

Big Data

La sessione prevede un intervento generale, sullo stato attuale e sulle prospettive di utilizzo dei Big Data in Istat, e due interventi specifici relativi a due tipologie di fonti Big, Internet as a Data Source e Social media.

Il primo intervento descrive la roadmap per l'utilizzo dei Big Data in Istat, con le azioni che l'Istat intraprenderà sul breve (dicembre 2016) e medio termine (dicembre 2020).

Il secondo intervento si focalizza sulle sfide metodologiche e tecnologiche relative all'uso di Internet as a Data Source. In particolare: uso di Web scrapers per l'accesso ai dati Web, memorizzazione e accesso a dati poco strutturati, tecniche di text processing e tecniche di classificazione supervisionata.

Infine, il terzo intervento illustra come le tradizionali statistiche mensili sul clima di fiducia dei consumatori, che datano dal 1985, potrebbero nei prossimi anni giovare delle tecniche di sentiment analysis applicate ai social media. Dall'Accademia dove sono nate, queste tecniche potrebbero presto entrare nella produzione corrente della Statistica ufficiale con il duplice fine di rendere più tempestivi gli indicatori prodotti e di raggiungere fasce di opinione pubblica sempre meno raggiungibili con le tradizionali tecniche d'indagine.