST0036751Unita_amministrative_Variazioni_territoriali_di_nome1861_2000+OCRottimizz.pdf.
- Istituto Nazionale di Statistica - Istat. 1992. *Anagrafe della Popolazione*. Metodi e Norme, Volume 5, N. 2: 161-174. Roma: Istat.
- Laaribi, A., and L. Peters. 2019. *GIS and the 2020 Census: Modernizing Official Statistics*. Redlands, CA, U.S.: Esri Press.
- Lenzerini, M. 2018. "Managing Data Through the Lens of an Ontology". *AI Magazine*. Volume 39, N. 2: 65-74.
- Lenzerini, M. 2011. "Ontology-based Data Management". In B. Berendt, A. de Vries, W. Fan, C. Macdonald, I. Ounis, and I. Ruthven. *CIKM '11: Proceedings of the 20th ACM international conference on Information and knowledge management*: 5-6. New York, NY, U.S.: Association for Computing Machinery <https://doi.org/10.1145/2063576.2063582>.
- Lipizzi, F. 2013. *Innovazioni di processo e di prodotto nelle fasi di aggiornamento delle basi territoriali 2010-2011*. Istat working papers, N. 2/2013. Roma: Istat. <https://www.istat.it/produzione-editoriale/innovazioni-di-processo-e-di-prodotto-nelle-fasi-di-aggiornamento-delle-basi-territoriali-2010-2011/>.
- Lipizzi, F., e S. Mugnoli. 2021. "Profili e dinamiche delle località abitate in Italia". *Nuova Verde Ambiente*, Anno I, N. 2: 27-32. Roma: Verdi Ambiente e Società - APS Onlus. <https://verdiambientesocieta.eu/WP2023/wp-content/uploads/2023/03/NVA-2-2021.pdf>.
- Lipizzi, F., e S. Mugnoli. 2017. "Profili e Dinamiche delle Località abitate in Italia". In Istituto Nazionale di

- Statistica - Istat. *Forme, livelli e dinamiche dell'urbanizzazione in Italia*: 39-58. Roma: Istat. <https://www.istat.it/it/files/2017/05/Urbanizzazione.pdf>.
- Liu, L., D. Xiaohui, Z. Xinyan, F. Liang, and G. Jun. 2020. "An Iterative Approach Based on Contextual Information for Matching Multi scale Polygonal Object Datasets". *Transactions in GIS*, Volume 24, N. 4: 1047-1072. <https://doi.org/10.1111/tgis.12625>.
- Mugnoli, S., D. Abbatini, R. Chiocchini, e F. Lipizzi. 2016. "La cartografia Istat come supporto per le emergenze territoriali". *GEOmedia*, Volume 20, N. 5. <https://mediageo.it/ojs/index.php/GEOmedia/article/view/1363/1245>.
- Mugnoli, S., and R. Chiocchini. 2014. "Istat data utilization to enhance LANDSAT 8 images classification process". *GEOmedia*, Volume 18, N. 5: 42-46. <https://mediageo.it/ojs/index.php/GEOmedia/article/view/1013>.
- Mugnoli, S., R. Chiocchini, S. Cruciani, A. Esposto, e F. Lipizzi. 2011. "Integrazione di dataset geografici di copertura del Suolo e Censuari per la realizzazione di una mappa statistica sintetica". In *Atti della 15ª Conferenza Nazionale ASITA*: 1633-1640. Milano: Federazione ASITA. <http://atti.asita.it/ASITA2011/Pdf/099.pdf>.
- Mugnoli, S., R. Chiocchini, A. Esposto, F. Lipizzi, G. Lombardo, and R. Minguzzi. 2014. *Land Cover and Census integration geographic datasets to realize a statistics synthetic map*. Paper presented at the *European Forum for Geography and Statistics*. Krakow, 22-24 October.
- Mugnoli, S., F. Lipizzi, and A. Esposto. 2018. "New ISTAT 'microzones' layer: a new way to read land cover statistics". *READING. Journal of Research and Didactics in Geography*, Volume 2: 95-104. <http://www.j-reading.org/index.php/geography/article/view/216>.
- Mugnoli, S., A. Sabbi, and F. Lipizzi. 2023. "The Land Cover/Use Code of the new Istat Census cartography". In di Bella, E., L. Fabbris, and C. Lagazio (eds.). *ASA 2022 Data-Driven Decision Making. Book of short papers*: 119-124. Firenze and Genova: Firenze University Press and Genova University Press. <https://doi.org/10.36253/979-12-215-0106-3>.
- Papadimitriou, C. H., and K. Steiglitz. 1982. "Combinatorial Optimization: Algorithms and Complexity". *IEEE Transactions on Acoustics, Speech and Signal Processing*, Volume 32, N. 6: 1258-1259.
- R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Regolamento (UE) 2023/138 della Commissione del 21/12/2022, che stabilisce un elenco di specifiche serie di dati di elevato valore e le relative modalità di pubblicazione e riutilizzo . (Testo rilevante ai fini del SEE). *Gazzetta ufficiale dell'Unione europea*, L 19, 20 gennaio 2023: 43-75. <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32023R0138>.
- Scanu, M. 2003. *Metodi statistici per il record linkage*. Metodi e norme, N. 16/2003. Roma, Italia: Istat.
- United Nations, Department of Economic and Social Affairs - UN. 2019. *The Global Statistical Geospatial Framework*. New York, NY, U.S.: UN. https://unstats.un.org/unsd/statcom/51st-session/documents/The_GSGF-E.pdf.
- United Nations, Department of Economic and Social Affairs - UN. 2009. *Handbook on Geospatial Infrastructure in Support of Census Activities*. Studies In Methods. Series F, N.103. New York, NY, U.S.: UN. <https://doi.org/10.18356/30560942-en>.
- United Nations, Statistics Division - UNSD. 2019. *The Global Statistical Geospatial Framework*. New York, NY, U.S.: United Nations. https://unstats.un.org/unsd/statcom/51st-session/documents/The_GSGF-E.pdf.
- United Nations Economic Commission for Europe - UNECE. 2024. *Guidance on the role of national statistical offices in achieving national climate objectives*. Geneva: UNECE. https://unece.org/sites/default/files/2024-04/2024_EC3_Guidance%20on%20the%20role%20of%20NSOs%20in%20achieving%20climate%20objectives.pdf.

- United Nations Economic Commission for Europe - UNECE. 2023. *Measuring hazardous events and disasters: set of core disaster-risk-related indicators*. Geneva: UNECE. https://unece.org/sites/default/files/2023-05/CES%202023%204%20E_0.pdf.
- United Nations Economic Commission for Europe - UNECE. 2021a. *Implementation Guidelines for the Conference of European Statisticians' Set of Core Climate Change-related Indicators and Statistics Using the System of Environmental-Economic Accounting*. Geneva: UNECE. <https://unece.org/statistics/publications/implementation-guidelines-CES-core-set-CCRSI>.
- United Nations Economic Commission for Europe - UNECE. 2021. *Guidelines for Assessing the Quality of Administrative Sources for Use in Censuses*. Geneva: United Nations. <https://unece.org/statistics/publications/CensusAdminQuality>.
- United Nations Economic Commission for Europe - UNECE. 2014. *Recommendations on Climate Change-Related Statistics*. Geneva: UNECE. <https://unece.org/statistics/publications/recommendations-climate-change-related-statistics>.
- United Nations Economic and Social Commission for Asia and the Pacific - UNESCAP. 2018. *Disaster-related Statistics Framework (DRSF)*. Bangkok, Thailand: UNESCAP. https://www.unescap.org/sites/default/d8files/event-documents/%21DRSF%20Manual_20220620_0.pdf.
- United Nations Office for Disaster Risk Reduction - UNDRR, and International Science Council - ISC. 2021. *Hazard Information Profiles. Supplement to UNDRR-ISC Hazard Definition & Classification Review - Technical Report*. Geneva: UNDRR. <https://www.undrr.org/publication/hazard-information-profiles-hips>.
- Vaste, G., M.P. Di Maio, P. Giorgetti, M. Passacantilli, e F. Dell'Orco. 2021. *Il sistema integrato dei registri (SIR): sperimentazione su strumenti di controllo di processo*. Poster presentato alla XIV Conferenza Nazionale di Statistica. Roma: Istat. <https://www.istat.it/storage/14-Conferenza-nazionale-statistica/poster/030.jpg>.
- Wallgren, A., and Wallgren, B. 2014. *Register-based Statistics. Statistical Methods for Administrative Data. Second Edition*. Chichester, UK: Wiley & Sons.

