

LA DIFFUSIONE DEL CENSIMENTO PERMANENTE DELLA POPOLAZIONE E DELLE  
ABITAZIONI TRA TRADIZIONE E INNOVAZIONE

Mariangela Verrascina<sup>1</sup>

**SOMMARIO**

Il Censimento della popolazione e delle abitazioni non sarà più un censimento tradizionale, decennale ed esaustivo, ma continuo, annuale e campionario. Il presente lavoro illustra la nuova strategia per il censimento permanente, che prevede l'uso di dati provenienti da fonti amministrative e da indagini campionarie, allo scopo di integrare le informazioni e di produrre le statistiche censuarie richieste a livello europeo e di interesse nazionale, con maggior frequenza e tempestività rispetto al passato. La diffusione ha subito evoluzioni dal 2001 al 2011 e ulteriori se ne prevedono per il futuro. La nuova impostazione, infatti, comporta una rimodellazione della diffusione, con un rilascio dei dati con cadenza maggiore e innovazione negli strumenti per una divulgazione tempestiva, al passo con i tempi e che soddisfi le esigenze degli utilizzatori finali. Nel lavoro si presenta la proposta di diffusione, in fase di progettazione, per il censimento permanente sul piano dei contenuti e su quello degli strumenti di pubblicazione. Tra gli obiettivi della produzione di statistiche censuarie c'è quello di ampliare l'offerta informativa e di facilitare l'accesso ai dati con livello territoriale minimo per pianificare, programmare, avviare e valutare politiche sociali e territoriali.

---

<sup>1</sup> Istat, Via Cesare Balbo 16, 00184 Roma, e-mail: [verrasci@istat.it](mailto:verrasci@istat.it).

## 1. Introduzione

Il censimento generale della popolazione e delle abitazioni italiane del 2011 è stato l'ultimo realizzato in modo tradizionale. Diversi sono gli elementi che portano a una riprogettazione delle statistiche censuarie. Da una parte c'è una crescente necessità da parte degli utenti di dati di censimento più frequenti e più tempestivi, dall'altra ci sono il budget ridotto stanziato per le rilevazioni e, soprattutto, la considerevole quantità di dati provenienti da fonti amministrative. La ridefinizione delle modalità di rilevazione e di diffusione censuaria è stata avviata anche a livello internazionale. I sistemi che consentirebbero una produzione più tempestiva di dati sono quelli fondati sui registri o gli approcci combinati, basati su fonti amministrative e supportati da indagini per raccogliere altre informazioni. Sono molti i paesi che dispongono di archivi amministrativi e ne prevedono un maggiore utilizzo nel prossimo censimento, che avrà come anno di riferimento il 2021. La legislazione lascia la libertà ai singoli paesi di scegliere le modalità di rilevazione più appropriate alle realtà nazionali, numerosi sono quelli che si stanno orientando verso un censimento combinato, che si aggiungono a quelli che già nei censimenti precedenti erano passati a un censimento basato sui registri. Anche in Italia la direzione va verso l'uso delle fonti amministrative disponibili e l'integrazione con le indagini campionarie per raccogliere informazioni non contenute negli archivi, necessarie per la produzione censuaria richiesta a livello europeo e di interesse nazionale. Così, il censimento italiano non sarà più tradizionale, cioè esaustivo e decennale, ma un censimento che si svolgerà ogni anno su un campione di comuni e famiglie, diverso di anno in anno. Cambia forma e cambia nome, diventando un censimento permanente. La nuova strategia richiederà qualche anno per la messa a regime, ma viene avviata nel 2018, con la prima rilevazione annuale campionaria prevista per ottobre.

La diffusione ha subito evoluzioni dal 2001 al 2011 e ulteriori se ne prevedono passando al censimento futuro, sia in Europa che in Italia. Nel primo caso, le modifiche si basano sull'analisi del lavoro svolto per i censimenti precedenti e sulle richieste da parte degli stakeholders su aspetti prioritari che solo la rilevazione censuaria può soddisfare e sono previste allo scopo di alleggerire sempre più il carico di lavoro sugli Istituti di statistica degli Stati Membri. A livello nazionale, le evoluzioni tengono conto delle modifiche nelle fonti (dati di input), della nuova impostazione censuaria, delle novità internazionali, degli strumenti adottati in passato in altri paesi e delle tecnologie ad oggi disponibili. Nel presente lavoro si illustrerà la strategia di pubblicazione che si sta progettando per il censimento permanente sul piano dei contenuti e su quello degli strumenti di divulgazione degli output. Si vedrà come la prossima diffusione sarà, ancora più di prima, volta al soddisfacimento delle esigenze degli utilizzatori dei dati (decisori, policy maker, specialisti); sarà orientata non solo alla ricerca e alle analisi a livello territoriale micro ma anche a un maggiore e più facile accesso ai dati necessari alla programmazione, all'avvio delle politiche e alla valutazione dei risultati.

Si comincia con una panoramica sulla produzione dei risultati, sul fronte italiano ed europeo, del 2001 e del 2011 (paragrafo 2), si presentano anche i risultati di un'analisi, svolta su alcuni paesi esteri, riguardante gli output censuari (paragrafo 3). Si descrivono poi le novità previste per il 2021 a livello europeo (paragrafo 4), la strategia futura in Italia, presentando la nuova impostazione del censimento permanente (paragrafo 5), il piano di diffusione che si sta predisponendo (paragrafo 6), come potranno essere diffusi in futuro i risultati censuari, seguendo un percorso che combina tradizione e innovazione (paragrafo 7) e le considerazioni di sintesi (paragrafo 8).

## 2. La tradizionale diffusione dei dati censuari: l'esperienza del 2001 e del 2011

La diffusione dei risultati del Censimento del 2001 rispetto al passato era stata innovativa pur mantenendo elementi della tradizione. La novità riguardava la disponibilità dei dati via *internet*, oltre alle consuete pubblicazioni cartacee. Era stato creato un *datawarehouse* (Istat, 2001) progettato appositamente per rendere accessibili *online* tavole statistiche con dati aggregati; l'utente aveva la possibilità di navigare tra le tavole in modo flessibile, di scegliere il dettaglio territoriale e un livello classificatorio tra quelli disponibili e di esportare gli incroci costruiti sul proprio pc. Sul sito erano messi a disposizione degli utenti i *report*

(comunicati stampa, note per la stampa, eccetera) relativi ai rilasci tematici, alcune basi di dati e una cartografia interattiva tematica. Le pubblicazioni cartacee prevedevano i fascicoli territoriali provinciali (Istat, 2005a), regionali (Istat, 2006b) e Grandi Comuni<sup>2</sup> (Istat, 2006c), due volumi nazionali contenenti i risultati definitivi relativi alle variabili demografiche il primo (Istat, 2005b) e alle variabili socio-economiche e alle abitazioni il secondo (Istat, 2006a), un volume tematico con un approfondimento sulla popolazione straniera residente in Italia al 21 ottobre 2001 (Istat, 2006d), un volume sul sistema di rilevazione e sul processo di produzione dei dati (Istat, 2006f), uno con tutta la documentazione predisposta per il 14° censimento generale della popolazione e delle abitazioni (atti normativi, questionari e modelli ausiliari utilizzati in occasione delle due indagini pilota, della rilevazione censuaria e dell'indagine di copertura) (Istat, 2006e), e, infine, uno sulla qualità dei dati (Istat, 2009). Era stata costruita e resa disponibile la “matrice del pendolarismo”, una matrice a livello comunale con informazioni dettagliate sugli spostamenti pendolari per motivi di studio o di lavoro. Erano stati poi predisposti due file per sezione di censimento, uno riservato agli Enti appartenenti al Sistema statistico nazionale (Sistan) con 279 variabili e uno destinato a tutte le categorie di utenti con 205 variabili; erano stati prodotti un “file di record individuali” (file di microdati rilasciabili agli Enti appartenenti al Sistan), un “file standard” (campione all'1% di dati elementari fruibili per fini di studio e di ricerca) e un “file per la ricerca” (campione al 5% realizzato specificatamente per esigenze di ricerca scientifica), con un maggior livello di dettaglio informativo rispetto al file standard. Infine, erano effettuate elaborazioni personalizzate a cura dell'Istituto per soddisfare particolari richieste di informazioni non diffuse *online*, non presenti nei volumi pubblicati, né sui supporti informatici ad essi allegati né tra gli altri prodotti predisposti (Dardanelli *et al.*, 2010).

Durante la produzione dei dati censuari italiani del 2001, era stata avviata la predisposizione delle tavole per la diffusione europea. Eurostat aveva richiesto ai singoli paesi europei la fornitura di 40 tavole statistiche, a livello nazionale, provinciale e comunale, cui erano state aggiunte successivamente due tavole per dettaglio territoriale di *urban areas*<sup>3</sup> (aree urbane). Le indicazioni per la predisposizione delle tavole erano contenute nel *Table Programme for the Community Programme of Population and Housing Censuses in 2001* che aveva rappresentato, a livello europeo, un *Gentlemen's Agreement* tra gli Paesi Membri. Per ogni tavola erano riportati le variabili da incrociare, le classificazioni da associare alle variabili e il dettaglio territoriale cui produrla. Il processo di produzione delle tavole europee è stato oneroso in termini di tempo e di risorse, per le difficoltà incontrate legate ad aspetti definitori e classificatori. Alcune informazioni non erano state rilevate, altre erano state rilevate in maniera differente; le classificazioni richieste, dettagliate e articolate, differivano dalle classificazioni utilizzate in ambito nazionale. Questo ha comportato il calcolo, laddove possibile, di nuove classificazioni, ottenute come trasformazioni di quelle già esistenti o con una ridefinizione a partire dalle variabili di rilevazione, ma non è stata possibile una completa ricostruzione delle tavole richieste e il rispetto dei tempi di consegna delle tavole a Eurostat<sup>4</sup>. Le difficoltà incontrate, non solo dall'Italia ma anche dagli altri paesi, nella produzione dei dati richiesti, le differenze di definizioni, classificazioni e modalità di rilevazione, non hanno garantito, in molti casi, la confrontabilità dei dati.

A differenza dei censimenti del 2001, per la tornata censuaria del 2010-2011, la Commissione Europea ha emanato un *Framework Regulation*, il Regolamento Quadro del Parlamento Europeo e del Consiglio relativo ai censimenti della popolazione e delle abitazioni n. 763/2008 (European Parliament and Council, 2008). Il Censimento europeo della popolazione e delle abitazioni del 2011 è stato il primo esercizio di censimento basato su una legislazione comune che copriva in dettaglio tutti gli aspetti relativi ai risultati del censimento. Per la prima volta, per i dati dei censimenti della popolazione e delle abitazioni condotti nei 32 paesi dell'UE e dell'EFTA<sup>5</sup>, è stato adottato un approccio di armonizzazione dei risultati (*output oriented*): i paesi erano liberi di produrre i dati richiesti utilizzando qualsiasi metodo e fonti di dati ritenuti appropriati in base alla

<sup>2</sup> Con popolazione residente superiore a 150.000 abitanti.

<sup>3</sup> Il concetto di area urbana non coincide con quello di comune (LAU 2 per la classificazione utilizzata da Eurostat) né con quello di area subcomunale, ma prescinde dai confini comunali. Le aree urbane sono state individuate all'interno dei confini regionali, come aggregazioni di sezioni di censimento.

<sup>4</sup> Secondo quanto previsto nell'accordo sottoscritto dai Paesi Membri per i censimenti del 2000-01 (*Gentlemen's agreement*), la trasmissione dei dati ad Eurostat doveva avvenire entro e non oltre il 30 giugno 2003.

<sup>5</sup> I paesi che appartengono all'Associazione Europea di Libero Scambio sono l'Islanda, il Liechtenstein, la Norvegia e la Svizzera.

disponibilità e alle preferenze nazionali (Dardanelli *et al.*, 2008). Tuttavia, i dati europei dovevano soddisfare pienamente gli standard definiti nella legislazione (definizioni, classificazioni, tavole e qualità). Il Regolamento Quadro ha consentito flessibilità agli Istituti nazionali di statistica nella scelta delle fonti di dati (comprese indagini campionarie e fonti amministrative) che potevano essere utilizzate per la raccolta di dati. Il Regolamento n. 763/2008 stabiliva la data entro cui rendere disponibili gli output censuari: gli Stati Membri hanno fornito a Eurostat i dati definitivi del censimento e i relativi metadati entro 27 mesi dalla fine dell'anno di riferimento della rilevazione (31 marzo 2014). Infine il Regolamento riportava, in allegato, le variabili (*core topics*) da includere obbligatoriamente nel piano di rilevazione e, per ciascuna variabile, il livello geografico cui doveva essere prodotta. In un'ottica di standardizzazione dell'output censuario, il Regolamento di attuazione n. 519/2010 (European Commission, 2010), avente per oggetto il programma di diffusione dei dati e i metadati, elencava gli incroci che dovevano essere prodotti e resi disponibili a Eurostat. Le variabili obbligatorie presenti negli ipercubi erano quelle riportate nell'Allegato del Regolamento Quadro; le classificazioni invece erano quelle del Regolamento n. 1201/2009 su *breakdowns* e *technical specifications* (European Commission, 2009). Gli *hypercubes* predisposti per il piano europeo sono stati complessivamente 60 (5 a livello nazionale, 36 a livello regionale, 10 provinciale, 5 comunale e 4 per luogo di lavoro), corrispondenti a 175 distribuzioni marginali principali; il programma di diffusione dei dati focalizzava l'attenzione prevalentemente sul dettaglio regionale. Il regolamento sugli ipercubi definiva anche i metadati testuali da fornire per le variabili di censimento. I dati sono stati prodotti in forma di tavole multidimensionali e accessibili tramite il sistema *online* Census Hub, che ha consentito agli utenti la flessibilità di definire tavole in base alle proprie esigenze, garantendo un facile accesso a una vasta gamma di incroci comparabili tra tutti gli Stati Membri.

Il piano di diffusione definito da Eurostat non corrispondeva al piano di diffusione nazionale; gli incroci contenuti nel programma degli ipercubi europeo, infatti, differivano dagli incroci contemplati nel piano di diffusione italiano per informazioni, classificazioni, definizioni, dettagli classificatori e livelli territoriali. È stato, pertanto, necessario produrre due distinte diffusioni: una europea, secondo quanto stabilito dai Regolamenti descritti in precedenza; l'altra italiana, per garantire la continuità delle serie storiche, la pubblicazione di informazioni relative a fenomeni di interesse nazionale (ad es. gli spostamenti pendolari per studio o lavoro) e relative ad aspetti indagati per la prima volta in Italia in occasione del censimento (ad es. il luogo di nascita della madre e del padre). Il piano di diffusione italiano era composto da più di 300 tavole (Istat, 2017). Per la diffusione dei dati definitivi del XV Censimento generale della popolazione e delle abitazioni non sono state previste pubblicazioni cartacee, la diffusione è avvenuta solo attraverso il web. È stato utilizzato I.Stat (Istat, 2011), il *corporate datawarehouse* statistico dell'Istituto, che ha permesso agli utenti di navigare tra le tavole non rigidamente strutturate, scegliendo le classificazioni e i dettagli territoriali di interesse. I dati sono stati presentati in tavole multidimensionali che gli utenti potevano esportare in diversi formati (Istat, 2016). Sono stati predisposti comunicati stampa che accompagnavano la prima uscita dei dati più rilevanti; sono stati poi diffusi dati per sezioni di censimento attraverso la pagina di basi territoriali, mappe tematiche di confronto con il censimento precedente attraverso BT Carto, la matrice di pendolarismo, dati in formato aperto attraverso LOD (*Linked Open Data*), oltre ai tradizionali prodotti (file standard, campioni di dati per la ricerca). E, infine, particolari richieste da parte dell'utenza sono state evase tramite elaborazioni *ad hoc* a cura dell'Istituto.

### **3. La produzione censuaria di alcuni stati esteri nel 2011**

#### *3.1. Il contesto della ricerca*

I Regolamenti europei emanati tra il 2008 e il 2010 per la tornata censuaria del 2010/2011 hanno comportato che alcuni dati dovessero obbligatoriamente essere messi a disposizione di Eurostat (attraverso il *Census Hub*), mentre per la diffusione italiana era necessario soddisfare il fabbisogno informativo degli utenti nazionali e garantire continuità alle serie storiche. È stata avviata una ricognizione di quanto effettuato

in altri paesi a livello nazionale, oltre alla produzione degli ipercubi sotto regolamento previsti dall'UE, con l'obiettivo di verificare se altri contesti europei avessero avuto necessità di “sdoppiare” la diffusione, così come accaduto in Italia. Pertanto sono stati presi in considerazione alcuni paesi esteri, con differenti approcci censuari, per esplorare la diffusione nazionale relativa al Censimento 2011, in base alla tecnica adottata, e confrontarla con la produzione censuaria italiana (Verrascina, 2017). La disamina della produzione censuaria degli altri paesi è stata effettuata considerando il tipo di censimento adottato nel 2011, la presenza a livello nazionale di pubblicazione di *output* censuario, il tipo di diffusione nazionale, i nuovi contenuti informativi (rispetto a quanto diffuso in Italia) e le modalità di diffusione adottate. La ricognizione è stata condotta analizzando il materiale (pagine web e documenti – comunicati stampa, volumi) disponibile *online* sui siti degli Istituti nazionali di statistica. Contestualmente alla tecnica di rilevazione adottata per l'ultimo censimento, è stato considerato il tipo di diffusione utilizzato per rendere disponibili agli utenti nazionali i risultati censuari. I paesi considerati nell'analisi sono stati classificati per tipo di censimento e suddivisi in: censimento basato su registri (*Register-based census*), censimento combinato (*Combined census*) e censimento tradizionale (*Traditional census*). Nel primo caso si tratta di censimento effettuato utilizzando esclusivamente informazioni contenute nei registri e nelle altre fonti amministrative, nel secondo caso si tratta di un approccio che combina informazioni da registri e informazioni che provengono da indagini (esaustive o campionarie), infine, nel terzo caso, si tratta del metodo tradizionale di enumerazione totale della popolazione. La ricerca ha permesso di esplorare forme e strategie di diffusione alternative a quelle tradizionalmente utilizzate in Italia, oltre che di individuare possibili nuovi contenuti informativi per il prossimo censimento. I paesi oggetto della ricognizione sono stati: Paesi nordici (Danimarca, Norvegia, Finlandia, Svezia), Austria, Slovenia per il *Register-based census*; Paesi Bassi, Spagna, Germania per il *Combined census*; Regno Unito (Inghilterra e Galles), Francia per il *Traditional census*.

Oltre a rendere disponibili i dati a livello europeo attraverso il *Census Hub*, non tutti i paesi diffondono *output* censuario nazionale. Alcuni fanno rientrare le statistiche censuarie nelle statistiche correnti tematiche (Danimarca, Svezia, Paesi Bassi). La Slovenia ha pubblicato solo i dati disponibili negli archivi con una rottura/interruzione delle serie storiche, ovvero senza più garantire la confrontabilità con il passato. In altri paesi, i risultati censuari sono stati ottenuti desumendo le informazioni da archivi e da indagini (laddove disponibili), mentre le informazioni non contenute negli archivi (o comunque di qualità non accettabile), e non rilevate da altre indagini, sono state raccolte attraverso una rilevazione con questionario su un campione e/o su sottopopolazioni non coperte in altro modo (Spagna e Germania e, in passato, i paesi nordici).

### 3.2. Approfondimento su alcuni paesi

Numerosi e differenti sono stati i modi di approcciarsi alla diffusione censuaria da parte dei paesi che producono risultati a livello nazionale (Norvegia, Finlandia, Austria, Spagna, Germania, Regno Unito-Inghilterra e Galles-, Francia). Le differenze riguardano i contenuti informativi, le classificazioni, più o meno dettagliate in base all'argomento, oppure ottenute combinando più informazioni (come risposte a quesiti presenti nel questionario di censimento o derivate da variabili contenute in archivi/registri) e, infine, anche le modalità con cui i risultati censuari sono stati messi a disposizione degli utenti finali.

Dal lato dei contenuti informativi, escludendo informazioni che tradizionalmente non sono indagate in occasione della rilevazione censuaria in Italia (razza, etnia, religione, salute), con la ricognizione sono state individuate poche nuove tematiche/variabili. Più in generale è emerso che informazioni che già tradizionalmente rientrano nel piano di rilevazione italiano possono essere aggregate e diffuse in modo diverso; non si è in presenza di novità di contenuto informativo ma di nuovi incroci che potrebbero essere inseriti nel piano di diffusione dei censimenti futuri. Dalla ricognizione effettuata emerge che, sul versante dei contenuti, l'offerta informativa italiana non ha pari tra i paesi esteri analizzati (fatta eccezione per Inghilterra e Galles) per contenuto, dettaglio classificatorio e dettaglio territoriale.

L'analisi delle tecniche di diffusione dei dati censuari usate da alcuni tra gli stati esteri considerati ha messo in luce modalità interessanti e diversificate di pubblicazione dei risultati che possono essere adottate e che rappresentano un'opportunità di miglioramento e perfezionamento della diffusione per i prossimi

censimenti. I paesi che adottano il *register-based census* hanno una popolazione totale contenuta e ciò, probabilmente, facilita la gestione degli archivi e, di conseguenza, l'integrazione tra le fonti amministrative. Meritano attenzione gli strumenti di diffusione utilizzati da alcuni paesi che per dimensione demografica, tenuta degli archivi, impostazione del censimento e diffusione censuaria, più si avvicinano all'Italia: Spagna, Germania, Inghilterra e Galles e Francia, a cui si aggiunge l'Austria, per il tipo di diffusione adottata. Si descrivono di seguito le caratteristiche per ogni paese.

### 3.2.1. Austria

L'Austria è l'unico paese che, a livello nazionale, non produce tavole ma mette a disposizione un *datawarehouse*. Tutti gli incroci sono possibili all'interno di un database tematico (*Persons, households, Families, Buildings, Conventional dwellings, Conventional dwellings with usual resident*); ogni db contiene anche classificazioni relative ad altri db e questo permette di legare i temi tra loro. Con lo strumento di interrogazione adottato è possibile combinare classificazioni a piacimento, creare gerarchie nelle tavole o rappresentare il risultato finale come figura. Per le tavole, c'è la possibilità di scegliere se visualizzare valori assoluti o percentuali (di riga o di colonna) e di creare classi di valori per evidenziare alcune celle. Ogni db ha la data di pubblicazione e di modifica. Infine, sul sito è riportato l'elenco di variabili che non possono essere più prodotte perché non presenti negli archivi.

### 3.2.2. Spagna

Qualche mese prima del censimento, l'Istituto di statistica spagnolo ha pubblicato un prospetto con la tempistica della diffusione e con le differenze attese rispetto al censimento precedente. Sul sito utilizzato per la diffusione 2011 e 2001 sono stati anche caricati i dati del censimento della popolazione e delle abitazioni del 1991, per facilitare il confronto dei risultati relativi ai tre censimenti. La diffusione dell'output censuario del 2011 è costituita da tavole definite *a priori* (*Pre-defined tables*) e da query personalizzate (*Advanced query systems*). Si tratta, per la seconda parte, di una forma di *datawarehouse* corredato da alcuni alert iniziali su come creare le query, alcune avvertenze su come leggere i dati provenienti dalla rilevazione campionaria e su come interpretare i dati nelle celle con piccoli numeri. Il risultato delle query personalizzate può essere rappresentato come tavola o come mappa. La Spagna ha puntato anche sulla rappresentazione cartografica dei dati; molti risultati sono stati diffusi in forma di mappa disaggregabile fino al più piccolo dettaglio territoriale, ovvero comunale. Sono state predisposte anche mappe con griglie da 1Km<sup>2</sup>. Nella sezione dedicata ai censimenti in Spagna, in primo piano si individuano una nota tecnica sulle differenze tra i dati censuari e i dati da Registro Comunale e alcune indicazioni per facilitare la corretta interpretazione dei dati.

### 3.2.3. Germania

Anche la Germania, nella pubblicazione dei dati relativi al Censimento del 2011, ha utilizzato una molteplicità di strumenti di diffusione. Sono presenti, nella pagina dedicata al censimento, una sezione dedicata alle tavole predefinite (*Easy way to results*), un'ampia sezione dedicata alle mappe (*Results cartographically and visually*) e una sezione dedicata alla creazione di analisi personalizzate (*Dynamic and individual results*). Quest'ultima è, come nel caso della Spagna, una forma di *datawarehouse* con indicazioni iniziali su come interpretare i dati provenienti dalla rilevazione campionaria; è possibile, infatti, effettuare delle query solo sulle famiglie e solo per i comuni con più di 10.000 abitanti. Anche la Germania ha puntato molto sulla rappresentazione cartografica dei dati. Risulta essere ben evidente la pagina dedicata al glossario, oltre alle note in cui si spiegano le differenze tra i numeri prodotti al censimento e i numeri derivanti da archivi o da altre indagini campionarie.

### 3.2.4. Inghilterra e Galles

L'impostazione adottata da Inghilterra e Galles per la diffusione dei risultati censuari del 2011 è improntata alla massima trasparenza e alla massima attenzione agli utenti finali. La strategia di diffusione dei dati censuari, infatti, ha previsto la produzione di tante tavole, tutte in formato xls, facilmente scaricabili. Le

informazioni sul contenuto del piano di diffusione e sulla tempistica dei rilasci sono riportate nel *2011 Census prospectus*, aggiornato costantemente. Sono presenti, inoltre, note tecniche con le differenze rispetto al censimento precedente. Ogni tavola diffusa riporta la data di pubblicazione e la versione del rilascio, oltre all'eventuale data di ricaricamento (l'informazione su ricaricamenti è segnalata anche in una pagina con gli aggiornamenti). Online è presente il volume "Glossario" contenente tutte le definizioni utilizzate nella diffusione. La divulgazione dei risultati è basata sul massimo utilizzo di nuove tecnologie e di nuovi strumenti di diffusione come *data visualisation* e *analyses* per permettere agli utenti di visualizzare i dati e le analisi nel modo preferito. L'ONS ha investito nella promozione del censimento e dei risultati censuari tramite media, autorità locali, mondo accademico e sui social media (Twitter e YouTube). Il rilascio dei dati è avvenuto in più fasi: prima la diffusione delle distribuzioni semplici percentuali e totali, in seguito per ogni variabile sono state pubblicate le classificazioni più estese e, infine, sono stati diffusi incroci sempre più complessi, per variabili coinvolte e per dettaglio classificatorio. La maggior parte delle tavole sono state diffuse a dettaglio territoriale più disaggregato. In aggiunta alle *Key Statistics*, *Quick Statistics*, *Detailed characteristics*, *Local characteristics*, per soddisfare le esigenze degli utenti, sono state pubblicate tavole che hanno oggetti differenti rispetto ai residenti abituali (*Alternative population bases: Workplace population, Workday population, Out-of-term population, Detailed UK Migration Statistics, Ad-hoc Tables*), e tavole che integrano i principali temi (*Specialist census products: Origin-destination data, Microdata, Small population groups*).

### 3.2.5. Francia

I risultati del censimento, calcolati sulle cinque indagini annuali più recenti, e riferiti all'anno medio, sono pubblicati ogni anno e in diverse forme, per soddisfare le esigenze di una varietà di utenti. La documentazione specifica (definizioni, variabili statistiche presenti, formule per calcolare gli indicatori, istruzioni per l'utente, suggerimenti) è inclusa in ogni database tematico; i database sono disponibili in formato Excel e in formato testo. I risultati essenziali del censimento (*Key figures and associated databases*) sono disponibili per livelli geografici dal comunale in su, con la possibilità anche di una ricerca tematica. Alcuni dati sono presentati come andamenti, confrontati con i risultati dell'ultimo censimento tradizionale (1999). Le tavole e i grafici forniscono i risultati principali per una determinata zona sotto forma di conteggi, percentuali e indicatori calcolati. Per le zone con meno di 2.000 abitanti, alcune tavole e grafici non vengono pubblicati. Nelle *Detailed tables and associated databases* sono contenuti dati aggiuntivi rispetto alle Key figures, sui temi *Famiglie e Nuclei*, *Formazione*, *Occupazione - forze lavoro*, *Nazionalità e immigrazione*, *Popolazione e luogo di residenza precedente*, *Abitazioni e Residenze Principali*; sono disponibili per aree con almeno 2.000 abitanti e il dettaglio delle tavole aumenta all'aumentare dell'ampiezza demografica di una zona, questo criterio garantisce la riservatezza e l'affidabilità dei risultati pubblicati. Le soglie di popolazione che consentono l'accesso a ulteriori livelli di dettaglio sono fissate a 10.000, 50.000 e 500.000 abitanti; oppure sono disponibili come database dettagliati (per tutti i comuni, compresi quelli con meno di 2.000 abitanti). Sono diffusi poi i dati sui viaggi pendolari verso il luogo di lavoro e di studio, nonché sulle migrazioni residenziali tra comuni (*Mobility databases*). I risultati subcomunali del censimento della popolazione (*Sub-municipal databases*) sono basati su IRIS<sup>6</sup> (*statistical block groups*); per coprire l'intero territorio, sono presenti dati a livello comunale per i comuni non divisibili in IRIS. I database sono organizzati in temi: *Popolazione*, *Lavoro*, *Formazione*, *Coppie-Nuclei-Famiglie*, *Abitazioni*. Infine, sono predisposti dei *Detailed files* con lo scopo di fornire i dati del censimento della popolazione a utenti più esperti (enti pubblici, autorità locali, grandi aziende, società di consulenza, ricercatori, ecc.); si tratta di file di grandi dimensioni organizzati in temi *Popolazione (2 files)*, *Lavoro (1 file)*, *Mobilità e migrazioni (5 files)* *Abitazioni (1 file)* che, per poter essere letti o elaborati, necessitano di un software statistico.

<sup>6</sup> IRIS (aggregated units for statistical information) è l'unità territoriale di base per la diffusione di dati subcomunali. Tutti i comuni di almeno 10.000 abitanti e la maggior parte dei comuni tra 5.000 e 10.000 abitanti sono divisi in IRIS. Ci sono tre tipi di IRIS: residential IRIS (92%), business IRIS (5%), other IRIS (3%). La residential IRIS ha popolazione tra 1.800 e 5.000, viene identificata come unità omogenea per ambiente di vita; i confini dell'unità sono basati sulle principali linee di divisione del tessuto urbano (strade principali, ferrovie, corsi d'acqua, ecc.)

### 3.3. Elementi rilevanti emersi

Particolare interesse hanno generato le strategie e gli strumenti di diffusione adottati da alcuni tra gli stati esteri considerati, strategie da tenere in considerazione nella progettazione delle modalità di pubblicazione dei risultati censuari italiani. Si riportano gli elementi interessanti emersi dall'analisi dal punto di vista degli strumenti utilizzati per la presentazione dei risultati.

#### 3.3.1. Rappresentazione dei dati

Spagna (*Advanced query systems: Map viewer*) e Germania (*Results cartographically and visually*) hanno puntato sulla rappresentazione cartografica dei dati: la Germania tramite la visualizzazione di mappe suddivise per tema, la Spagna tramite visualizzazione di mappe tematiche semplici, indicatori per sezioni con cartografia digitalizzata, dati e indicatori per griglie da 1Km<sup>2</sup>. Inghilterra e Galles (*interactive content*) hanno utilizzato mappe manipolabili per evidenziare le aree di interesse, anche per il confronto con i risultati del censimento precedente.

#### 3.3.2. Utilizzo di nuove tecnologie e nuovi strumenti di diffusione

Per la rappresentazione dei dati, Inghilterra e Galles hanno utilizzato nuove tecnologie e nuovi strumenti di diffusione, come la *data visualisation* e le *analyses*; si tratta di cinque principali approcci per la produzione e la presentazione dei dati del censimento (*Stories*: progettate specificamente per la pubblicazione web. *Summaries*: articoli concisi che illustrano solo i principali risultati di un'analisi. *Infographics*: rappresentazioni grafiche dei principali risultati dell'analisi con alcune informazioni contestuali o annotazioni, progettate per l'utilizzo in siti web esterni. *Video podcasts*: comprende brevi animazioni PowerPoint -meno di cinque minuti- con audio per spiegare i principali risultati dell'analisi. *Interactive content*: include mappe, grafici e immagini che l'utente può manipolare per evidenziare le aree di interesse).

#### 3.3.3. Diffusione differenziata per ampiezza demografica dei comuni

La Spagna ha prodotto tavole con classificazioni più o meno dettagliate a seconda dell'ampiezza demografica dei comuni; alcune tavole sono state diffuse solo per comuni con una certa ampiezza; infine, nelle query personalizzate, in base all'ampiezza dei comuni, c'è la possibilità di incrociare più o meno variabili. Per la Germania, le query personalizzate di dati sulle famiglie sono consentite solo per comuni con più di 10.000 abitanti. In Francia, alcuni dati sono diffusi solo per comuni con più di 2.000 abitanti; sono prodotte tavole più dettagliate per i comuni più grandi.

#### 3.3.4. Analisi personalizzate (e Statistical disclosure control)

In Austria il *datawarehouse* permette l'incrocio di variabili e classificazioni per soddisfare le esigenze degli utenti; in Germania è stata predisposta una sezione dedicata alla creazione di analisi personalizzate (*Dynamic and individual results*); in Spagna è stato messo in piedi un sistema che permette query personalizzate (*Advanced query systems: Customised table query*). In Germania, le query sono differenziate in base alle unità statistiche selezionate; c'è la possibilità di scegliere tra valori assoluti o percentuali; infine, per le variabili sulle famiglie, i dati sono disponibili solo per comuni con più di 10.000 abitanti. In Spagna, c'è la possibilità di creare query differenziate a seconda del numero di unità del campione di famiglie coinvolto; le distribuzioni marginali risultanti dalle query possono differire da quelle delle tavole predefinite, i valori sono arrotondati a 5 unità (e multipli), i valori più piccoli di 5 sono contrassegnati da un asterisco, e, infine, i risultati delle query possono essere rappresentati in forma tabellare o cartografica.

#### 3.3.5. Trasparenza e attenzione rivolta agli utenti

La rassegna dei siti dei paesi analizzati ha rilevato la presenza di prospetti con la tempistica della diffusione, della data di pubblicazione dei dati, della versione del rilascio e dell'eventuale data di



ricaricamento. I risultati sono corredati di metadati e materiali di supporto per guidare gli utenti e aiutarli nella corretta interpretazione del dato, anche rispetto ai censimenti precedenti. In diversi paesi è, inoltre, presente un elenco delle informazioni diffuse nei censimenti precedenti che non sono state più pubblicate perché non rilevate più al censimento o non incluse nei registri/archivi utilizzati per la produzione delle statistiche censuarie. Sono, infine, pubblicate note per spiegare differenze tra i numeri del censimento e i numeri derivanti da archivi o da altre indagini campionarie. In Austria, ogni database ha la data di pubblicazione e di modifica (data dell'ultimo aggiornamento del cubo). Nell'introduzione di ogni db sono riportate le avvertenze e le definizioni relative agli oggetti contenuti. Inoltre, ogni classificazione e modalità che necessita di spiegazioni ha una (i) accanto che riporta la descrizione: i metadati sono sempre direttamente associati ai dati e sono facilmente reperibili. Sul sito è riportato l'elenco di variabili che non possono essere più prodotte perché non presenti negli archivi. In Spagna il *datawarehouse* è corredato da alcuni alert iniziali su come creare le query, alcune avvertenze su come leggere i dati provenienti dalla rilevazione campionaria e come interpretare i dati nelle celle con piccoli numeri. È ben visibile nel sito, nella sezione dedicata ai censimenti, il link al glossario, oltre alle comunicazioni sulle differenze rispetto ai dati diffusi al censimento precedente, sulle differenze tra i dati censuari e i dati da Registro Comunale e alle note per facilitare la corretta interpretazione dei dati. In Germania, è presente un link a "Important information" (Correzioni e modifiche nel database censimento, Numero di abitanti, Errore standard Relativo, Ottimizzazione del metodo di calcolo dei dati di base del censimento 2011 per gli aggiornamenti di popolazione intercensuaria). Nella sezione mappe, ampio spazio è dedicato ai metadati (descrizione di ciascun indicatore rappresentato in mappa) oltre ad alcune informazioni di corredo (*Introduction on the 2011 Census, Legal bases, Method, Optimisation of the method of calculating the basic data from the 2011 Census for intercensal population updates, Symbols*). Risultano essere ben evidenti la pagina dedicata al glossario e le note in cui si spiegano le differenze tra i numeri prodotti al censimento e i numeri derivanti da archivi o da altre indagini campionarie (con descrizione delle caratteristiche metodologiche e definitorie dell'indagine/archivio e del censimento). Per Inghilterra e Galles, le informazioni sul contenuto del piano di diffusione e sulla tempistica dei rilasci sono riportate in un prospetto. Sono presenti, inoltre, note tecniche con le differenze rispetto al censimento precedente. È disponibile materiale di riferimento insieme a ogni rilascio (*description, definition and context around each topic, comparability with the 2001 Census outputs, updates to the user guide, and further sources of information where applicable*). Ogni tavola diffusa riporta la data di pubblicazione (*Release date*) e la versione del rilascio (*Release*), oltre all'eventuale data di ricaricamento (*Revised*). Nel caso di aggiornamento delle informazioni, gli utenti sono stati informati tramite *Census News Alerts*, ovvero con notifiche sul sito web e aggiornamenti in una pagina dedicata. *Online* è presente anche il volume *Glossary* contenente tutte le definizioni utilizzate nella diffusione.

#### **4. Le novità a livello europeo: il programma per il 2021**

Una Task Force europea composta da 18 Stati Membri dell'UE, tra cui l'Italia, e 3 paesi candidati ha lavorato per la preparazione della nuova legislazione. La raccolta di dati del censimento UE del 2021 si basa sul Regolamento Quadro (n. 763/2008) che ha rappresentato la base giuridica per la tornata censuaria del 2011. L'attuale regolamento quadro sui censimenti, pertanto, rimane in vigore ma sono stati sviluppati nuovi regolamenti di attuazione per definire in dettaglio il censimento del 2021, comprese le definizioni delle variabili, le disaggregazioni, il programma degli incroci, i metadati e i quality report. Sebbene la legislazione per il 2021 fornisca le definizioni e le classificazioni che si applicano alle variabili, l'elenco delle variabili e il relativo dettaglio territoriale cui devono essere prodotte non può essere modificato. Il Regolamento n. 543/2017 riguarda le classificazioni e le specifiche tecniche delle variabili da applicare per la rilevazione del 2021 (European Commission, 2017a). Sono state introdotte numerose modifiche rispetto al 2011, allo scopo di semplificare, laddove possibile, definizioni e classificazioni, di rimuovere modalità opzionali. Inoltre le definizioni e le classificazioni sono state aggiornate per riflettere le nuove esigenze di dati, i cambiamenti nella società e per tenere conto delle nuove tecniche di rilevazione statistica. Il Regolamento n. 712/2017

riguarda l'anno di riferimento, il piano degli incroci e i metadati testuali per la rilevazione censuaria UE del 2021 (European Commission, 2017b). Il programma degli ipercubi per il 2021 conserva la ricchezza del censimento come fonte di dati e copre i temi necessari a soddisfare le esigenze prioritarie, riducendo al minimo il carico per gli Istituti nazionali di statistica. Il lavoro di definizione delle nuove tavole di diffusione per il 2021 si è basato su una serie di principi chiave, tra cui la semplificazione dei dati (derivanti dall'esperienza e dai risultati del 2011) con una riduzione del numero, della dimensione e della complessità delle tavole, la rimozione degli incroci o classificazioni poco utilizzati nel 2011 e un focus su variabili considerate prioritarie per gli utenti. In particolare, i dati prioritari sono quelli relativi alle variabili legate alle migrazioni (sulla migrazione e su altre questioni relative alla mobilità, compresi i legami tra mobilità, istruzione e occupazione). Mentre il programma degli incroci del 2011 includeva 60 hypercubes e 175 distribuzioni marginali principali, il programma per il 2021 si basa su 119 set di dati più piccoli che ciascuno Stato Membro deve rendere disponibile a Eurostat entro 27 mesi dalla fine dell'anno di riferimento della rilevazione (31 marzo 2024). I 119 incroci proposti per il censimento del 2021 sono più piccoli delle distribuzioni marginali principali del 2011, ma sono molto più piccoli rispetto agli ipercubi completi del 2011. Infatti, i 119 incroci possono essere raggruppati in circa 40 tavole più grandi (hypercubes) nell'ottica di semplificare la fornitura di dati attraverso il Censu Hub, che sarà usato anche per la diffusione dei dati europei del Censimento Europeo 2021.

I dati del censimento UE del 2021 possono essere suddivisi in due aree: i dati "principali", la cui diffusione è stata descritta sopra, e un insieme aggiuntivo di dati georeferenziati per griglia standard da 1Km<sup>2</sup>. Questo è un nuovo sviluppo per il censimento del 2021, distinto dal programma principale per la tornata censuaria prevista nell'UE. La raccolta di questi dati non è basata sul Regolamento n. 763/2008, ma utilizzerà una "Temporary Statistical Action" basata sul Regolamento sulle Statistiche Europee (European Commission, 2018). I dati dettagliati geograficamente del censimento sono essenziali per pianificare, sviluppare e valutare le politiche regionali e di coesione. La raccolta di dati sulla popolazione a livello di griglia da 1Km<sup>2</sup> è un'area in cui vi è una significativa domanda da parte degli utenti e che sta subendo un rapido sviluppo in molti Istituti nazionali di statistica. Il valore potenziale di unire i dati geo-spaziali con le statistiche ufficiali è quello di fornire migliori informazioni sociali e ambientali. La rilevazione censuaria del 2021 sarà un'importante opportunità per l'integrazione di dati statistici e geo-spaziali. I dati georeferenziati saranno prodotti in un'unica tavola bidimensionale; 13 numeri<sup>7</sup> saranno raccolti per ciascuna cella della griglia, che rappresenta il numero di persone in ciascuna delle categorie. Si prevede che anche questi dati saranno diffusi tramite Censu Hub.

## 5. La nuova strategia italiana: il censimento permanente

Il nuovo censimento della popolazione e delle abitazioni in Italia non sarà più decennale ma permanente. L'impianto si basa sul Sistema Integrato del Censimento permanente della popolazione e delle Indagini Sociali (SICIS), un complesso processo statistico che integra le informazioni derivanti dai registri e dalle indagini sociali. L'impianto prevede anche un *Master Sample*, basato su un campione in due fasi, composto da un insieme di indagini bilanciate e coordinate tra loro. Il nuovo disegno ha gli obiettivi di dare supporto ai registri statistici che l'Istat sta mettendo in piedi, aumentare la quantità di informazioni statistiche prodotte e migliorare il livello di copertura e qualità. Ai fini del censimento della popolazione e delle abitazioni, al momento, sono in costruzione tre registri statistici: il Registro della popolazione, contenente le variabili censuarie sesso, età, luogo di nascita, stato civile e cittadinanza; il Registro del lavoro, con informazioni sugli occupati: posizione nella professione, settore di attività economica, impiego a tempo pieno o parziale, contratti a tempo determinato e il Registro degli luoghi, che contiene i dati relativi a edifici e abitazioni. La prima fase di Master Sample si basa su due diverse componenti campionarie: la componente A (Areale) e la

---

<sup>7</sup> Popolazione totale, Sesso (Maschi, Femmine), Età (Minori di 15 anni, 15-64 anni, Con più di 65 anni), Occupati, Luogo di nascita (nel paese, in un altro paese UE, fuori dall'UE), Dimora abituale un anno prima (invariata, nel paese, fuori dal paese).

componente L (da Lista). La componente A - basata su un campione areale di sezioni di censimento e/o di indirizzi - è progettata per stimare i tassi di sotto-copertura e di sovra-copertura del registro di popolazione a livello nazionale e locale, per diversi profili di sottopopolazioni individuati sulla base di sesso, classi di età e cittadinanza. La componente L - basata su un campione di famiglie estratte da una lista - permetterà di integrare, da un punto di vista tematico, le informazioni che non sono desumibili dai registri. Le variabili di censimento sono state classificate in totalmente sostituibili, parzialmente sostituibili o non sostituibili. Il primo gruppo include variabili per le quali le fonti amministrative forniscono le informazioni proxy corrispondenti; sono considerate complete, poiché sono disponibili per tutte le unità nei registri tematici e sono precise, con un buon livello di copertura e qualità. Le variabili parzialmente sostituibili sono considerate complete e accurate solo per una parte della popolazione target; per il resto della popolazione, le informazioni sono mancanti o non possono essere considerate accurate (ad esempio, per il grado di istruzione il Ministero dell'Istruzione italiano non fornisce informazioni sulle lauree conseguite all'estero). Nell'ultimo gruppo (variabili non sostituibili), ci sono le variabili non ancora disponibili da fonti amministrative. Pertanto, il censimento permanente italiano è un censimento combinato che integrerà le informazioni contenute nei registri statistici, per le variabili sostituibili, con le informazioni ottenute tramite il SICIS per le altre variabili (Falorsi, 2016).

A sette anni dall'ultimo censimento tradizionale italiano, la prima indagine del nuovo impianto di censimento è prevista per ottobre 2018. Le componenti A e L coinvolgono circa 1 milione 400 mila famiglie e 2.800 comuni con diverse dimensioni della popolazione. Per il campione Areale si tratta di una rilevazione porta a porta totalmente *paperless* con la tecnica CAPI (intervista assistita da computer). Il questionario sarà compilato da rilevatori con uno specifico software su tablet. Anche per il campione da Lista l'indagine è totalmente priva di questionari cartacei; il modello di rilevazione sarà auto-compilato dalle famiglie tramite web (CAWI). Il primo passo è l'invio di lettere di invito a partecipare alla rilevazione con credenziali per accedere ai questionari web; in seguito, le famiglie che non hanno compilato il questionario entro una data stabilita saranno intervistate dai rilevatori che andranno sul campo per il recupero delle mancate risposte (CAPI). Il questionario è lo stesso per entrambe le componenti e include non solo variabili parzialmente o non sostituibili, ma tutte le variabili necessarie per la diffusione censuaria. L'obiettivo è di utilizzare le informazioni raccolte anche per testare la qualità e la copertura dei dati già disponibili nei registri.

## **6. Il piano di diffusione del censimento permanente**

Il piano di diffusione italiano è progettato per garantire le serie storiche, per soddisfare le esigenze degli utenti, per fornire dati su fenomeni di interesse nazionale. Uno dei principali scopi della nuova strategia di censimento italiano è la tempestività delle informazioni; alla tradizionale diffusione dei dati ogni 10 anni saranno affiancati piani degli incroci da pubblicare con una maggiore frequenza, per far fronte alle esigenze degli stakeholders e dei policy makers. Pertanto, l'Istat sta lavorando alla preparazione di diversi programmi di diffusione. Il Piano di diffusione esteso, in linea con l'UE, avrà come anno di riferimento il 2021. Anche in Italia è stata effettuata una lettura critica del piano di diffusione del 2011, attraverso l'analisi dell'uso dei dati pubblicati su I.Stat (accessi e scarico delle tavole) e delle richieste di elaborazioni personalizzate. L'intento era quello di semplificare cubi particolarmente complessi, prediligere tavole a dettaglio comunale o provinciale, riducendo il livello territoriale di alcuni incroci in origine previsti per regione e di introdurre nuovi incroci per descrivere fenomeni sociali emergenti. La bozza di piano decennale prevede al momento circa 250 tavole, che coprono varie aree tematiche e che garantiscono, a un fine dettaglio territoriale e un elevato dettaglio classificatorio, sia l'aggiornamento delle serie storiche che l'ampliamento dell'offerta informativa (nuovi approfondimenti su anziani, figli, tipologie familiari emergenti, NEET). Considerata l'offerta informativa disponibile grazie ai Registri e alla rilevazione campionaria annuale (MS) su oltre un milione di famiglie, è stata prevista anche la diffusione, con periodicità annuale, a partire dal 2019/2020, di un insieme di tavole a livello comunale (circa 30, rimodulabili nel tempo). Gli incroci comunali al momento predisposti riguardano la popolazione, le famiglie e i nuclei familiari (struttura della popolazione per sesso,

età e cittadinanza; migrazioni interne; grado di istruzione; condizione professionale o non professionale; spostamenti per motivi di studio o lavoro; tipologia di famiglie e di nuclei)<sup>8</sup>. E ancora, è stato individuato un ulteriore insieme di tavole che si può definire intermedio tra la proposta decennale e quella annuale. Si tratta infatti di un sottogruppo di tavole da diffondere per livello territoriale provinciale/area metropolitana (e grandi comuni<sup>9</sup>), ovvero di incroci di variabili non incluse nella proposta annuale o di variabili con classificazioni più dettagliate di quelle previste a livello comunale. Il set di tavole dovrebbe essere pubblicato con una periodicità intermedia, al momento ancora da definire (2, 3 o 5 anni). Il contenuto della rilevazione annuale può evolvere a mano a mano che cambiano le esigenze degli utenti e aumenta la disponibilità dei dati da fonte amministrativa. Nel tempo, è presumibile che la distinzione tra la diffusione censuaria annuale e decennale si riduca. Alcuni dei dati che ora sono considerati solo per la pubblicazione decennale potrebbero, in futuro, essere diffusi su base annuale o con una frequenza intermedia.

## **7. La diffusione dei dati censuari tra tradizione e innovazione**

Alla ristrutturazione dell'impianto censuario corrisponderà una nuova strategia di diffusione dei risultati, volta ad ampliare la platea degli utilizzatori degli output censuari, a caratterizzare la divulgazione a seconda dei soggetti che adoperano i dati, diversificando l'offerta informativa sul piano delle modalità di pubblicazione dei risultati e a soddisfare le esigenze degli utilizzatori (i comuni in via primaria) dei risultati del Censimento, necessarie per pianificare, sviluppare e avviare programmi e politiche e valutarne i risultati. La disamina della produzione censuaria di altri paesi europei ha permesso di esplorare strategie di diffusione, diversificate e alternative a quelle tradizionalmente utilizzate in Italia, alcune delle quali potrebbero essere adottate per il censimento futuro. L'Istat sta lavorando alla presentazione dei risultati del Censimento Permanente prevedendo la produzione di diversi output e l'utilizzo di diversi canali di diffusione.

Il punto di partenza è la preparazione di tavole predefinite, si sta predisponendo un piano di diffusione annuale, pluriennale, decennale (2021) con incroci definiti a priori. È in fase di studio la possibilità di una diffusione annuale comunale "minima" uniforme per tutti i comuni (indipendentemente dall'ampiezza demografica) e di un'ulteriore diffusione a seconda dell'ampiezza demografica dei comuni. È in fase di valutazione l'ipotesi di una diffusione differenziata per ampiezza demografica dei comuni, cioè di pubblicare incroci solo per i comuni di una certa dimensione, e/o tavole con classificazioni più dettagliate per i comuni sopra una soglia di popolazione e più aggregate per quelli sotto la soglia. È in fase di studio anche una diffusione annuale subcomunale, laddove le numerosità lo permettano, e una diffusione più articolata per numero di variabili e/o per complessità delle classificazioni a livello provinciale.

La diffusione in forma tabellare andrà integrata con nuovi strumenti di presentazione dei risultati o con il potenziamento di strumenti già utilizzati per il censimento del 2011 o già utilizzati in Istituto ma per la prima volta da adottare per il Censimento. Sul fronte della rappresentazione dei dati, quindi, si sta considerando la possibilità di ricorrere a tecnologie avanzate per la rappresentazione dei dati in generale e dell'output censuario nello specifico e a nuovi strumenti di presentazione dei risultati del censimento. Gli strumenti di *data visualization* permettono di trasformare tavole di numeri (a volte poco comprensibili) in tavole visive che possono catturare l'attenzione e spiegare un fenomeno. Si tratta di potenziare la rappresentazione dei dati attraverso strumenti di cartografia, ovvero potenziare l'uso di mappe, a dettaglio territoriale comunale, subcomunale, laddove possibile, e per griglie da 1Km<sup>2</sup>. La caratteristica delle griglie è che sono stabili nel tempo rispetto ai confini amministrativi (motivo per cui anche a livello europeo stanno investendo molto), pertanto si garantisce la confrontabilità nel tempo e nello spazio, superando il concetto di dimensione demografica dei comuni. Si punterà poi sulle infografiche, ovvero rappresentazioni grafiche dei principali risultati dell'analisi con alcune informazioni di contesto, come strumento di sintesi per la presentazione dei

---

<sup>8</sup> È in fase di studio l'ipotesi di una diffusione uguale per tutti i comuni e di un'ulteriore diffusione a seconda dell'ampiezza demografica dei comuni.

<sup>9</sup> Tradizionalmente, alla diffusione per livello territoriale provinciale è associata la diffusione per i grandi comuni. Nel 2011 erano considerati grandi comuni quelli con più di 150.000 abitanti.

dati. L'Istat ha da qualche anno avviato la produzione di infografiche statiche e dinamiche (interattive, videografiche, video), differenziate in base al contenuto e al canale da utilizzare. Si sta inoltre studiando la possibilità di consentire analisi personalizzate, che potranno avvenire o attraverso un sistema di query personalizzate nel caso in cui le informazioni che si stanno cercando non si trovino nelle tavole predefinite, o attraverso un datawarehouse che permetta l'incrocio di variabili e classificazioni in base alle esigenze degli utilizzatori<sup>10</sup>. Infine, sarà opportuno implementare il sistema di monitoraggio dell'uso dei dati. Le informazioni derivanti dal monitoraggio potrebbero permettere, infatti, di delineare i profili degli utenti che accedono ai dati e, per ciascun profilo, individuare le tavole e i temi di maggior interesse. Questo permetterebbe anche di modificare in itinere la diffusione a seconda degli aspetti più visualizzati e meno visualizzati e in base agli utilizzatori. Individuare i profili dei diversi segmenti di utenza potrebbe consentire di offrire prodotti e servizi personalizzati e di adottare una strategia di diffusione multicanale.

## 8. Considerazioni di sintesi

La diffusione, fase conclusiva delle rilevazioni statistiche, ha un ruolo cruciale essendo il risultato finale dell'intero processo censuario, il prodotto reso disponibile ai decisori e agli utenti. Il nuovo censimento permanente, che si svolgerà ogni anno su un campione di comuni e famiglie, comporta l'opportunità di un rilascio più frequente dei dati, con informazioni derivanti da archivi amministrativi, e di una divulgazione tempestiva. La necessaria rimodellazione della diffusione sarà al passo con i tempi, nell'ottica di ampliare la platea e di soddisfare le esigenze degli utilizzatori. Il sistema di diffusione sarà articolato su una periodicità diversa in funzione degli obiettivi e evolverà sulla base della disponibilità di nuovi contenuti, derivanti da archivi o rilevati tramite indagini campionarie. Quindi anche l'output cambia, rafforzando il suo ruolo al servizio della conoscenza. Si pone la necessità di utilizzare forme più sintetiche ma anche di più facile comprensione del patrimonio informativo dell'Istituto, orientate a favorire l'accessibilità e la leggibilità della produzione statistica per sviluppare una conoscenza della realtà, sociale ed economica dell'Italia, ai diversi livelli territoriali. Anche gli strumenti con cui i dati saranno rilasciati saranno innovativi. Si combineranno, tra tradizione e innovazione, piani di diffusione in linea con il passato con nuove tecniche per la pubblicazione e con forme alternative di visualizzazione e di interrogazione dei dati. L'Istat si sta muovendo verso un sistema dinamico e più flessibile di interrogazione e analisi *ad hoc* (*query on the fly*). La nuova impostazione censuaria garantirà la massima disponibilità di dati a livello micro per analisi e modelli, ai fini della programmazione oltre a quelli classici della ricerca. La nuova offerta informativa permetterà analisi trasversali e longitudinali, grazie alla possibilità di seguire nel tempo gli individui e le storie di vita. Sarà possibile studiare fenomeni attraverso analisi territoriali e spaziali su aree anche diverse dalle tradizionali Unità Amministrative Territoriali (Sistemi locali del lavoro, distretti socio-sanitari, unità sanitarie locali, distretti scolastici, contesti urbani). Da un punto di vista statistico, l'indagine censuaria ha rappresentato, fino a oggi, l'unica rilevazione in grado di fornire un livello di dettaglio minimo per le analisi del territorio; ha permesso di produrre dati per le più piccole partizioni territoriali (sezioni di censimento), che non sono disponibili per le indagini campionarie. La sfida dell'Istat è quella di ottenere risultati censuari a livello

---

<sup>10</sup> Inghilterra e Galles, che nel 2011 avevano diffuso tavole semplici e facilmente scaricabili in formato xls, stanno riflettendo sull'approccio alla diffusione per il prossimo censimento, prendendo in esame due modalità: produrre un gran numero di piccoli insiemi di dati oppure produrre un piccolo numero di grandi quantità di dati (open data) e dare la possibilità agli utenti di creare tavole attraverso uno strumento di "table builder". Hanno pertanto avviato il "table builder project" (con particolare attenzione rivolta allo Statistical Disclosure Control). Il primo approccio è quello adottato nel 2011 e per censimenti precedenti, sono stati prodotti più di 600 diversi set di dati, ciascuno disponibile per una o più aree geografiche. Questo approccio si sposa più facilmente con una metodologia di controllo della diffusione che ha permesso la coerenza all'interno e tra le serie di dati. Ma ha avuto alcuni inconvenienti: tempi lunghi per la definizione, la produzione e il controllo della diffusione dei dati e con l'impiego di molte risorse. Il secondo approccio può fornire agli utenti più facile accesso alle informazioni e maggiore flessibilità per creare insiemi di dati su misura, che soddisfino le esigenze degli utenti. Questa strategia garantirebbe maggiore facilità di ri-uso i dati in un formato aperto e di combinazione con altri insiemi di dati open. Ma la protezione della riservatezza dei dati in questo approccio può comportare l'adozione di forme di perturbazione, complesse da implementare e che possono portare a incongruenze tra le query o alla perdita di alcuni dati in celle piccole.

territoriale micro attraverso l'utilizzo di informazioni contenute nelle fonti amministrative, integrate con le informazioni raccolte in indagini campionarie. Ci vorrà tempo per avviare l'intero sistema ma, una volta messo a regime, sarà possibile avere dati annuali dettagliati (in termini di contenuto, classificazioni e territorio) che aiuteranno a conoscere, investigare, monitorare il territorio e le politiche che si attuano. Uno degli obiettivi principali della nuova diffusione è quello di soddisfare le esigenze degli utilizzatori dei dati, in via primaria i comuni, che adoperano i dati censuari (a livello territoriale minimo) per pianificare, programmare, avviare politiche sociali e territoriali, offrendo dati e strumenti di analisi e valutazione. Lo scopo è di favorire i processi decisionali di tutti i soggetti della società (cittadini, amministratori, ecc.) e non solo di policy maker e di specialisti dei dati (ricercatori, analisti, giornalisti).

## 9. Bibliografia

- European Parliament and Council (2008), Regulation (EC) No 763/2008 of the European Parliament and of the Council of 9 July 2008 on population and housing censuses, *Official Journal L 218*, 13/08/2008, P. 0014 - 0020.
- European Commission (2009), Commission Regulation (EC) No 1201/2009 of 30 November 2009 regards the technical specifications of the topics and of their breakdowns.
- European Commission (2010), Commission Regulation (EU) No 519/2010 of 16 June 2010 adopting the programme of the statistical data and of the metadata for population and housing censuses.
- European Commission (2017a), Commission Implementing Regulation (EU) 2017/543 of 22 March 2017 regards the technical specifications of the topics and of their breakdowns.
- European Commission (2017b), Commission Regulation (EU) 2017/712 of 20 April 2017 establishing the reference year and the programme of the statistical data and metadata for population and housing censuses.
- European Commission (2018), Commission implementing regulation deciding on a temporary direct statistical action in accordance with Regulation (EC) No 223/2009 to disseminate selected topics of the 2021 population and housing census geocoded to a 1Km<sup>2</sup> grid (in fase di pubblicazione).
- Dardanelli S., S. Mastroluca, A. Sasso e M. Verrascina (2008), *La progettazione dei censimenti generali 2010–2011. Novità di regolamentazione internazionale per il 15° Censimento generale della popolazione e delle abitazioni*. Roma: Istat. Dati e prodotti, Pubblicazioni scientifiche, Documenti n. 1/2009.
- Dardanelli S., Sasso A., Verrascina M. (2010), La diffusione dei dati del censimento della popolazione e delle abitazioni del 2011 alla luce di alcune novità introdotte a livello nazionale e internazionale, XXXI Conferenza Italiana di Scienze Regionali, Aosta, 20-22 Settembre.
- Falorsi S. (2016), "Il Sistema integrato del Censimento permanente della popolazione e delle indagini sociali (SICIS)", Istat (documento interno).
- Istat (2001). *14° Censimento della popolazione: dati definitivi*. <http://dawinci.istat.it/MD/>
- Istat (2005a). *Popolazione residente e abitazioni nelle province italiane*. Roma: Istat,
- Istat (2005b). *Struttura demografica e familiare della popolazione residente, Fascicolo nazionale I, Italia*. Roma: Istat.
- Istat (2006a). *Struttura socioeconomica e condizione abitativa della popolazione residente, Fascicolo nazionale II, Italia*. Roma: Istat.
- Istat (2006b). *Popolazione residente e abitazioni nelle regioni italiane*. Roma: Istat.
- Istat (2006c). *Popolazione residente e abitazioni nei grandi comuni italiani*. Roma: Istat.
- Istat (2006d). *Gli stranieri in Italia: analisi dei dati censuari*. Roma: Istat.

- Istat (2006e). *I Documenti*. Roma: Istat.
- Istat (2006f). *Il Piano di rilevazione e il Sistema di produzione*. Roma: Istat.
- Istat (2009). *La qualità dei dati*. Roma: Istat.
- Istat (2011). *15° Censimento della popolazione: i risultati*. <http://dati-censimentopopolazione.istat.it/Index.aspx>
- Istat (2016). *I sistemi informatici a supporto delle operazioni censuarie*. Roma: Istat.
- Istat (2017). *I contenuti informativi della rilevazione, la validazione e diffusione dei dati*. Roma: Istat.
- Verrascina M. (2017), *La diffusione dei dati censuari della popolazione in alcuni stati esteri: contenuti e strategie*, Istat Working Paper (in fase di pubblicazione).

## **ABSTRACT**

The population and housing Census will be no longer a traditional census, decennial and exhaustive, but continuous, annual and by sample. The present work illustrates the new strategy for the permanent census, which involves the use of data from administrative sources and sample surveys, in order to integrate information and produce census statistics, required at European level and of national interest, with greater frequency and timeliness than in the past. Census dissemination has undergone developments from 2001 to 2011 and further changes are planned for the future. The new approach, in fact, involves a reshaping of the dissemination, with a release of data with greater incidence and innovation in the tools for a timely publication, moving with the time and meeting the end users' needs. The work presents the proposal, at the design stage, for the permanent census dissemination on the informative content and on the publication tools. Among the aims of the production of census statistics is to broaden the informative provision and to facilitate access to data with a minimum territorial level to plan, adopt and evaluate social and territorial policies.