

L'uso dei dati di telefonia mobile per migliorare le stime dei flussi di popolazione intercomunali

L'Istat avvia una sperimentazione nell'ambito del progetto "Persons and places – popolazione insistente e city users"

25.11.2015



Nel contesto degli studi dedicati all'utilizzo dei Big Data per le finalità della statistica ufficiale verrà sviluppato un progetto teso a sperimentare l'uso integrato di dati anonimizzati di telefonia mobile e di dati amministrativi per migliorare la qualità delle stime dei flussi di popolazione intercomunale. Il test sarà condotto dall'[Istat](#) nell'ambito del progetto *Persons and Places - Popolazione insistente e city users*, utilizzando congiuntamente dati da sistema informativo statistico e dati che saranno forniti dalla compagnia telefonica Wind.

L'ipotesi allo studio è che l'integrazione informativa realizzata considerando le due fonti - dati da sistema informativo statistico e dati da traffico telefonico – possa migliorare la qualità delle stime dei flussi intercomunali di popolazione e, in particolare, contribuire alla stima di altri dati, tra cui l'incidenza del pendolarismo giornaliero sul totale dei flussi intercomunali.

I dati di telefonia mobile, da acquisire presso Wind, sono i Call Detail Records (CDR), relativi alle chiamate effettuate dagli utenti in un territorio predefinito e in un arco di tempo stabilito. I dati saranno relativi al seguente insieme di variabili: progressivo (anonimizzato) corrispondente all'utente che effettua le chiamate; codice del Comune all'interno del quale si trova la cella da cui proviene la chiamata; data e ora della chiamata.

Sulla base delle definizioni di dato personale (art. 4, comma 1, lett. b del d.lgs. n. 196/2003) e di dato anonimo (art. 4, comma 1, lett. n del d.lgs. n. 196/2003), i dati richiesti a Wind sono anonimi e non personali. Infatti, essi non permettono di individuare il singolo utente di telefonia mobile né direttamente (il progressivo è totalmente indipendente dall'identificativo dell'utente) né indirettamente (non sono disponibili informazioni che possano consentire un'identificazione indiretta del singolo utente).

In particolare, viene esplicitamente richiesto al gestore Wind di assegnare un codice ad ogni CDR e successivamente provvedere a cancellare la tabella di corrispondenza tra codici assegnati e identificativi dei CDR, prima di trasmetterli all'Istat.

Per meglio garantire l'impossibilità di identificazione indiretta, inoltre, verrà chiesto a Wind di fornire, anziché le coordinate della cella di chiamata, il solo codice del comune cui la cella appartiene.

Per la realizzazione del progetto saranno utilizzati anche i microdati provenienti dal Sistema informativo statistico dell'Istat denominato *Persons and Places*, relativo alla *Popolazione insistente e city users*. Ovviamente, il confronto sarà possibile solo a livello aggregato, non essendo disponibile per i dati di telefonia mobile alcuna informazione che permetta di stabilire corrispondenze a

livello micro. In particolare, i dati relativi ai luoghi di residenza, studio e lavoro permetteranno di calcolare gli aggregati utili per valutare la qualità di quelli corrispondenti, ottenuti dai dati di telefonia mobile.

I dati forniti da Wind verranno elaborati dall'Istat con una metodologia già sperimentata dal [Cnr-Isti](#) e dall'[Università di Pisa](#) per la stima dei flussi di mobilità sul territorio, così come descritta nel contributo [Use of mobile phone data to estimate mobility flows](#), pubblicato negli atti della 47ª Riunione scientifica della Società italiana di statistica ([Sis](#)).

Tale metodologia si basa sulla caratterizzazione dei singoli utenti di telefonia mobile mediante le indicazioni spazio-temporali (data, ora e Comune) associate alle chiamate effettuate, in modo da stimare le componenti di interesse della popolazione "insistente" su un dato Comune. Poiché nella sperimentazione effettuata da Cnr-Isti e Università di Pisa il territorio considerato è quello del Comune di Pisa, in una prima fase verranno utilizzati i dati relativi a quattro settimane di traffico telefonico in provincia di Pisa, durante il mese di ottobre del 2011. Dopo questa sperimentazione, con dati già in possesso di Cnr-Isti e Università di Pisa, sarà presa in considerazione la possibilità di elaborazioni estese a territori più ampi, a livello regionale e nazionale.

Per approfondire gli aspetti tecnici e metodologici, è possibile consultare il [documento di progetto](#).

Roberta Vivio