



00

Basi statistiche e presentazioni generali

1789-1700

Ufficio federale di statistica Strategia «Data Innovation»

Scopo, obiettivi strategici e fasi di attuazione



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Dipartimento federale dell'interno DFI
Ufficio federale di statistica UST

Neuchâtel 2017

Editore:	Ufficio federale di statistica (UST)
Informazioni:	Bertrand Loison, UST, tel. 058 463 67 70, bertrand.loison@bfs.admin.ch
Contenuto:	Bertrand Loison, UST; Diego Kuonen, Statoo Consulting
Serie:	Statistica della Svizzera
Settore:	00 Basi statistiche e presentazioni generali
Testo originale:	inglese
Traduzione:	Servizi linguistici dell'UST
Layout:	sezione DIAM, Prepress/Print
Titolo:	UST; concezione: Netthoevel & Gaberthüel, Bienne; foto: © vinnstock – Fotolia.com
Copyright:	UST, Neuchâtel 2017 Riproduzione autorizzata, eccetto per scopi commerciali, con citazione della fonte.
Ordinazione di versioni cartacee:	Ufficio federale di statistica, CH-2010 Neuchâtel, tel. 058 463 60 60, fax 058 463 60 61, order@bfs.admin.ch
Prezzo:	gratuito
Download:	www.statistica.admin.ch (gratuito)
Numero UST:	1789-1700

Indice

Nota introduttiva	5
Introduzione	7
Il nostro punto di vista sulla «Data Innovation»	9
Base giuridica in vigore	9
Dati primari e secondari	10
Ciclo di ragionamento induttivo-deduttivo	11
Primo obiettivo strategico	11
Secondo obiettivo strategico	13
Considerazioni generali	13
Realizzazione degli obiettivi strategici	14

Nota introduttiva

L'oggetto principale del presente documento è incrementare e/o completare la produzione statistica ufficiale di base dell'Ufficio federale di statistica (UST) cui sia sensato applicare la «data innovation» come definita qui di seguito. Come esposto più avanti, lo scopo di questa strategia è limitato ad alcuni elementi dell'innovazione statistica. Pertanto non tutti gli aspetti innovativi saranno utilizzati integralmente all'UST, né si intende stabilire un ordine di priorità. Inoltre, questa strategia sarà oggetto di revisioni e aggiornamenti man mano che gli obiettivi saranno raggiunti attraverso le fasi descritte qui di seguito.

Introduzione

La rivoluzione digitale in corso si fonda sui dati: i termini come big data, cloud, internet degli oggetti, internet di ogni cosa, quarta rivoluzione industriale, città intelligenti ed economia basata sui dati non sono più solo parole sulla bocca di tutti, ma veri e propri concetti che stanno modificando le abitudini delle persone, dei consumatori e delle imprese. La trasformazione digitale è cominciata con la prima ondata di digitalizzazione, ovvero quella tecnica, che ha convertito contenuti e servizi analogici in formato digitale (senza modifiche sostanziali ai processi che li supportano) e ciò ha portato alla rivoluzione dei (big) dati. I big data rappresentano però un'infrastruttura di gestione dei dati (corredata da relativi hardware, software e architettura) e non dovrebbero essere utilizzati soltanto per fini museali. Già nel 1942 l'esperto di statistiche ufficiali, W. Edwards Deming, ha constatato che la finalità ultima della raccolta di dati è fornire una base o una raccomandazione d'azione. È necessaria quindi una seconda ondata di digitalizzazione per consentire di trarre informazioni da questi big data e generare un valore aggiunto dalla rivoluzione che hanno determinato e che ha interessato tutta la società: questo fenomeno è noto all'UST come **«data innovation»**. In questo senso, la maggiore sfida è rappresentata dalla veridicità del «pedigree dei dati», ovvero la **fiducia in questi dati** e nella loro affidabilità, funzionalità, validità e qualità, nonché nella loro trasparenza e processo di produzione. La **veridicità dei dati** è più importante che mai in un mondo politico in cui la verità è diventata irrilevante e le false notizie spopolano.

Le statistiche ufficiali hanno sempre utilizzato diverse fonti di dati. Per fornire servizi statistici affidabili e sicuri, l'UST ha costantemente esaminato, fin dalla fondazione nel 1860, la veridicità del pedigree dei propri dati. Tuttavia, l'utilizzo o il riutilizzo di fonti di dati esistenti e/o l'uso di nuove fonti di dati consentirebbe all'UST di innovare, incrementare e/o completare le statistiche pubbliche, rendendole più efficienti ed efficaci, nonché più attuali e complete.

È importante osservare che la rivoluzione dei dati non è rilevante soltanto nell'ambito delle statistiche ufficiali.

A livello internazionale, l'accesso e l'utilizzo di nuove fonti di dati rappresentano i temi di discussione delle conferenze regionali e mondiali. Gli obiettivi di sviluppo sostenibile dell'Agenda 2030 cambiano le regole del gioco: i dati e la statistica ufficiale acquistano molta più attenzione dal punto di vista politico. A causa di queste richieste, si deve discutere, esaminare e, se possibile, concedere l'accesso a nuove fonti di dati e il loro utilizzo. Il piano d'azione globale per lo sviluppo sostenibile, adottato a marzo 2017 durante la 48esima sessione della Commissione statistica Onu a Città del Capo, richiede

diverse azioni per quanto concerne le nuove fonti di dati. Ad esempio, è necessario identificare ed eliminare le barriere che impediscono l'utilizzo di nuove fonti di dati, compresi i dati dei registri, quelli amministrativi o altri dati derivanti da nuove fonti e bisogna coordinare le iniziative per integrarle nella produzione statistica ufficiale. Il lavoro svolto dall'UST contribuirà a questa azione.

Negli ultimi anni sono state definite a livello nazionale varie strategie che mirano a estendere l'accesso e l'utilizzo di nuove fonti di dati e di dati dei registri (p.es. «Digital Switzerland Strategy», «National Data Policy»). Tuttavia, queste strategie sono per lo più basate sulla tecnologia e non danno molta importanza all'utilizzo di questi dati per finalità statistiche. Poiché le statistiche ufficiali, e in particolare l'UST, sono sottoposti a una pressione crescente da parte del Parlamento per ridurre l'onere delle persone interrogate attraverso l'uso di dati tratti dai registri e di dati amministrativi, nonché di altre fonti di dati potenziali derivanti dai progressi della digitalizzazione, il lavoro dell'UST contribuirà sia a posizionare le statistiche ufficiali rispetto all'implementazione di queste iniziative strategiche, sia a elaborare soluzioni efficienti per rispondere alle pressioni politiche e alle relative aspettative.

Il nostro punto di vista sulla «Data Innovation»

Considerando «**l'analisi**» come la scienza che consente di apprendere attraverso i dati (oppure di trarre un significato da quest'ultimi), l'UST definisce

la «**data innovation**» come l'applicazione di *metodi di analisi complementari* (ad es. analisi predittiva ricorrendo agli approcci della statistica avanzata, scienza dei dati e/o apprendimento automatico) alle fonti di dati esistenti (o tradizionali) e/o nuove (o non tradizionali),

per sostenere il ruolo della statistica ufficiale nel processo democratico svizzero, garantendo che le informazioni fornite siano sempre accertate, trasparenti e degne di fiducia.

In quanto tale, la «data innovation» potrebbe offrire opportunità per interventi più economici e fornire nuove prospettive, che potrebbero non essere state considerate nel quadro di approcci e metodi analitici tradizionali. L'obiettivo della «data innovation» oltre ad aumentare la tempestività e il grado di dettaglio delle statistiche ufficiali, è quello di provare a migliorare la qualità, la portata e il rapporto costi-benefici dei prodotti statistici, nonché di ridurre l'onere per le economie domestiche e le imprese.

Base giuridica in vigore

Secondo quanto previsto dalla legge («Legge sulla statistica federale»), l'UST e gli altri servizi di statistica ufficiale si impegnano in modo professionale e indipendente a fornire risultati rappresentativi in merito alle condizioni e all'evoluzione di temi importanti così da potersi basare su dati concreti per prendere le decisioni e definire le politiche riguardanti popolazione, economia, società, formazione, ricerca e pianificazione del territorio e dell'ambiente in Svizzera. Per raggiungere questi obiettivi, l'UST utilizza già un ventaglio di fonti di dati. Ai sensi di legge, l'UST deve innanzitutto verificare se i dati amministrativi e dei registri della Confederazione, dei Cantoni, dei Comuni o di altri enti pubblici siano disponibili nella qualità necessaria. Soltanto se così non fosse e qualora, spetti all'UST, conformemente al programma pluriennale, fornire informazioni statistiche, sarebbe quest'ultimo a raccogliere i dati attraverso censimenti e rilevazioni, il cui numero e tipologia sarebbero limitati allo stretto necessario. I dati raccolti dall'UST hanno il vantaggio di esserlo stato per un chiaro scopo statistico e di presentare veridicità e qualità garantite dall'Ufficio federale di statistica.

Dati primari e secondari

Nel quadro della «data innovation» l'UST distingue le fonti dei dati in primarie e secondarie.

I cosiddetti **dati primari** («ottenuti o preparati») sono stati raccolti (e preparati) dall'UST a fini statistici per spiegare e verificare la validità di precise «idee» esistenti emerse attraverso l'applicazione di concetti teorici. L'apprendimento desunto dai dati primari è noto come analisi primaria (o «dall'alto», ovvero con carattere esplicativo e confermativo). Il modello di analisi corrispondente è il «ragionamento deduttivo», secondo cui «l'idea (la teoria) viene prima di tutto». I dati ricavati dalle fonti tradizionali, quali i censimenti e le rilevazioni che l'UST realizza a fini statistici, costituiscono un esempio di dati primari.

D'altro canto, i cosiddetti **dati secondari** (tratti dall'osservazione oppure «scoperti o organici») sono stati raccolti (e trattati) per altre ragioni (e spesso non sono sotto il controllo dell'UST) e potrebbero essere utilizzati per creare nuove idee («teorie»). L'apprendimento desunto dai dati secondari è noto come analisi secondaria (o «dal basso», ovvero con carattere sperimentale e predittivo). Il modello di analisi corrispondente è il «ragionamento induttivo», secondo cui «i dati vengono prima di tutto». I dati ricavati dalle fonti non tradizionali, come le informazioni tratte dai registri interni ed esterni, quelli amministrativi e altri dati digitali provenienti da dispositivi, macchine, sensori, satelliti, droni e social media, sono esempi di dati secondari. Poiché non sono concepite per essere direttamente utilizzate nel sistema di produzione statistico ufficiale, le fonti di dati secondarie devono essere «rese idonee a raggiungere l'obiettivo di inferenza statistica», ovvero bisogna giungere, attraverso un ragionamento deduttivo, a una conclusione conforme allo scopo delle statistiche ufficiali.

Per farlo, è possibile classificare i dati secondari in **identificabili e non identificabili**. I primi possono essere associati a una singola unità in un luogo e momento preciso, quali un individuo, un'istituzione, un prodotto o una localizzazione geografica (per es. dati amministrativi e dei registri, immagini satellitari, informazioni geospaziali e codici a barre dei prodotti). Per quanto concerne i dati non identificabili, non è possibile identificarli ad alcun livello (ad esempio i dati di Google Trends, i feed di Twitter e altre forme di social media). Contrariamente ai dati secondari non identificabili dei quali non è possibile verificare la veridicità, i dati secondari identificabili possono essere approntati a fini di inferenza statistica purché sia stato possibile appurarne la veridicità, come nel caso dei dati dei registri interni attualmente utilizzati dall'UST.

Ciclo di ragionamento induttivo-deduttivo

È importante sottolineare che i due approcci all'analisi, ovvero il ragionamento induttivo e deduttivo, sono *complementari e devono essere usati in modo iterativo* sia per ottenere un miglioramento continuo, sia per prendere decisioni e definire politiche basate sui dati. Ciò implica che sarà necessario affiancare ai metodi di analisi attualmente utilizzati dall'UST quelli di analisi complementari.

Ad esempio, le informazioni (conoscenze) elaborate attraverso un ragionamento induttivo non possono spiegare se e perché queste scoperte siano utili e in quale misura siano valide. È necessario utilizzare gli strumenti di conferma del ragionamento deduttivo per confermare tali scoperte e valutare la qualità delle decisioni basate su questi risultati.

Primo obiettivo strategico

Il primo obiettivo strategico del programma pluriennale della statistica federale (2016–2019) è fornire una produzione statistica di base che sia costante e risponda alle esigenze degli utenti. Tra le varie fasi di attuazione è stato proposto di elaborare linee guida sui potenziali usi delle nuove fonti di dati, sensibilizzando il pubblico sul fatto che la «data innovation» **è un fattore strategico fondamentale**.

Inoltre, il primo obiettivo strategico sarà progressivamente riesaminato e aggiornato in funzione dell'attuazione delle sue fasi descritte qui di seguito.

Obiettivo strategico 1: elaborare linee guida sulla «data innovation» e indagare se i metodi di analisi complementari possono essere applicati alle fonti di dati esistenti (o tradizionali) e/o a fonti nuove (o non tradizionali), incrementando e/o completando la produzione statistica di base esistente alla quale sia sensato applicare la «data innovation».

A tal fine, nell'ambito di progetti pilota interni all'UST, la **prima fase** consiste nell'applicare metodi di analisi complementari alle fonti di dati primarie interne all'UST che esistono già (tradizionali) e, se esistono, alle fonti di dati secondari identificabili già appaiate. Pertanto, si tratta di aumentare e/o integrare ciò che esiste, ovvero utilizzare in modi nuovi le fonti di dati esistenti per acquisire esperienza pratica e di preparare un inventario, per esempio, delle sfide, risorse, competenze e tecnologie esistenti. Questo permetterà di applicare la «data innovation» e ottenere rapidi benefici.

Se ciò è possibile, segue una **seconda fase** nella quale si incrementa e/o integra la produzione statistica esistente dell'UST con l'ausilio della «data innovation» risultante dall'applicazione di metodi di analisi complementari ai dati secondari supplementari già utilizzati dall'UST.

Nella **terza fase** i metodi di analisi complementare possono essere applicati solo ai nuovi dati secondari (finora inutilizzati dall'UST) per esaminare e creare nuove informazioni statistiche e indagini in settori statistici particolari.

La sequenza di fonti di dati in ordine decrescente da preferire per la strategia di digitalizzazione dell'UST è la seguente:

1. fonti di dati primarie interne all'UST e fonti di dati secondarie identificabili e già appaiate (a condizione che siano già utilizzate nella produzione statistica attuale dell'UST);
2. fonti di dati secondarie supplementari già in uso all'UST;
3. nuove fonti di dati secondarie (finora inutilizzate dall'UST).

In generale, lo scopo di questa sequenza è accrescere l'efficienza e ridurre i costi e l'onere amministrativo che grava sulle imprese e le persone, consentendo una migliore comprensione dei metodi, delle tecnologie e degli strumenti senza creare inutili complicazioni e problemi legali, tecnologici (IT) e di capacità che sorgerebbero se si cominciasse direttamente con la terza fase. Oltre a ciò, l'abituarsi ai metodi di analisi complementari offre la possibilità di perfezionare l'abilità di reperire nuove fonti di dati che si adattino allo scopo specifico della produzione statistica ufficiale.

Secondo obiettivo strategico

L'innovazione dei dati avrà bisogno di un cambiamento di paradigma grazie alla combinazione dei ragionamenti deduttivo e induttivo, che inevitabilmente determinerà nuove produzioni e comunicazioni delle statistiche ufficiali.

È essenziale sottolineare che forse non si troverà un approccio metodologico generico comune per tutte le possibilità d'uso nella produzione statistica ufficiale degli approcci di «data innovation» menzionati precedentemente. Le sfide e opportunità sono specifiche per ciascuna applicazione di «data innovation» e tipologia di dati. Sarà quindi necessario adattare e ampliare le strutture degli attuali processi all'UST e i relativi modelli.

In questo contesto, **comunicare** in merito alle questioni qui trattate diventa fondamentale. Questa riflessione conduce al secondo obiettivo strategico.

Obiettivo strategico 2: sviluppare e attuare le misure di comunicazione interne ed esterne all'UST per accrescere la consapevolezza che la «data innovation», insieme al relativo nuovo paradigma, costituisce un valore aggiunto per le statistiche ufficiali.

Considerazioni generali

Le considerazioni generali dell'UST per realizzare le fasi di «data innovation» descritte precedentemente (nonché tutte le considerazioni su questa nozione) devono rispettare i valori fondamentali seguenti: confidenzialità dei dati, politiche di protezione dei dati, sicurezza dei dati, gestione dei dati e fiducia del pubblico.

Realizzazione degli obiettivi strategici

Per conseguire questi obiettivi strategici e concretizzare la l'innovazione dei dati sono necessarie le seguenti azioni specifiche:

- valutare la necessità di adattare e ampliare i criteri concettuali correnti dell'UST per ciò che concerne la qualità dei dati al fine di renderli compatibili con la «data innovation» (per es. con i metodi di analisi complementari quali l'analisi predittiva);
- esaminare l'eventualità di richiedere nuove competenze al personale dell'UST per quanto riguarda i metodi di analisi complementari, le tecnologie e gli strumenti necessari (IT) per realizzare la «data innovation». Di conseguenza, la prima fase e i relativi progetti pilota summenzionati sono fondamentali per definire le competenze necessarie per la «data innovation». Non appena sarà effettuata questa valutazione, l'UST potrà (se necessario):
 - per esempio, ideare in collaborazione con società private e/o università nuovi corsi di formazione (formazione continua) per il personale interno all'UST, che comprendano i metodi di analisi complementari, le competenze trasversali, quali creatività, collaborazione, creatività, risoluzione dei problemi, comunicazione e narrazione (forza lavoro qualificata);
 - definire e introdurre l'uso di un laboratorio di dati interno e/o esterno all'UST («sandbox») per sviluppare nuovi set di conoscenze necessarie all'utilizzo di metodi di analisi complementari;
 - stabilire i requisiti per adattare le tecnologie (IT), le infrastrutture e gli strumenti interni e/o esterni all'UST;
- studiare le questioni legali connesse alla «data innovation», per es. i requisiti e gli incentivi legali per ottenere il consenso del settore privato (proprietà dei dati) sui dati secondari senza compromettere gli interessi economici dei rispettivi proprietari (nel quadro della summenzionata terza fase);
- discutere e scambiare informazioni con tutte le parti in causa in seno all'UST, in seno ai sistemi statistici e alla comunità di ricerca.

Per l'UST, i partenariati all'interno della Confederazione, del settore industriale e accademico nonché della comunità statistica costituiscono un fattore critico per raggruppare tutte le competenze ed esperienze che saranno necessarie ad affermare la strategia della «data innovation».

Ordinazioni

Tel. 058 463 60 60

Fax 058 463 60 61

order@bfs.admin.ch

Prezzo

gratuito

Download

www.statistica.admin.ch (gratuito)

Numero UST

1789-1700

**La statistica
conta per voi.**

www.la-statistica-conta.ch