



STUDENTE FUTURO STUDENTE LAUREATO DOCENTE  
Mi@Lumsa A+ / A- Contatti English

## CORSI

Sei in: Home » Ricerca » Big Data » Prof. [Giorgio Alleva](#), Presidente [Istat](#) (Istituto Nazionale di Statistica)

Prof. [Giorgio Alleva](#), Presidente [Istat](#)  
(Istituto Nazionale di Statistica)

di Luigi Crimella



### I Big Data per potenziare la ricerca statistica e allargare la “democrazia digitale”

*Nella mole dei dati che l'Istat, Istituto nazionale di statistica, sforna ogni anno, ci sono rilevazioni che si basano anche sull'utilizzo di informazioni estratte dai Big Data. Nell'intervista che segue al Presidente prof. Giorgio Alleva emerge che anche per il futuro della statistica applicata al benessere sociale proprio i Big Data possono svolgere un ruolo importante. Grazie a iniziative innovative, quali il “Big Data Contest” promosso insieme a Google, l'Istat indica la strada della innovazione nella ricerca statistica in collaborazione con realtà di “data scientists”. Nel sito ufficiale dell'ISTAT cresce lo spazio dedicato agli “open data” rivolti sia a specialisti sia a un pubblico indifferenziato. Il prof. Alleva, ordinario di Statistica presso la facoltà di economia dell'Università Sapienza di Roma, in precedenza ha operato presso il Centro Studi Investimenti Sociali (Censis), ha insegnato statistica all'Università Luiss “Guido Carli” di Roma e all'Universidad del Salvador di Buenos Aires. È autore di numerose pubblicazioni su tematiche quali l'analisi dei dati territoriali e i metodi di statistica multivariata, la valutazione della redditività di investimenti pubblici e privati, la programmazione per obiettivi e la valutazione della performance, i modelli di domanda di trasporti.*

#### In questi ultimi anni sembra che i Big Data stiano assumendo una importanza superiore a quella delle analisi statistiche tradizionali. Come valuta questa tendenza? Possiede dei presupposti di valore scientifico?

**ALLEVA** – Precisiamo subito che i Big Data non andranno a sostituire le fonti tradizionali ma contribuiranno ad arricchire l'informazione statistica disponibile. I Big Data rappresentano sicuramente una sfida non facile per gli istituti di statistica perché sono un tipo d'informazione con la quale non siamo abituati a trattare, ma allo stesso tempo rappresentano un'opportunità, perché possono ampliare le capacità di analisi e la tempestività delle informazioni. Per loro stessa natura, i Big Data hanno caratteristiche molto differenti rispetto alle fonti tradizionali. Tant'è che è necessario adeguare i sistemi di raccolta e trattamento dei dati e costruire un framework per la valutazione della qualità di queste nuove fonti. Su questi aspetti la comunità statistica internazionale si sta già confrontando.

#### L'Istat ha mostrato una crescente attenzione all'adozione di metodiche legate all'IT, ai Big Data, all'utilizzo di programmi in grado di analizzare campionature sempre più complesse. Ci può illustrare sinteticamente queste novità?

**ALLEVA** – Alla stregua di altri istituti di statistica, i più innovativi, l'Istat sta portando avanti un programma d'integrazione di tutte le fonti disponibili (survey tradizionali, dati amministrativi, nuove fonti). Siamo convinti di ottenere in questo modo grandi guadagni in termini di accuratezza, coerenza e completezza delle informazioni statistiche prodotte e al contempo di ridurre il disturbo statistico che grava sui rispondenti delle nostre indagini. L'uso dei Big Data nella produzione delle statistiche ufficiali rientra proprio in questa strategia. Il loro utilizzo richiede però, da una parte, la risoluzione di complessi problemi di carattere metodologico, dall'altra, un forte investimento in strumenti IT in grado di trattare insieme di dati di dimensione incomparabilmente superiore a quella che caratterizza le tradizionali rilevazioni statistiche. Ad esempio, l'Istat ha contribuito allo sviluppo della piattaforma Sandbox, promossa dall'High Level Group of Modernization of Official Statistics dell'UNECE e basato sulla cooperazione con circa 20 organizzazioni, ottenendo importanti indicazioni su metodologie e strumenti utili all'elaborazione di grandi insiemi di dati. Inoltre abbiamo lanciato un nuovo laboratorio interno per la sperimentazione e la messa a punto di nuove tecnologie e sistemi, l'Istat Innovation Lab (uno spazio fisico per lanciare, promuovere e incentivare nuove idee) su cui puntiamo molto. Più in generale, vorrei segnalare che stiamo anche predisponendo l'apertura nei prossimi mesi di un'area di diffusione delle “Experimental Statistics”, che comprendono non solo indicatori ottenuti dall'uso dei big data, ma anche nuovi prodotti costruiti sull'integrazione delle fonti e analisi statistiche basate su modelli, a integrare l'informazione che proviene dalle statistiche tradizionali.

#### Ci può dire qualcosa in merito alle vostre ricerche sul rapporto tra cittadini, imprese e nuove tecnologie?

**ALLEVA** – Da quindici anni l'Istat documenta lo sviluppo della Società dell'informazione in Italia attraverso due indagini specifiche, sugli individui e sulle imprese, realizzate in maniera coordinata con gli Istituti di statistica degli altri paesi europei. I risultati sono utilizzati dalla Commissione europea per valutare il progresso degli Stati membri verso gli obiettivi dell'Agenda digitale dell'Unione europea. Nel nostro Paese scontiamo un ritardo storico nella diffusione dell'uso di Internet tra i cittadini e delle opportunità offerte dalla rete. Naturalmente le percentuali cambiano secondo le caratteristiche

demografiche e socio-economiche degli individui intervistati. In particolare, a fare ostacolo è l'età (ovviamente) ma, a parità d'età, conta l'istruzione. Per quanto riguarda le imprese, l'uso delle tecnologie ICT è, invece, almeno in termini d'accesso, in linea con la media europea. Permangono tuttavia opportunità di maggiore sfruttamento delle potenzialità offerte dalle nuove tecnologie, come ad esempio nel commercio elettronico. Le nostre analisi mostrano che parte del ritardo potrebbe riflettere vincoli organizzativi all'interno delle aziende. L'Italia è per esempio il Paese europeo con la percentuale più elevata d'impresе – quasi un terzo contro il 22% per l'insieme dell'Unione – che ritiene non valga la pena sostenere i costi di vendere online, perché non ripagati dai benefici. La debolezza dell'offerta e della domanda, che originano da un comune substrato culturale, tendono quindi a influenzarsi reciprocamente.

### **I sostenitori delle potenzialità di innovazione sociale dei Big Data fanno riferimento alle ricadute attese nei servizi pubblici, nell' "internet of things" e nelle smart cities. Avete effettuato qualche studio e ricerca specifica?**

**ALLEVA** – Da tempo abbiamo avviato studi sulla misurazione della smartness a livello territoriale, con obiettivi di definizione, valutazione e monitoraggio. In questo contesto i Big Data possono sicuramente svolgere un ruolo importante. Se ci riferiamo alle dimensioni della smartness, la mobilità e i trasporti pubblici ne rappresentano un elemento fondamentale. È estremamente utile conoscere i flussi casa/lavoro e casa/studio con elevato dettaglio territoriale; in questo campo, i dati di telefonia mobile e quelli derivanti dai GPS collocati su alcuni veicoli costituiscono una fonte ideale per finalità di sicurezza stradale, programmazione dei trasporti, persino per il coordinamento delle attività di supporto pubblico in situazioni di calamità naturali. Se consideriamo invece la dimensione "ambiente", un esempio è dato dal possibile utilizzo degli smart meters ovvero quei device elettronici che segnalano a intervalli ravvicinati i consumi di gas, acqua ed energia elettrica delle famiglie: il loro utilizzo può servire non solo a fini statistici (per la stima dei consumi energetici) ma anche per ridurre sprechi e favorire soluzioni alternative.

### **Recentemente avete promosso insieme a Google una iniziativa denominata "Big Data Contest". Cosa è emerso?**

**ALLEVA** – Il contest è stato vinto da università ed enti di ricerca (Università degli Studi di Milano, CNR di Pisa, Università di Torino), ma anche da soggetti singoli. Alcune idee interessanti ed innovative hanno avuto seguito in progetti che sono stati poi lanciati in [Istat](#), come ad esempio l'uso del social network Twitter per indagare in modo dinamico la percezione di benessere personale e relazionale all'interno della società. Questo contest ha rappresentato un esempio positivo di partnership pubblico-privato: Google è un grosso provider di Big Data ma [Istat](#) è il soggetto che fornisce statistiche "certificate". Credo che il tema delle partnership sia fondamentale per potenziare la linea di ricerca sui Big Data e per migliorare le nostre analisi grazie al valore generato dall'utilizzo e dall'integrazione di dati privati per uno scopo pubblico. A questo riguardo, vorrei qui ricordare anche la partnership che [Istat](#) ha sviluppato con Wind per i dati di telefonia mobile o con la Grande Distribuzione per l'uso degli scanner data.

### **"Big Data" non è sinonimo di "Open Data". Per questi ultimi, in vista di favorire quella che è chiamata "democrazia digitale", [Istat](#) sta facendo qualcosa?**

**ALLEVA** – Riguardo gli Open Data, [Istat](#) è uno degli attori principali, oltre che fornitore, sul panorama nazionale e internazionale. I dati pubblicati tramite il nostro data warehouse di diffusione [Istat](#) sono in formato open e corredati da un insieme d'informazioni esplicative (metadati) che ne facilitano l'utilizzo e la comprensione. Recentemente abbiamo anche pubblicato un portale dedicato ai Linked Open data, LOD in gergo tecnico, che sono una forma particolarmente avanzata di dati aperti. La pubblicazione di dati in formato LOD, oltre che aderire alle direttive per la valorizzazione del patrimonio informativo pubblico, pubblicate dall'Agenzia per l'Italia Digitale (AGID), risponde alle esigenze sempre più vaste espresse dalle comunità di utilizzatori di disporre di dati standardizzati e interoperabili su scala nazionale. I Linked Open Data, infatti, consentono di navigare dati in formato open sulla base di tecnologie e standard del web semantico. I LOD sono accessibili dagli utenti attraverso interfacce grafiche e possono essere direttamente interrogati da applicazioni esterne in modo indipendente dalle tecnologie adottate. [Istat](#) partecipa anche a diverse iniziative internazionali per promuovere la pubblicazione di dati nei formati open più avanzati. Ad esempio, ha guidato nel 2016, insieme all'Istituto di Statistica francese, un progetto della divisione economica europea delle Nazioni Unite (UNECE-United Nations Economic Commission for Europe) per facilitare l'integrazione dei dati pubblicati secondo modelli e formati open (progetto Implementing ModernStats Standards).

### **A chi vi rivolgete in particolare?**

**ALLEVA** – La piattaforma LOD dell'Istat è costituita da un ambiente di navigazione e interrogazione relativo sia alle basi territoriali sia alle variabili censuarie disaggregate a livello di sezioni di censimento, località abitate, aree di censimento e aree sub-comunali. Vengono rese disponibili informazioni sulla popolazione residente in Italia al 9 ottobre 2011 - classificata per sesso, età, stato civile, cittadinanza, grado di istruzione, condizione professionale e spostamenti pendolari per motivi di studio o lavoro - sulle famiglie, le abitazioni e gli edifici. Sono a disposizione anche file pre-confezionati in formato csv. Si tratta di informazioni di grande interesse per ragioni di lavoro e di studio. Il portale dei LOD sembra essere molto apprezzato: i numeri degli accessi sono, infatti, molto interessanti.

### **L'indice del Bes (Benessere equo e sostenibile) sta facendosi strada cercando di entrare in competizione con il Pil (Prodotto interno lordo). Ve ne siete già occupati e risulta misurabile con metodi statistici tradizionali o debbono entrare in gioco i Big Data?**

**ALLEVA** – Il BES è un progetto ormai consolidato, che si colloca nel filone d'iniziativa realizzate in molti Paesi europei (e non) di misurazione del benessere sociale e individuale e della sua sostenibilità. Il suo obiettivo non è però quello di individuare misure di sintesi che competano con il PIL, ma di offrire un quadro informativo multidimensionale, capace di dare un'adeguata rappresentazione a tutti gli aspetti che caratterizzano il benessere e la sua sostenibilità, come la salute, il lavoro, l'istruzione, l'ambiente, la qualità della vita individuale. Al fine di costruire questo quadro di informazioni [Istat](#) ha intrapreso un percorso che, attraverso il coinvolgimento della società civile e di esperti, ha portato alla definizione dell'attuale set di indicatori contenuti nel BES e alla loro successiva diffusione annuale in un rapporto dedicato. È importante sapere che l'insieme degli indicatori non sarà immutabile, così come non lo è nel tempo il concetto di benessere. Senza dubbio la disponibilità di nuove fonti potrà rappresentare un importante stimolo al miglioramento e alla selezione anche degli indicatori contenuti nel Bes. Come esempio voglio citare la possibilità di realizzare delle misure di sentiment/opinione attraverso l'applicazione di tecniche di data mining sui social network. Sarà interessante verificare e studiare se i giudizi dei cittadini rilevati attraverso le tradizionali survey saranno coerenti con quelle che emergeranno dall'analisi delle nuove fonti di dati.

### **A proposito dei Big Data avete fatto assunzioni mirate e come avete strutturato tale nuovo campo di ricerca all'interno dell'Istat?**

**ALLEVA** – La nostra strategia prevede due fasi principali. Quella iniziale è di sperimentazione e tipicamente è realizzata da una partnership con altri organismi competenti, come altri enti di ricerca e università. La fase successiva è invece di valutazione e consente di capire quali sperimentazioni siano

così promettenti da aggiungere alle statistiche ufficiali nuovi indicatori o da migliorare la capacità di quelli attuali. In questo passaggio, assume particolare rilievo la collaborazione con l'Autorità garante della privacy: l'introduzione nella produzione statistica corrente di nuove fonti richiede anche un aggiornamento del quadro legale. Per la fase di sperimentazione abbiamo bisogno di diverse professionalità. [L'Istat](#) è partner di corsi post-laurea e specializzazione nel campo della "data science" con alcune università, proprio per costruire e programmare insieme le competenze necessarie. Tipicamente si tratta di professionalità statistiche e tecnologiche, ma servono anche competenze di tipo umanistico per capire bene il significato nella lettura dei fenomeni. Abbiamo constatato come per la professione del "data scientist" la dinamica occupazionale sia molto positiva rispetto a quella generale: i giovani che scelgono queste specializzazioni trovano facilmente lavoro. [L'Istat](#) ha vincoli sul reclutamento, naturalmente, ma abbiamo visto che possiamo ottenere buoni risultati con programmi mirati di formazione. Del resto, quello dei Big Data è un campo di ricerca sul quale i nostri ricercatori e metodologi hanno da sempre mostrato un grande interesse. Ma certamente il tema della competenza è una delle sfide principali per promuovere l'uso dei Big Data negli istituti nazionali di statistica.

#### Documenti di riferimento

- > [Il valore dei dati nell'era dei Big Data](#)
- > [Data Innovation in Official Statistics: the Leading Role of Open Data](#)

CORSI DI LAUREA

MASTER E POST LAUREA



Trasparenza

Mappa del sito

Intranet

Valutazione didattica

Available on the

ANDROID APP ON

UNIVERSITÀ