
HUB DELLA STATISTICA PUBBLICA

DOCUMENTO DI PROGETTO

1. L'idea

1.1 Oggetto e finalità

Il progetto “*Hub della statistica pubblica*” rappresenta un sistema per la condivisione, integrazione e diffusione di macrodati prodotti da soggetti appartenenti al Sistan o da altri enti produttori di statistiche che svolgono funzioni o servizi di interesse pubblico, caratterizzato dallo sviluppo di un *data warehouse* distribuito quale unico punto di accesso per la navigazione di tutte le informazioni e liberamente interrogabile dagli utenti esterni tramite un'interfaccia web.

L'idea progettuale deriva dalla sempre più sentita necessità di garantire una stretta cooperazione tra l'Istituto nazionale di statistica e gli altri soggetti produttori di dati di interesse per la collettività, al fine di promuovere l'utilizzo di concetti, di classificazioni e di metodi condivisi, nonché di assicurare maggiore coerenza e migliore comparabilità fra statistiche prodotte a livello nazionale, offrendo, altresì, agli utenti finali la possibilità di interrogare in tempo reale le informazioni residenti nei singoli data base dei soggetti aderenti al progetto.

Il coordinamento dell'intero progetto è a cura dell'Istat, che sovrintende alla mappatura dei dati di interesse e allo sviluppo del sistema di diffusione, monitorando la rispondenza dei dati e dei metadati ai requisiti di qualità propri della statistica ufficiale.

1.2 Caratteristiche tecniche

L'Hub della statistica pubblica è progettato come un sistema federato di basi di dati di diffusione distribuite ed è composto da:

- un nodo centrale, denominato *hub*, per l'interconnessione, la ricerca e presentazione dei dati;
- un insieme di nodi locali (o periferici) in cui sono memorizzati i dati da diffondere e la cui gestione e sviluppo è affidata alla responsabilità dei singoli enti partecipanti al progetto.

Il sistema si basa sullo standard “Statistical Data and Metadata eXchange” (ISO IS-17369), sviluppato da sette organizzazioni internazionali (Bank for International Settlements, European Central Bank, Eurostat, International Monetary Fund, Organisation for Economic Cooperation and Development, United Nations Statistical Division, the World Bank) e utilizzato già da diversi anni nell'ambito del Sistema Statistico Europeo per lo scambio di dati tra gli Istituti nazionali di statistica, le altre agenzie statistiche nazionali ed Eurostat.

L'accesso all'informazione avviene attraverso il nodo centrale hub che, su richiesta degli utenti, interagisce con i nodi periferici per estrarre i dati e i metadati richiesti, presentandoli in forma leggibile (tabelle, grafici, mappe, ecc.). Il sistema è progettato in modo tale da permettere la generazione di tabelle statistiche sintetiche, anche di livello territoriale, utili per soddisfare le esigenze dell'utenza anche non specialistica.

Il sistema, oltre a permettere l'accesso agli utenti attraverso un'interfaccia grafica di navigazione, implementa anche l'accesso via API (Application Program Interface), consentendo inoltre l'estrazione delle informazioni in vari formati (SDMX-ML, CSV, JSON, RDF Data Cube, StatDCAT) e facilitando così il percorso verso il livello alto degli Open Data.

Il progetto viene realizzato secondo un approccio “progressivo”, così da garantire risultati in tempi certi e ridurre al minimo i fattori di rischio.

1.3 Obiettivi e risultati attesi

L'obiettivo principale del progetto è quello di favorire la standardizzazione dei processi di diffusione di informazioni statistiche di interesse per la collettività e agevolare l'accesso ai relativi dati e metadati. Allo stesso tempo il progetto persegue l'obiettivo di garantire la qualità dei dati messi a disposizione, attraverso un "governo della qualità" che assicura un'attività di selezione e controllo dei requisiti dei dati prima delle loro immissione nelle pertinenti aree del sistema.

I risultati attesi che il progetto si prefigge sono dunque quelli di:

- realizzare un unico punto di accesso (Hub) in cui ricercare e visualizzare dati prodotti da enti diversi;
- implementare l'interoperabilità semantica, così come sollecitato nell'Agenda Digitale, attraverso armonizzazione dei dati e metadati;
- standardizzare il processo di diffusione utilizzando lo standard SDMX (ISO IS-17369);
- favorire il percorso verso gli *open (statistical) data* permettendo un accesso all'informazione secondo le necessità delle diverse tipologie di utenti;
- condividere le *best practices* sviluppate a livello nazionale ed internazionale.

2. Descrizione delle attività e impegni delle parti

2.1 Attività del progetto

L'applicazione dello standard SDMX con i moduli software già disponibili nella comunità statistica internazionale riduce l'esigenza di sviluppo di software complessi e garantisce standard elevati di qualità, affidabilità e riusabilità.

Costituiscono attività essenziali del progetto:

- configurazione dell'infrastruttura informatica;
- l'analisi ed il controllo della qualità dell'informazione;
- la modellazione dei dati da diffondere;
- l'eventuale mappatura dei dati nei nodi locali;
- la definizione del piano integrato di diffusione;
- l'attivazione di forme strutturate per la gestione dei servizi all'utenza.

Le attività, che saranno svolte nell'ambito del progetto, saranno articolate nelle seguenti macro fasi:

- *definizione dei contenuti ed integrazione con il sistema di diffusione dell'ISTAT I.Stat*: prevede la condivisione dei contenuti da diffondere e del relativo piano integrato della diffusione, nonché degli standard fissati dall'ISTAT in merito ai metadati per la diffusione (classificazioni, denominazioni variabili, ecc.);
- *governo della qualità*: comprende l'analisi strutturata e la gestione coordinata delle attività di selezione e di controllo dei requisiti di qualità dei dati statistici al fine della loro immissione nelle pertinenti aree del sistema;
- *implementazione dell'infrastruttura informatica, e delle funzionalità e modalità di navigazione dei dati*: attiene allo sviluppo dei software necessari sia per il nodo hub che per i nodi locali. Tale sviluppo potrà essere condotto in collaborazione fra le Parti, secondo le direttive tecniche e la supervisione dell'ISTAT;
- *configurazione dei nodi locali e gestione del processo di diffusione*: potrà avvenire anche attraverso modalità di mappatura dei dati e metadati, in formato SDMX. Nell'ambito di tale macro fase, sulla base delle linee guida sulla modellazione e mappatura dei dati che saranno rilasciate dall'ISTAT, i responsabili dei nodi locali: a) cureranno l'insieme dei metadati che dovranno essere esposti; b) provvederanno al caricamento nei database dei dati e metadati da diffondere; c) acquisiranno il know-how sia per la gestione del nodo locale sia per la modellazione e mappatura;

- *definizione degli aspetti di comunicazione e di governace dei servizi all'utenza*: comprende la progettazione delle modalità di comunicazione (web design, layout grafico, ecc.) e delle specifiche *utility* da rilasciare all'utenza (Faq, tutorial, ecc.).

Le attività sopra specificate sono svolte sotto il coordinamento tecnico-organizzativo dell'ISTAT, secondo gli standard per la diffusione adottati da quest'ultimo e tenuto conto delle indicazioni del Comitato di coordinamento le cui funzioni sono descritte nel paragrafo 4.

2.2 Impegni delle parti

Ciascun ente si impegna a partecipare mettendo a disposizione le necessarie risorse tecniche e di personale, essenziali alla realizzazione delle varie fasi previste nel progetto.

Gli oneri finanziari derivanti dalle attività svolte da ciascun ente sono a carico dell'ente stesso.

L'ISTAT si impegna alla realizzazione dei nodi periferici l'ISTAT renderà disponibile i software capaci di gestire la connessione con il nodo hub e/o di realizzare un database di diffusione (quest'ultimo software per quegli enti che attualmente diffondono i dati su supporto cartaceo o attraverso file, quali Excel, CSV, SAS, ecc.).

Inoltre, e solo per un periodo transitorio, l'ISTAT potrà ospitare in modalità *cloud* quegli enti che vogliono iniziare una fase sperimentale, non avendo ancora a disposizione l'architettura informatica necessaria.

Ciascun ente aderente al progetto si impegna a:

- predisporre l'architettura informatica, composta principalmente da un server di database e un server web esposto su Internet, oltre gli opportuni collegamenti fra i server;
- rendere disponibile al nodo Hub l'accesso al database attraverso web service;
- selezionare, in collaborazione con l'ISTAT, i contenuti da diffondere;
- modellare i contenuti selezionati secondo le linee guida definite dall'ISTAT, orientate all'armonizzazione dei concetti statistici e delle relative liste di codici;
- mappare i dataset da diffondere con i metadati definiti in fase di modellazione;
- supervisionare lo stato dei propri server al fine da garantire continuità operativa.

3. Requisiti e modalità di adesione al progetto

3.1 Requisiti dei soggetti interessati

Possono aderire al progetto gli enti facenti parte del Sistema statistico nazionale e qualsiasi soggetto interessato le cui attività soddisfino i requisiti e i criteri di seguito descritti:

- a. svolgimento di funzioni e servizi di interesse pubblico;
- b. svolgimento di attività connesse al raggiungimento degli obiettivi del progetto, nonché utili al potenziamento e/o incremento della capacità informativa ed organizzativa del sistema;
- c. svolgimento di attività di raccolta dati in nuovi settori di informazione, ovvero in grado di rendere disponibili informazioni complementari ed integrabili con altre già disponibili nel sistema;
- d. capacità di contribuire significativamente alla creazione e gestione di sistemi informativi statistici.

3.2 Procedura di ammissione

Gli interessati dovranno presentare all'Istat una formale istanza di adesione al progetto (allegati 1 e 2), corredata degli elementi informativi necessari per la valutazione della sussistenza dei requisiti e dei criteri di ammissibilità richiesti.

La Direzione centrale dell'ISTAT per la pianificazione strategica, l'indirizzo del Sistema statistico nazionale, le relazioni istituzionali e gli affari internazionali (DCPS) cura l'istruttoria dell'istanza, raccogliendo ed esaminando la documentazione fornita dallo stesso richiedente. Qualora esigenze istruttorie lo rendano opportuno, la medesima Direzione provvede a integrare le informazioni disponibili richiedendo

all'interessato ogni altro elemento di valutazione ritenuto utile. Sulla base degli elementi raccolti in sede istruttoria.

A seguito dell'accoglimento della richiesta di adesione, viene dettagliato il programma operativo di avvio delle attività, in linea con quanto indicato al paragrafo 2 e con quanto dichiarato nella lettera di adesione relativamente al contributo informativo che si ritiene di poter fornire al progetto. In particolare, l'interessato si impegna allo svolgimento delle attività come definite nel precedente paragrafo 2, garantendo l'adesione ai principi del Codice italiano delle statistiche ufficiali, nonché il rispetto della normativa sul segreto statistico e sulla tutela della riservatezza nei trattamenti, a fini statistici, dei dati personali.

Nella lettera di adesione dovranno essere comunicati i nominativi di due rappresentanti (uno con competenze relative agli aspetti dei contenuti informativi, l'altro con competenze informatiche) che parteciperanno al Comitato di coordinamento, di cui al successivo paragr. 4.

È facoltà dell'Ente procedere alla sostituzione dei propri rappresentanti, dandone tempestiva comunicazione all'Istat.

4. Comitato di coordinamento

Al fine di assicurare la realizzazione delle attività previste nel progetto, è istituito un Comitato di coordinamento composto da quattro rappresentanti dell'ISTAT e due rappresentanti di ciascun ente partecipante al progetto (uno con competenze relative agli aspetti dei contenuti informativi, l'altro con competenze informatiche). Le funzioni di coordinatore saranno svolte da un membro scelto tra i rappresentanti dell'ISTAT.

Il Comitato di coordinamento ha il compito di:

- a) valorizzare le esperienze di ciascun ente attraverso la condivisione di *best practices*;
- b) monitorare l'andamento dei lavori al fine di verificare che gli stessi procedano nel rispetto degli impegni assunti da ciascun soggetto aderente e della tempistica concordata;
- c) definire le iniziative di pubblicizzazione e di valorizzazione dei risultati del progetto a livello nazionale e territoriale, volte ad attirare l'interesse dell'utenza e degli altri Enti ed Uffici di statistica del Sistema statistico nazionale al fine di accrescere progressivamente nel tempo l'ampiezza dell'informazione diffusa, in un'ottica di crescente integrazione e coerenza.

Il Comitato, nel caso lo riterrà necessario, potrà procedere alla costituzione di gruppi di lavoro e/o nuclei locali anche su base regionale con il compito di condividere know-how ed expertise e facilitare l'inserimento dei vari enti, tenendo conto delle competenze tecniche richieste per le attività da svolgere. La costituzione di ciascun gruppo di lavoro e/o nucleo locale, comprensiva dell'indicazione dei nominativi dei componenti e dei compiti ad esso attribuiti, sarà oggetto di specifica formalizzazione attraverso verbalizzazione della seduta del Comitato in cui è avvenuta.

È facoltà del Comitato invitare a partecipare alle proprie riunioni, a titolo gratuito, esperti individuati in relazione all'argomento da trattare.

Il Comitato si riunisce su richiesta di almeno una delle parti, e comunque almeno ogni tre mesi.

La partecipazione al Comitato è a titolo gratuito. Ciascuna Parte si fa carico delle eventuali spese di missione relative ai propri rappresentanti. Le videoconferenze costituiscono lo strumento privilegiato per garantire la presenza di membri o esperti esterni la cui sede di lavoro non sia ubicata nella città in cui si svolgono le riunioni.

5. Disposizioni finali

5.1 Diritti proprietà

La titolarità dei diritti di proprietà sugli strumenti di diffusione (sia fisici che digitali) sviluppati nell'ambito del Progetto rimangono di proprietà di ciascuna parte aderente.

5.2 Trattamento dei dati e segreto statistico

I trattamenti di dati rientranti nell'ambito del progetto sono effettuati nel rispetto delle disposizioni in materia di segreto statistico di cui all'art. 9 del d.lgs n. 322 del 1989.

Le attività previste non comportano la comunicazione tra le Parti di dati personali, né di dati comunque coperti dal segreto statistico. L'eventuale trattamento di dati coperti dal segreto statistico è effettuato esclusivamente dal soggetto che li detiene e che ne è titolare.

Le informazioni statistiche saranno diffuse secondo standard comuni e nel rispetto delle linee guida per la diffusione dei dati statistici prodotti dal Sistema statistico nazionale.

5.3 Risoluzione recesso

Ciascun soggetto aderente, qualora intervengano fatti o provvedimenti che modifichino la situazione esistente all'atto della richiesta di adesione, ha facoltà di recedere in qualunque momento dagli impegni assunti.

In tali ipotesi, gli interessati rappresentano le proprie motivazioni mediante comunicazione scritta da inviare all'Istat con un preavviso di almeno 30 giorni.

In caso di gravi inadempienze agli impegni assunti da parte degli aderenti al progetto, il comitato di coordinamento di cui al paragrafo 7 potrà invitare la Parte inadempiente a provvedere all'adempimento entro un termine prefissato. Decorso il predetto termine, se la Parte inadempiente non avrà provveduto, gli impegni assunti tra le Parti si intenderanno risolti.

Quadro normativo di riferimento

- *Regolamento (UE) 2015/759 del Parlamento europeo e del Consiglio del 29 aprile 2015 che modifica il regolamento (CE) n. 223/2009, relativo alle statistiche europee;*
 - *Legge 7 agosto 1990, n. 241 e s.m.i. “Nuove norme in materia di procedimento amministrativo e di diritto di accesso ai documenti amministrativi”, in particolare l’art. 15 in materia di “Accordi tra pubbliche amministrazioni”;*
 - *Legge 6 novembre 2012, n. 190 s.m.i. “Disposizioni per la prevenzione e la repressione della corruzione e dell’illegalità nella pubblica amministrazione”;*
 - *Decreto legislativo 6 settembre 1989, n. 322, e successive modificazioni e integrazioni, “Norme sul Sistema statistico nazionale e sulla riorganizzazione dell’Istituto nazionale di statistica”;*
 - *Decreto del Presidente della Repubblica 7 settembre 2010, n. 166, “Regolamento recante il riordino dell’Istituto nazionale di statistica”;*
 - *Decreto legislativo 30 giugno 2003, n. 196, e successive modifiche e integrazioni, “Codice in materia di protezione dei dati personali”;*
 - *“Codice di deontologia e di buona condotta per i trattamenti di dati personali a scopi statistici e di ricerca scientifica effettuati nell’ambito del Sistema statistico nazionale” (All. A.3 del Codice in materia di protezione dei dati personali - d.lgs. 30 giugno 2003, n. 196);*
 - *Direttiva n. 10 del Comitato di indirizzo e coordinamento dell’informazione statistica (COMSTAT), con la quale è stato adottato il Codice italiano delle statistiche ufficiali il quale stabilisce i principi e i criteri da adottare per assicurare elevati standard di qualità delle statistiche ufficiali.*
-