



**Audizione dell'Istat presso la Commissione per la garanzia della qualità  
dell'informazione statistica**

**Qualità e diffusione dei microdati di fonte amministrativa e controllo e  
promozione della qualità dei dati statistici prodotti nelle ONAs e nel  
Sistan**

**Commissione per la garanzia della qualità dell'informazione statistica  
Roma, 28 settembre 2018**

## Indice

1	Introduzione.....	4
2	La revisione della lista delle ONAs .....	5
3	Il programma di audit.....	7
4	Iniziative formative per gli enti del Sistan .....	9
4.1	Il sistema informativo dell’Istat per la condivisione delle statistiche geografiche ufficiali .....	10
4.2	Il progetto ARCH.I.ME.DE.....	10
4.3	Video-lezioni su temi di interesse .....	11
4.4	Corso di alta formazione .....	12
5	Hub della statistica pubblica.....	12
5.1	Lo strumento per la diffusione di qualità delle statistiche del Sistan .....	12
5.2	L’evoluzione del progetto .....	13
6	Qualità dei dati amministrativi e dei registri statistici .....	14
6.1	Programmazione e acquisizione, conservazione e rilascio dei dati amministrativi.....	14
6.2	Documentazione della qualità dei dati amministrativi.....	15
6.3	Qualità dei registri statistici costruiti di fonti amministrative .....	21
7	Il miglioramento della trasparenza all’Istat.....	22
7.1	Avanzamenti delle politiche di revisione .....	22
7.2	I Quality report e le Schede standard di qualità .....	23
8	Miglioramenti nelle politiche di accesso ai microdati .....	25
8.1	Conservazione e accessibilità dei microdati .....	26
8.2	L’informazione sui microdati .....	27
8.3	Il Laboratorio ADELE e l’accesso da remoto .....	27
8.4	Le Linee guida accesso ai microdati.....	28



## 1 Introduzione

Uno dei cardini della politica per la qualità dell'Istat è il Codice delle statistiche europee<sup>1</sup>. L'adozione del Codice da parte del Sistema statistico europeo (SSE) nel maggio 2005 e le sue revisioni nel settembre 2011 e nel novembre 2017 hanno segnato un passo importante per promuovere i valori comuni nell'ambito dell'SSE e rafforzare la fiducia nelle statistiche.

Per valutare l'ottemperanza da parte degli Istituti nazionali di statistica (Ins) a tali principi, Eurostat ha promosso due round di Peer review (2006-2008 e 2014-2015).

In particolare, il secondo round, che in Italia si è svolto a giugno 2015 ha dato luogo a una serie di raccomandazioni per ognuna delle quali gli Istituti nazionali di statistica hanno individuato una lista di azioni di miglioramento, oggetto di un monitoraggio annuale da parte di Eurostat.<sup>2</sup> Per quanto riguarda l'Italia, la Peer review ha evidenziato 21 raccomandazioni e altrettante azioni di miglioramento suddivise in sotto-azioni (62 in totale).

L'Eurostat ha completato il monitoraggio sulla attuazione delle azioni di miglioramento agli inizi del 2018 su tutti i paesi dell'UE. Da tale monitoraggio emerge che l'Italia è tra i paesi ad aver completato tali azioni come programmato. Il rapporto di monitoraggio dell'Eurostat è stato sottoposto al Comitato del Sistema Statistico Europeo nella sua riunione del 17 maggio 2018.

In considerazione della revisione del Codice delle statistiche europee e alle esigenze di continuare a monitorare l'attuazione del codice europeo nel Comitato dell'SSE si sta discutendo per l'avvio di una nuova fase di Peer review presumibilmente nel 2019.

A seguito della lista delle Raccomandazioni, l'Istat ha allargato lo sguardo oltre le statistiche europee, cercando di rispettare non solo i *milestones* di verifica proposti da Eurostat, ma anche di applicare (o adattare) alcune delle raccomandazioni alle esigenze di monitoraggio della qualità delle statistiche nazionali.

Pertanto, ripercorrere alcune delle azioni svolte per rispondere alle raccomandazioni emerse dalla peer review consente di illustrare le principali azioni intraprese dall'Istat per il miglioramento della qualità non solo delle statistiche europee ma di tutte le statistiche prodotte, non solo dall'Istat ma anche dal Sistan.

---

<sup>1</sup> Le statistiche europee sono quelle, come previsto dalla legge statistica europea, stabilite nel Programma statistico europeo e rese ulteriormente operative nel Programma di lavoro annuale, che sono quindi trasmesse alla Commissione Europea (Eurostat).

<sup>2</sup> I rapporti delle risultanze delle Peer review e la lista delle azioni di miglioramento per ciascun paese sono pubblicate sul sito di Eurostat (<http://ec.europa.eu/eurostat/web/quality/peer-reviews>).

Il documento ripercorre e illustra, concentrandosi sulle specifiche esigenze informative espresse dalla Commissione, alcune di queste attività.

In particolare:

- ✓ revisione della lista delle ONAs (**Improvement action 2 Further initiatives on coordination** Re-profile authorities responsible for the production of statistics relevant at European (ONAs), national and territorial level);
- ✓ il programma di audit (**Improvement action 4 Strengthening knowledge and know-how in the ONAs on quality management, measurement and assessment** Start out tailor-made audits on specific products / processes through a pilot phase with a view to plan at least 3 yearly audits);
- ✓ il programma di formazione (**Improvement action 4** Definition of a targeted plan of initiatives for Sistan relevant bodies, focused on transferring knowledge and know-how on quality management, measurement and assessment),
- ✓ l'Hub della statistica pubblica (**Improvement action 5 Development and implementation of Integrated dissemination system** "Istat will develop a plan of initiatives for the interconnection and the integrated dissemination of statistics produced by Sistan bodies, as well as for guaranteeing the access to products of the National Statistical Program");
- ✓ programma di modernizzazione e in particolare attività di acquisizione, conservazione, rilascio e qualità dei dati amministrativi e qualità dei registri statistici costruiti da fonti amministrative (**Improvement action 6 Implementation of Modernisation programme**).
- ✓ il miglioramento della trasparenza all'Istat (Improvement action 19 Development of revision policy e Improvement action 20 Improvement of Quality documentation);
- ✓ conservazione e accessibilità dei microdati (Improvement action 21 Release of a website devoted to microdata to enhance their accessibility (data archive));
- ✓ la realizzazione delle Linee guida di accesso ai microdati (**Improvement action 21 Review of dissemination policy** Proposal for amending national relevant legislation in order to promote access to microdata, in accordance with the Data Protection Authority, through virtual labs among others).

## 2 La revisione della lista delle ONAs

L'Istat ha avviato, nel luglio 2017, una ricognizione tra le ONA, attuali e potenziali, con lo scopo di identificare i reali produttori di statistiche europee, cioè coloro che sono responsabili, a livello nazionale, della produzione di una parte specifica di statistiche europee e della loro trasmissione a Eurostat (e non ad altre Direzioni Generali della Commissione o a organismi internazionali) sia diretta che indiretta (cioè attraverso l'Istat o altra amministrazione nazionale).

La ricognizione ha avuto luogo a seguito di incontri bilaterali con le diverse Amministrazioni centrali e i diversi Enti appartenenti al Sistan; ha coinvolto 18 soggetti di cui 15 già nella lista delle ONA e 3 individuati a seguito delle riunioni tra l'Istat e gli enti. L'esclusione di quattro soggetti dalla lista ONA è dipesa dalla soppressione dell'ente stesso, dall'accorpamento in altro ente o dalla verifica della mancata produzione di statistiche europee.

Tale processo si è reso necessario in relazione ai seguenti elementi:

- ✓ la necessità di verificare la lista delle ONA che risale al 2009. Sulla base dell'art.5 (2) della legge statistica, la Commissione (Eurostat) pubblica sul suo sito la lista delle ONA designate dagli Stati membri, quali responsabili dello sviluppo e produzione e diffusione delle statistiche europee. Spetta allo Stato Membro aggiornare tale lista e comunicarla via Rappresentanza UE all'Eurostat;
- ✓ l'opportunità di dare un riscontro al Comstat della situazione sulle ONAs e sui necessari adeguamenti;
- ✓ l'opportunità di una riflessione sul ruolo e sulle caratteristiche dei reali produttori di statistiche europee, nonché una riflessione accurata su ruolo e caratteristiche dei fornitori di dati da fonti amministrative per la produzione di statistiche europee. A differenza del produttore di statistiche europee, il fornitore detiene una fonte di dati e fornisce tali dati all'Istat o ad altra amministrazione del Sistan per il loro trattamento a fini statistici per la produzione di statistiche europee;
- ✓ la necessità di avviare l'analisi dei flussi dei dati in ottemperanza ai regolamenti europei e la sensibilizzazione su un processo di monitoraggio per la compliance ai regolamenti europei di competenza e il ruolo di coordinamento dell'Istat secondo quanto stabilito dalla Legge statistica europea.

Per la ricognizione è stato progettato e realizzato da parte dell'Istat un questionario online (modulo ad hoc, ciascuno abbinato a un atto normativo europeo) volto a raccogliere informazioni relative alla produzione e/o fornitura di dati, alla modalità e frequenza di trasmissione e i costi di produzione. In particolare la ricognizione ha rilevato:

- ✓ quali statistiche europee vengono prodotte dalle potenziali e attuali ONA; quali di queste vengono trasmesse direttamente o indirettamente a Eurostat e con quale modalità (piattaforma Edamis, file excel, PDF);
- ✓ quali dati sono forniti dagli enti coinvolti all'Istat o ad altre amministrazioni che le utilizzano per la produzione di statistiche europee;
- ✓ quali dati vengono trasmessi o forniti alle altre Direzioni Generali della Commissione o alle agenzie dell'UE;
- ✓ I costi per la produzione delle statistiche europee;
- ✓ Frequenza della produzione e trasmissione dei dati.

Dei 18 enti 15 sono risultati *produttori* di statistiche europee che trasmettono direttamente o indirettamente a Eurostat. La lista aggiornata è stata inviata ad Eurostat a dicembre 2017 secondo la procedura indicata. Tale lista è quindi stata pubblicata nello stesso mese sul sito di Eurostat<sup>3</sup>.

### 3 Il programma di audit

Per la definizione del programma di audit metodologico sulle statistiche europee prodotte dalle Altre Autorità Nazionali (ONAs), si è proceduto durante tutto il 2017 alla predisposizione degli strumenti idonei a questo obiettivo<sup>4</sup>. Tali strumenti sono:

- ✓ lo standard di riferimento per la conduzione dell'audit, ossia le "Linee guida per la qualità delle statistiche del Sistema Statistico Nazionale";
- ✓ il questionario di valutazione per la verifica dell'ottemperanza verso i principi delle linee guida;
- ✓ lo standard per la relazione dei risultati della procedura di audit (template del report finale di valutazione).

Le linee guida contengono i principi da soddisfare per assicurare la qualità delle statistiche prodotte e, per ciascun principio, un insieme di suggerimenti di attività e indicatori standard di qualità per il monitoraggio e la valutazione, all'interno dei quali possono essere scelti quelli più idonei al contesto produttivo dell'ente.

Al fine di produrre un manuale applicabile agli enti del Sistan, la prima versione delle linee guida è stata diffusa presso un insieme allargato di ONAs, attraverso una consultazione che aveva l'obiettivo di raccogliere feedback sulla chiarezza, applicabilità e completezza del manuale.

Il questionario di audit è lo strumento che viene utilizzato per rilevare le pratiche adottate, la qualità del processo e delle statistiche prodotte. Esso rispecchia da vicino la struttura e il contenuto delle linee guida. Infatti, esso contiene: i principi delle linee guida e, a seguire, i quesiti per la valutazione.

Oltre ai quesiti per la rilevazione delle pratiche adottate dal processo, il questionario contiene delle domande chiave, chiaramente identificabili attraverso la grafica. Si tratta di quesiti di valutazione che esplorano le diverse sfaccettature della qualità: la valutazione degli errori campionari e non campionari che si generano nelle fasi del

---

<sup>3</sup> [https://ec.europa.eu/eurostat/documents/747709/753176/20180719\\_List\\_ONAs\\_LV.pdf](https://ec.europa.eu/eurostat/documents/747709/753176/20180719_List_ONAs_LV.pdf)

<sup>4</sup> Tutti gli strumenti sono presentati e messi a disposizione sul sito dell'Istat al seguente indirizzo: <https://www.istat.it/it/metodi-e-strumenti/strumenti-per-la-qualit%C3%A0/strumenti-per-la-qualit%C3%A0-nel-sistan>

processo, la valutazione della qualità delle fasi del processo produttivo statistico e la valutazione della qualità del prodotto, ossia delle dimensioni della qualità.

Il questionario di valutazione è stato testato su due processi di due enti:

- ✓ la “Statistica annuale della produzione e consumo di energia elettrica”, prodotta da Terna SpA (Novembre 2017)
- ✓ il “Sistema dei conti pubblici territoriali” prodotta dall’Agenzia per la Coesione Territoriale (Dicembre 2017).

Quest’ultimo ente è stato scelto al di fuori delle ONA sia per la peculiarità dei processi trattati che per la disponibilità fornita.

Le interviste sono state condotte dal team della qualità dell’Istat e finalizzate a raccogliere feedback sulla comprensione di quesiti, la loro adeguatezza e più in generale sull’efficacia del questionario nella valutazione dell’ottemperanza ai principi delle linee guida.

A conclusione delle interviste di audit, il team dell’Istat ha predisposto, in via sperimentale, il report finale di valutazione, secondo il template predefinito, con la sintesi delle evidenze emerse dall’intervista e l’identificazione delle azioni di miglioramento da avviare per superare le criticità. I report sono stati trasmessi e condivisi con gli enti che hanno partecipato alla sperimentazione.

A seguito della sperimentazione, il questionario ha subito delle rilevanti modifiche e alcune semplificazioni. Alcuni miglioramenti hanno interessato le stesse linee guida.

Come pianificato, nel 2018 è stato avviato il programma di audit.

I tre processi e relativi enti da coinvolgere nel primo round di audit sono stati individuati sulla base di considerazioni connesse alla disponibilità e rilevanza. Requisito per la selezione del processo è la produzione di dati statistici previsti in ottemperanza di regolamenti europei settoriali.

Per ogni audit si è costituito un team di esperti dell’Istat composto da: un rappresentante del settore della qualità, un rappresentante del settore tematico del processo oggetto di audit, un rappresentante del settore delle metodologie, due rappresentanti del settore affari internazionali e per l’indirizzo del Sistema Statistico Nazionale.

Gli audit si sono svolti a Maggio, Giugno e Luglio rispettivamente presso il Ministero dello Sviluppo Economico (Mise), il Gestore dei Servizi Energetici (GSE) e il Ministero delle Politiche Agricole Alimentari e Forestali (Mipaaf). Le interviste di audit hanno evidenziato che i processi statistici valutati vengono svolti con elevata professionalità e accuratezza e seguendo buone pratiche. In generale, alcune criticità sono state riscontrate nelle procedure interne di documentazione, nella predisposizione di



istruzioni per alcune fasi e nel calcolo di indicatori di monitoraggio, attribuibili a mancanza di tempo e risorse. Inoltre aree di miglioramento sono state identificate nelle procedure e nella documentazione a corredo della diffusione. Le problematiche evidenziate sono state inserite nei report di valutazione accompagnate da relative azioni di miglioramento e una tempistica per la loro realizzazione. Le azioni di miglioramento verranno monitorate dall'Istat nell'arco dei 18 mesi successivi alla procedura.

Il programma di audit deve essere esteso agli altri enti e ai vari processi. Tale progetto non solo rafforza il ruolo di coordinamento dell'Istat ma agevola anche l'accreditamento dell'Ufficio di Statistica dell'ONA all'interno della propria amministrazione.

È allo studio una attività di promozione e coordinamento del *quality reporting*, orientati agli utilizzatori esterni, per le statistiche prodotte dalle ONAs ed eventualmente da altri produttori nazionali rilevanti da affiancare all'attività di audit. I quality report sono dei documenti che accompagnano la diffusione, che descrivono alcuni aspetti salienti delle statistiche diffuse e la loro qualità. Essi migliorano la comprensione e utilizzabilità dei dati diffusi da parte degli utenti, pertanto aumentano la fruibilità delle statistiche prodotte. Così come per l'Istat, i report di qualità da produrre a supporto delle statistiche prodotte da altri enti del sistema, devono tenere conto degli standard europei e dei regolamenti settoriali europei. La proposta sarà sviluppata nel 2019.

#### **4 Iniziative formative per gli enti del Sistan**

L'offerta formativa destinata agli enti del Sistan si inserisce nell'ambito dei compiti assegnati all'Istituto dal decreto legislativo n. 322 del 1989 (art.15). Nel corso del 2016 è stato, inoltre, sottoscritto un Protocollo d'intesa ISTAT-ANCI-UPI che ha, tra i suoi obiettivi, anche quello di sviluppare percorsi formativi per il personale degli uffici di statistica, con finalità prevalentemente informative e di circolazione della conoscenza su alcuni temi di interesse per gli operatori del Sistan a livello provinciale e comunale.

In particolare sono stati organizzati i seguenti eventi:

1. Corso "GISTAT: il sistema informativo dell'Istat per la condivisione delle statistiche ufficiali"
2. Corso "Il progetto ARCH.I.ME.DE (ARCHivi Integrati di Microdati Demografici ed Economici)"
3. Una serie di video-lezioni dedicate a temi di interesse per gli uffici di statistica.

Le prime due iniziative formative rientrano nell'ambito del Protocollo d'Intesa ISTAT ANCI UPI. La terza, invece, è stata organizzata per la XIII Conferenza Nazionale di Statistica.

Tra le diverse metodologie didattiche si è deciso di privilegiare la modalità e-learning che all' Istat si sta sperimentando da qualche anno. Nel 2014 è stata sviluppata la Piattaforma per la formazione statistica, che ospita numerosi contenuti didattici e-learning (video-lezioni, videotutorial, webinar, materiale didattici, ecc.). La piattaforma è stata realizzata allo scopo di arricchire le competenze degli operatori del Sistema statistico nazionale

La progettazione delle iniziative formative è stata realizzata da un *team* costituito da:

- ✓ rappresentanti di ANCI-USCI e UPI che hanno portato le esigenze formative degli uffici di statistica (per quanto riguarda i primi due corsi);
- ✓ esperti tematici dell'ISTAT;
- ✓ esperti di formazione dell'ISTAT.

#### 4.1 *Il sistema informativo dell'Istat per la condivisione delle statistiche geografiche ufficiali*

Sono state realizzate due edizioni del corso "GISTAT: il sistema informativo dell'Istat per la condivisione delle statistiche ufficiali" nel 2017.

Obiettivo del corso è stato quello di presentare i vantaggi offerti dalla tecnologia GIS in ambito di produzione, analisi e diffusione statistica. In particolare sono stati illustrati i concetti di base dell'integrazione tra informazione statistica e geo-spaziale, e presentati alcuni applicativi per la georeferenziazione.

L'iniziativa è stata organizzata in tre moduli: un modulo introduttivo, che si è svolto attraverso un webinar; un modulo dedicato al sistema GIS, che si è articolato in due videolezioni; un modulo sugli strumenti GIS per l'aggiornamento del geo-database dell'Istat.

Sulla piattaforma sono stati inoltre resi disponibili tutti i materiali presentati dai docenti durante i tutorial e la registrazione del webinar. Hanno partecipato al corso 114 risorse. Molto positivi sono stati i feedback forniti dagli iscritti con un giudizio complessivo medio pari a 8,5 (su una scala valoriale che va da 1 a 10), che corrisponde a una valutazione molto buona. Stimolanti anche i commenti e suggerimenti proposti.

#### 4.2 *Il progetto ARCH.I.ME.DE*

Si tratta, in particolare, di una iniziativa di *knowledge sharing* il cui obiettivo è stato quello di offrire una panoramica complessiva dell'archivio e delle relative basi di dati.

Sono state realizzate complessive 6 videolezioni. La prima di introduzione generale al progetto: cos'è il progetto ARCH.I.ME.DE, quali sono le basi di dati disponibili, le criticità e i rischi dell'uso dei dati amministrativi integrati, gli output del progetto. Tre videolezioni sono state dedicate alle singole basi di dati: Precarietà lavorativa, Condizioni

socio economiche delle famiglie e Popolazioni che usano un territorio e mobilità. È stata, inoltre, presentata una lettura statistica dedicata al tema “Studenti e scuole dell’istruzione primaria e secondaria in Italia”. È stato, infine, realizzato un *videotutorial* sulla procedura per richiedere i file integrati prodotti per ogni singola base dati. Oltre alle videolezioni, sono disponibili sulla piattaforma note metodologiche, alcune letture statistiche e link utili.

Hanno partecipato al corso 42 risorse. Anche in questo caso positivi sono stati i feedback forniti, con un giudizio complessivo pari a 8,3.

#### 4.3 *Video-lezioni su temi di interesse*

Per la XIII Conferenza Nazionale di Statistica sono state realizzate le seguenti videolezioni:

1. La storia sbagliata dei NEET
2. Gli indicatori di benessere nella politica economica
3. La qualità dei processi statistici: l’errore di misurazione
4. Commitment on Confidence
5. Le seconde generazioni
6. Progetto HUB della statistica pubblica
7. Il rapporto sulla conoscenza 2018: evidenze empiriche

La selezione degli argomenti è stata orientata da diversi elementi: rilevanza strategica della produzione statistica; enfasi sul profilo istituzionale della statistica ufficiale; esperienze maturate nel campo, argomenti che hanno suscitato interesse in altre occasioni recenti di formazione.

In occasione della Conferenza sono stati aperti ai visitatori anche alcuni webinar su tematiche di innovazione statistica, realizzati in Istat e tenuti da docenti universitari (italiani e internazionali). In particolare:

1. Business Process Automation and Data Processing Workflows. Prof. M. Dumas – University of Tartu (Estonia)
2. How to use big data for Official Statistics. Prof. Li-Chun Zhang - University of Southampton and Statistisk Sentralbyrå
3. Dealing with non-ignorable nonresponse in survey sampling. Prof.ssa Ranalli - Università di Perugia

In generale sulla piattaforma per la formazione statistica - nell’area dedicata agli enti del SISTAN –sono disponibili numerosi webinar realizzati in ISTAT negli ultimi anni su diverse tematiche. Molti webinar sono aperti a tutti (non richiedono, quindi, una previa registrazione).

#### 4.4 Corso di alta formazione

Accanto alle iniziative realizzate per gli uffici del Sistan operanti a livello territoriale, a partire dal 2017, in collaborazione con la Scuola Nazionale dell'Amministrazione (SNA) è stato progettato e sono state finora realizzate 3 edizioni del Corso di alta formazione sulla statistica ufficiale. Si tratta di un corso particolarmente impegnativo: circa 70 ore di lezione in tre mesi, con esercitazioni e casi di studio.

Il programma<sup>5</sup> è stato disegnato elaborando l'impostazione del Codice della statistica ufficiale e quindi dividendo la didattica in tre moduli, trattando materie riconducibili al profilo istituzionale, al sistema di produzione e alla diffusione delle statistiche ufficiali. Nel complesso, inoltre, è stata sottolineata, trasversalmente e in tutte le sezioni formative, l'esigenza di disegnare, produrre e diffondere statistiche di qualità.

Le prime due edizioni hanno visto la partecipazione di 57 persone provenienti da 14 tra Ministeri e altre amministrazioni centrali. La partecipazione è stata buona e costante. La valutazione espressa è stata elevata. Tra i suggerimenti proposti emerge l'esigenza di fornire non solo nozioni, ma anche strumenti e competenze applicabili ai processi ordinari in esercizio negli uffici.

Dopo l'edizione iniziale, inoltre, è emersa la possibilità di allargare l'offerta formativa anche oltre i confini delle amministrazioni ministeriali, a tutte le amministrazioni centrali, alle regioni e ad altre amministrazioni locali di rilievo (ad esempio: le città metropolitane). Il corso è stato riproposto nel 2018 (due edizioni) ed è in programma anche per il 2019.

## 5 Hub della statistica pubblica

### 5.1 Lo strumento per la diffusione di qualità delle statistiche del Sistan

Nella seduta del 5 dicembre 2012 il COMSTAT ha discusso ed espresso pieno sostegno al progetto di costruzione di un data warehouse nazionale per la diffusione dei macrodati prodotti dal Sistema statistico nazionale con i seguenti obiettivi:

- ✓ garantire una stretta cooperazione tra l'Istituto nazionale di statistica e gli altri soggetti produttori di statistiche, al fine di migliorare la qualità nella diffusione in ambito Sistan, promuovendo l'utilizzo di concetti, di classificazioni e standard internazionali (ISO IS-17369 *Statistical Data and Metadata eXchange* standard);
- ✓ assicurare maggiore coerenza e migliore comparabilità fra statistiche prodotte a livello nazionale;

---

<sup>5</sup> [http://www.sistan.it/index.php?id=88&no\\_cache=1&tx\\_ttnews%5Btt\\_news%5D=6487](http://www.sistan.it/index.php?id=88&no_cache=1&tx_ttnews%5Btt_news%5D=6487)

- ✓ offrire agli utenti finali la possibilità di interrogare in tempo reale le informazioni residenti nei singoli data base dei soggetti aderenti al progetto, attraverso un unico punto di accesso, ricerca e presentazione.

La sperimentazione è stata avviata nel 2014 e vi hanno partecipato, sulla base di apposite convenzioni, la Regione Lombardia, la Regione Sicilia, il Comune di Vicenza e Unioncamere Lombardia. Si sono aggiunti, poi, su base volontaria INPS e Banca d'Italia.

Tra il 2016 e 2017 sono state organizzate due sessioni di formazione a cui hanno partecipato: il Ministero della Giustizia, il Ministero delle Infrastrutture e dei Trasporti, INPS, Banca d'Italia, Agenzia per la Coesione Territoriale, Presidenza del consiglio dei ministri (DIPE), Regione Sicilia, Regione Lombardia, Unioncamere Lombardia.

Inoltre, a seguito di incontri bilaterali, sono state organizzate apposite giornate informative a cui hanno partecipato: Ministero dell'Interno, Ministero dell'Economia e delle Finanze (Dip. Finanze), Ministero dell'Economia e delle Finanze (Dip. Tesoro), Ministero della Salute, Ministero Beni culturali, Ministero dell'Istruzione dell'Università e della Ricerca.

La richiesta della partecipazione al progetto avviene attraverso l'invio di una lettera, da parte di un Ente interessato, all'Istat. I dettagli sulle modalità di adesione e il modello di lettera sono accessibili sul sito del Sistan.

Gli enti che hanno finora aderito sono: Roma Capitale, Provincia Pesaro-Urbino, Regione Lombardia, Regione Sicilia, Unioncamere Lombardia, Comune di Roma, Comune di Palermo, Unioncamere Campania, Comune di Brescia.

All'indirizzo <http://sistanhub.istat.it> è possibile accedere sia al prototipo dell'*Hub* sia al *repository* dei metadati a supporto.

Nella XIII conferenza Nazionale di Statistica è stato presentato, in forma prototipale, l'accesso attraverso l'*Hub* a un nodo, a gestione Istat, con tabelle statistiche pubblicate da Enti centrali. L'idea è quella sia di facilitare la partecipazione degli Enti centrali stessi al progetto (un Ente centrale che decide di partecipare troverà il lavoro già avviato), sia gli Enti territoriali che, attraverso l'*Hub*, potranno accedere a tutte quelle statistiche di cui hanno bisogno per inserirle nei propri annuari.

## 5.2 *L'evoluzione del progetto*

Nel corso di questa prima fase di sperimentazione si è constatata una crescente difficoltà degli enti centrali ad aderire al progetto. Si intersecano infatti problematiche che coinvolgono tutta l'amministrazione, i rapporti tra i vari uffici e in particolare con i fornitori dei sistemi informativi. Questo nonostante si sia offerta la disponibilità di prodotti e moduli liberamente utilizzabili, standardizzati in ambito Eurostat e OCSE,

assistenza nella fase di impianto e diverse possibilità di caricamento dei dati (anche in cloud presso l'Istat).

Si sta pertanto mettendo a punto una evoluzione del progetto che, partendo da una revisione della scheda PSN che riguarda i prodotti che verranno diffusi, richiede agli enti di formalizzare sin dall'inizio l'indicazione delle tavole principali che verranno prodotte (o i sistemi informativi), seguendo le regole di qualità e metadazione previste per l'HUB (in particolare, le variabili di classificazione utilizzate, la struttura della tavola, la URL dove è possibile scaricare la tavola, le date degli aggiornamenti, ecc.) In sede di Stato d'attuazione (cioè in fase di dichiarazione di quanto effettivamente prodotto), le tavole dovranno essere fornite con le caratteristiche richieste.

Questo non solo permetterà di verificare a posteriore che quanto dichiarato a livello di scheda sia poi effettivamente effettuato, ma anche faciliterà l'alimentazione del "nodo Enti centrali" accessibile dall'*Hub* e, nel tempo, un consistente miglioramento della qualità complessiva dei rilasci.

## **6 Qualità dei dati amministrativi e dei registri statistici**

Come ricordato nella Relazione di fine mandato del Presidente Giorgio Alleva, un punto di attenzione del Programma di modernizzazione riguarda la capacità di misurazione e la lettura dei nuovi fenomeni d'interesse della statistica ufficiale mediante l'integrazione delle diverse fonti e un massivo ricorso ai dati amministrativi.

Ne deriva in maniera diretta una rinnovata attenzione alla qualità degli archivi.

### *6.1 Programmazione e acquisizione, conservazione e rilascio dei dati amministrativi*

Come già descritto nella relazione per l'Audizione dell'Istat presso la Commissione per la garanzia della qualità dell'informazione statistica del 22 settembre 2017, il processo di gestione dei dati amministrativi segue un iter procedurale che prevede ogni anno la raccolta dei fabbisogni informativi presso le Direzioni di produzione statistica e la predisposizione del Documento di programmazione delle acquisizioni di dati amministrativi.

Nel 2018 è programmata l'acquisizione di 477 forniture relative a 194 archivi di titolarità di 51 Enti. La dinamica della domanda e dell'offerta di dati amministrativi è stata notevole negli ultimi anni e ciò ha richiesto importanti innovazioni nel processo di gestione dei dati amministrativi.

Il controllo della qualità del processo di gestione dei dati viene presidiato attraverso la progressiva standardizzazione delle procedure. L'acquisizione da parte dell'Istat dei dati amministrativi avviene tramite un portale unico di raccolta informatica dei dati di titolarità di enti pubblici e privati denominato "Sistema di acquisizione degli archivi

amministrativi - Arcam” (<https://arcam.istat.it>). L’utilizzo di tale Portale rappresenta l’esito di un processo innovativo fortemente perseguito dall’Istituto, finalizzato a garantire una maggiore sicurezza nella trasmissione delle informazioni, l’integrità dei flussi di dati verso l’Istituto e l’utilizzo di tecnologie standard conformi ai requisiti prescritti dalla legislazione in materia di trattamento dei dati personali.

I dati amministrativi acquisiti dall’Istituto vengono utilizzati esclusivamente come input dei diversi processi di produzione di titolarità dell’Istat; in nessun caso vengono diffusi o comunicati all’esterno in forma elementare.

Un *repository* centralizzato all’interno della rete intranet dotato di tutte le misure di sicurezza e di protezione previste dall’Istituto garantisce la conservazione dei dati.

Alcuni archivi (circa un centinaio) vengono conservati nella forma originale in un *file system* organizzato a cui è possibile accedere esclusivamente attraverso il sistema unico di rilascio dei dati amministrativi<sup>6</sup> e dai soli utenti autorizzati nel rispetto della vigente normativa in materia di protezione dei dati personali<sup>7</sup>.

Altri archivi (circa 70) dopo opportune procedure di trasformazione e standardizzazione dei formati di input, vengono memorizzati in un database di tipo relazionale (il dbms ORACLE standard d’Istituto) denominato SIM – Sistema integrato di microdati amministrativi, già descritto nella relazione per l’Audizione del 22 settembre 2017.

L’accesso ai dati amministrativi è consentito ai soli utilizzatori interni e a seguito dell’espletamento di una procedura di autorizzazione che prevede la verifica da parte della struttura di gestione del processo di acquisizione e rilascio dei dati amministrativi (RDD) che l’archivio sia stato correttamente indicato nel progetto del Programma Statistico Nazionale in vigore per cui si richiede e che le persone a cui devono essere consegnati i dati siano stati indicati nella Delibera di incarico al trattamento dei dati personali che ogni struttura è tenuta a predisporre e a mantenere aggiornata.

## 6.2 Documentazione della qualità dei dati amministrativi

La documentazione della qualità dei dati amministrativi acquisiti da enti esterni, appartenenti o meno al Sistan, viene gestita dal portale QRCA (*Quality Report Card* dei dati Amministrativi) in fase di rilascio nella intranet. La QRCA è rivolta agli utenti interni Istat e ha i seguenti obiettivi:

---

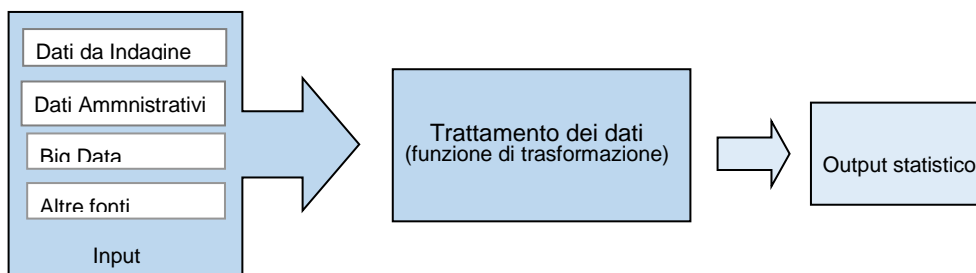
<sup>6</sup> <https://edi.istat.it>

<sup>7</sup> La normativa vigente, infatti, consente l’acquisizione dei dati amministrativi solo per le finalità statistiche indicate nel Programma Statistico Nazionale (PSN) in corso di validità o per le finalità statistiche connesse ad adempimenti derivanti da Regolamenti Europei, ovvero per altre finalità statistiche legate a progetti di ricerca di particolare rilevanza.

1. valutare la qualità delle forniture dei dati amministrativi per: a) supportare il processo di acquisizione; b) misurare l'usabilità statistica dei dati amministrativi acquisiti; c) monitorare le variazioni inattese della qualità/usabilità dei dati acquisiti;
2. controllare la coerenza del sistema di integrazione;
3. produrre il sistema di documentazione della qualità dei dati amministrativi e del processo di integrazione.

La definizione di qualità dei dati statistici si è evoluta negli ultimi anni per comprendere le diverse tipologie di produzione non più esclusivamente basate sui dati di indagine. Laddove il processo prevede l'uso di dati esterni (processi multifonte) si è inserito il concetto qualità dell'input, considerando che la valutazione è effettuata a posteriori non avendo l'utilizzatore la visione del processo di acquisizione e non potendo intervenire su di esso (si pensi, ad esempio, ai dati della modulistica reddituale acquisiti dall'Istat). Accanto al concetto di qualità dell'input del processo, si affianca la qualità del trattamento o *throughput* e, infine, la qualità dell'output.

Figura 1 Componenti della qualità del processo di produzione delle statistiche



Ovviamente la qualità dell'output dipende dalla qualità delle singole fasi e ne ingloba gli effetti. Occorre sottolineare che, per quanto riguarda i dati amministrativi, il concetto di qualità a scopo statistico non sempre si identifica con il concetto di qualità a scopo amministrativo: dati amministrativi che adempiono perfettamente agli scopi istituzionali o gestionali derivanti dalle normative o dalle procedure per le quali sono generati, possono avere una scarsa qualità statistica in termini di usabilità nei processi di produzione.

La valutazione della qualità statistica dei dati amministrativi rende possibile, quindi, definirne la loro usabilità come input nei processi di produzione (obiettivo 1b). L'altro aspetto importante che si vuole sottolineare è il monitoraggio della qualità, ovvero la verifica della continuità del grado di usabilità dei dati acquisiti (obiettivo 1c). Le variazioni di qualità possono derivare sovente da mutazioni della normativa che genera i dati amministrativi presso le istituzioni pubbliche preposte a svolgere specifici



adempimenti (fiscali, previdenziali, di controllo, ...). Variazioni della normativa possono produrre modifiche della modulistica e, di conseguenza, nella struttura dei dati amministrativi, sia in termini di variabili che in termini di unità (specifiche deduzioni fiscali che vengono introdotte e soppresse, modifiche della platea delle imprese che devono pagare una determinata imposta, etc.). Anche ristrutturazioni dei DB di gestione presso gli Enti possono determinare un impatto nel processo di produzione delle statistiche. Al fine di prevenire impatti negativi, sono previsti controlli di qualità in termini di serie storica per intercettare variazioni inattese. In occasione della comunicazione continua con gli enti titolari fornitori di dati è sempre attiva la funzione di informazione sugli scopi statistici dei dati acquisiti e tale canale di comunicazione permette, per le fonti amministrative più rilevanti, di collaborare nella gestione delle variazioni suddette e prevenire gli impatti. Nei limiti dei vincoli legati alle funzioni istituzionali, viene anche proposta la standardizzazione delle classificazioni e delle definizioni statistiche. La documentazione dei dati amministrativi nella QRCA supporta questo obiettivo, nelle modalità che si ritengono più idonee in relazione alla predisposizione alla collaborazione dell'Ente.

Per standardizzare questi concetti, nell'ambito del progetto internazionale BlueEts<sup>8</sup> a cui l'Istat ha partecipato, sulla base dell'esperienza di *Statistics Netherlands*, è stato definito un *Framework* con specifiche dimensioni della qualità che differiscono dalle dimensioni classiche della qualità dei dati statistici relative, invece, alla qualità dell'output (rilevanza, accuratezza, tempestività e puntualità, accessibilità e chiarezza, coerenza e comparabilità). Il *Framework* comprende un sistema gerarchico di iperdimensioni e dimensioni della qualità, indicatori e metodi di misura. Nello schema di seguito (Tabella 1) sono riportate le iperdimensioni, Fonte, Metadati e Dati, e le relative Dimensioni della qualità dei dati amministrativi utilizzati a fini statistici implementate nella QRCA. Ad esempio gli indicatori di Rilevanza, nell'Iperdimensione Fonte, misurano l'estensione di uso di una determinata fonte nell'ambito dei processi produttivi dell'Istat, nell'Iperdimensione dei Metadati sono riportati gli aspetti concettuali dei dati, in termini di unità e di variabili. La dimensione della Completezza, nell'Iperdimensione dei Dati, comprende sia gli indicatori della presenza di missing values nelle variabili amministrative che gli indicatori di copertura in termini di popolazioni statistiche di riferimento.

---

<sup>8</sup> Daas et al. (2009) Checklist for the Quality evaluation of AD Sources. Discussion paper 09042, Statistics Netherlands. Daas et al. (2011) Reports on methods preferred for the quality indicators of administrative data sources, Deliverable 4.2 of Workpackage 4 of the BLUE-ETS project. CBS, Netherlands, SSB, Norway, Istat, Italy, SCB, Sweden.

Tabella 1 - Framework della qualità dei dati amministrativi utilizzati a scopo statistico

IPERDIMENSIONE	DIMENSIONE
FONTE	Informazioni sull'archivio
Informazioni necessarie a gestire il processo di acquisizione dei dati con lo scopo di valutare e migliorare la qualità dei dati acquisiti	Rilevanza della fonte (utilizzi statistici)
	Questioni connesse alle norme di uso (accessibilità)
	Relazioni e feedback con i titolari delle fonti
METADATI	Chiarezza e interpretabilità dei dati
Informazioni per la valutazione della qualità a livello concettuale e di processo presso il titolare della fonte	Comparabilità tra concetti amministrativi e concetti statistici (unità e variabili)
	Descrizione del processo di acquisizione e trattamento (da parte del titolare)
DATI	Forniture acquisite e accordi
Valutazione della qualità dei dati acquisiti	Technical checks
	Integrabilità/Integrazione
	Accuratezza e coerenza interna
	Completezza
	Aspetti temporali

Rispetto alla formulazione originaria, definita nell'ambito del progetto internazionale BlueEts, in Istat è stata modificata la dizione della Dimensione dell'Integrabilità che prevede indicatori della qualità delle variabili di *linkage*, aggiungendo anche la parte di Integrazione con lo scopo di riportare informazioni sul processo di integrazione operato nell'ambito della DCRD. In tal modo è possibile documentare il processo di gestione dei dati amministrativi, che prevede anche la fase di integrazione dei dati nel SIM. Altre modifiche di minore entità hanno adattato lo schema alla realtà dell'Istat<sup>9</sup>.

<sup>9</sup> Per ulteriori approfondimenti, si veda Cerroni, F., Di Bella, G., & Galiè, L. (2014). Evaluating administrative data quality as input of the statistical production process, Rivista di statistica ufficiale 1-2/2014, 117-146 [https://www.istat.it/it/files/2014/10/Articolo-7\\_Evaluating-administrative....pdf](https://www.istat.it/it/files/2014/10/Articolo-7_Evaluating-administrative....pdf).

Ulteriore documentazione è anche disponibile (VENTURI M., DI BELLA G. (2018) The Istat platform prototype for the documentation of administrative data used as input of the official statistics, Paper per Q2018 - European Conference on Quality in Official Statistics, Kraków, Poland. 26-29 June 2018. DI BELLA G. (2018) L'interoperabilità conviene: documentare la qualità dei dati amministrativi utilizzati a scopi statistici in Istat, Presentazione della Relazione alla 13° Conferenza Nazionale di Statistica, Roma, 4 -6 luglio 2018. VENTURI M., DI BELLA G. (2017) "Supporting the use of administrative data in official statistics. The System of Integrated Microdata – SIM and the Quality Report Card of Administrative data - QRCA", Presentazione alla NTTS 2017 Conference, Bruxelles, 13 – 17 March 2017. VENTURI M., DI BELLA G. (2016) "New data collection organization in the Italian Statistical Institute and the IT tools to collect and manage administrative data", Presentazione al Workshop on access to administrative sources nell'ambito del progetto internazionale ADMIN ESS.VIP, Bruxelles 13-14 settembre 2016. DI BELLA G. (2017) Dati amministrativi per finalità statistiche: usi e accessibilità interna, Indicatori di qualità sui dati di fonte amministrativa. Corso Istat sulla qualità "Introduzione alla qualità nella Statistica Ufficiale". Roma, 23 - 24 Ottobre 2017).

Il progetto di produzione della QRCA attraversa ora la fase di implementazione. La situazione è favorevole in occasione del processo di modernizzazione che rafforza l'importanza dei dati amministrativi per la produzione delle statistiche.

Occorre premettere che la sfida per la progettazione del sistema ha dovuto tenere in considerazione il fatto che:

- ✓ gli indicatori di qualità devono essere prodotti per tutte le fonti amministrative e le oltre 450 forniture periodiche acquisite ogni anno dall'Istat;
- ✓ le caratteristiche dei dataset amministrativi hanno un'elevata variabilità;
- ✓ i dataset amministrativi sono spesso molto grandi in termini di bytes;
- ✓ in presenza di dati personali, occorre operare nel rispetto della normativa in termini di riservatezza e trattamento dei dati.

La difficoltà iniziale di costruire e, soprattutto, mantenere aggiornato il sistema con alti livelli di performance è stata superata adottando l'innovativa strategia che prevede il riuso dei metadati dei sistemi che gestiscono i dati amministrativi acquisiti in Istat per la produzione degli indicatori di qualità. In particolare, le metriche vengono calcolate in modo automatico utilizzando i metadati presenti in Arcam (il sistema di acquisizione e repository dei dati amministrativi) e in SIM (il Sistema integrato di Microdati amministrativi). A tale scopo i due sistemi, nati in due fasi distinte, sono stati resi interoperabili al fine di poter "seguire" il ciclo di vita delle forniture. Questa strategia garantisce l'aggiornamento automatico della QRCA in occasione delle nuove acquisizioni o dei nuovi trattamenti.

Nel prototipo che si pubblicherà entro l'anno, sono implementate le prime metriche per le quali, già da ora, sono disponibili le informazioni per il loro calcolo e aggiornamento automatico. In particolare, una volta effettuata l'autenticazione nel sistema con le proprie credenziali uniche di accesso alla intranet (single sign on), è possibile visualizzare la lista degli archivi amministrativi acquisiti dall'Istat, selezionare l'archivio di interesse e accedere alle seguenti informazioni:

- ✓ Identificazione della Fonte, dell'Ente titolare
- ✓ Rilevanza della fonte in termini di estensione di uso in Istat
- ✓ Elenco dei lavori PSN che utilizzano la fonte
- ✓ Regolamenti comunitari il cui adempimento dipende dalla fonte
- ✓ Identificazione delle forniture, periodicità, riferimenti temporali
- ✓ Puntualità dell'ente fornitore
- ✓ Tempestività dei dati riferita all'Ente (intervallo temporale tra la data dell'ultimo evento registrato nella fornitura e la data di acquisizione)
- ✓ Tempestività dei dati complessiva (intervallo temporale tra la data dell'ultimo evento registrato nella fornitura e la data di disponibilità)
- ✓ Lista delle variabili amministrative
- ✓ Classificazioni amministrative per le variabili categoriche

- ✓ Tipi di unità presenti (Individui, Unità economiche, Luoghi)
- ✓ *Technical checks* per la verifica della conformità dei dati ricevuti rispetto a quelli richiesti e per il monitoraggio della qualità: confronti in serie storica delle misure.
- ✓ Percentuale di valori mancanti per le variabili (anche in serie storica)
- ✓ Controlli di completezza delle decodifiche delle classificazioni
- ✓ Misure della qualità delle variabili di *linkage* (variabili identificative disponibili nella fonte e percentuale di valori mancanti)
- ✓ Misure del monitoraggio della qualità delle procedure di record *linkage* (misure deterministiche)

Nelle sue release successive, è previsto che il sistema si arricchisca con le metriche che via via si renderanno disponibili una volta messe a punto le interconnessioni con altri sistemi che possono alimentare la QRCA. Il prossimo passo consiste nella definizione di una connessione tra il sistema Arcam di acquisizione dei dati amministrativi con il DB del PSN. Poiché l'accesso ai dati amministrativi è vincolato alla dichiarazione di uso all'interno di ciascun lavoro PSN dell'Istat, risulta estremamente utile disporre di un collegamento tra la lista degli archivi amministrativi dichiarati nel PSN e la lista degli archivi amministrativi acquisiti in Istat, l'obiettivo operativo è di condividere in un unico spazio le informazioni comuni, permettendo la gestione automatica delle variazioni: nuove fonti acquisite dal lato Arcam che incrementano la lista, nuovi archivi amministrativi utilizzati nei processi e dichiarati nel PSN che si abbinano alla lista o la incrementano. Attualmente il processo prevede una fase di controllo manuale della conformità del processo di acquisizione. È stata recentemente avviata questa nuova attività in collaborazione con la Direzione Istat di riferimento (Direzione centrale per la pianificazione strategica, l'indirizzo del Sistema statistico nazionale, le relazioni istituzionali e gli affari internazionali) e le prime innovazioni saranno implementate nel prossimo aggiornamento del PSN con l'aggiuntiva di una fase di validazione dei nuovi inserimenti rispetto alla lista Arcam che verrà gestita con il supporto della QRCA.

Una considerazione conclusiva riguarda la funzione centralizzata di trattamento dei dati amministrativi gestita nel "Servizio Fonti amministrative e integrazione dei registri" che determina l'output verso i servizi di produzione. Come detto, una parte degli archivi (circa il 60%) viene acquisito e rilasciato ai settori di produzione in forma grezza, il restante 40% (costituito dagli archivi più rilevanti) viene sottoposto ad un'analisi E/R, e le forniture vengono progressivamente caricate e integrate nel sistema SIM. I controlli di qualità di questo output riguardano la verifica della conformità dei dati ricevuti rispetto alla richiesta ed il monitoraggio della qualità in serie storica in termini macro (numero di record, numero di missing, andamenti delle distribuzioni di frequenza). Non vengono applicati piani di check sulle singole variabili o su insiemi di variabili e, tantomeno operazioni di correzione. Tale attività sarebbe, invero, troppo onerosa e

dovrebbe accogliere i fabbisogni dei vari utenti in relazioni agli specifici e differenti target di produzione (si pensi alle diverse esigenze delle statistiche strutturali, congiunturali o di contabilità nazionale). L'intenzione futura è di connettere la QRCA con il Sistema di documentazione delle indagini, SIDI, al fine di associare agli archivi documentati nella QRCA ulteriori indicatori, come gli indicatori di copertura o gli indicatori che misurano la distanza tra concetti amministrativi e concetti statistici come funzione di trasformazione del dato, rispetto agli specifici target dei processi statistici di SIDI. A questo scopo occorre che i due sistemi effettuino dei passi di convergenza sulle questioni comuni.

### 6.3 *Qualità dei registri statistici costruiti di fonti amministrative*

Con la costituzione di un sistema integrato di registri (SIR) per la produzione statistica l'Istat si sta attrezzando per la definizione di standard di documentazione, indicatori per il monitoraggio e misure per la valutazione della qualità dei registri stessi, dal momento che questi sono alimentati estensivamente da dati di origine amministrativa oltre che di indagine.

Già da diverso tempo l'Istat ha avviato varie attività relativamente alla valutazione della qualità della produzione di statistiche basata su dati di fonte amministrativa.

I principali registri del SIR e i processi che utilizzano dati amministrativi sono correntemente documentati nel sistema informativo della qualità dell'Istat (SIQual), e alcuni indicatori di base, in particolare quelli comuni alle altre tipologie di processo, sono calcolati e archiviati nel sistema (per es. indicatori sul controllo e correzione, indicatori di tempestività, coerenza e comparabilità).

Nel 2016, sono state prodotte e pubblicate le Linee Guida per la qualità dei processi statistici che utilizzano dati amministrativi<sup>10</sup>, e sono stati condotti due audit sperimentali su due diverse tipologie di processo.

Partendo dall'esperienza maturata nell'ambito del registro delle statistiche strutturali sulle imprese (denominato FRAME-SBS), l'Istat ha già definito un quadro metodologico di riferimento per la misurazione della qualità dei dati nel caso di registri ottenuti integrando diverse fonti amministrative, e ha identificato un set di indicatori per il monitoraggio di alcune fasi tipiche del processo produttivo alla loro base. Tali indicatori comprendono misure di qualità dell'input simili a quelle sviluppate per gli archivi amministrativi acquisiti e descritti nel precedente paragrafo, ma si differenziano dai precedenti in quanto, nel caso dei registri statistici, la valutazione è orientata al monitoraggio della qualità di dati e processi finalizzati alla produzione di *specifici* output statistici. Inoltre, gli indicatori definiti prevedono una serie di misure per la valutazione

---

<sup>10</sup> <https://www4.istat.it/it/files/2010/09/Linee-Guida-fonte-amministrativa-v1.1.pdf>

e il monitoraggio della qualità dei processi di trasformazione del dato amministrativo in informazione statistica, di integrazione delle fonti, di trattamento delle eventuali incoerenze nei dati integrati, di stima degli aggregati di interesse.

Importanti studi sono stati avviati, in particolare, per la valutazione della qualità dei dati e delle stime basate sui registri in funzione del prossimo censimento permanente della popolazione italiana, con particolare riferimento alle statistiche sull'occupazione, sulle caratteristiche demografiche degli individui della popolazione, e sui luoghi.

Sotto l'impulso e la supervisione del Comitato per la Gestione del Sistema dei Registri (CGR), inoltre, ha preso l'avvio una riflessione strutturata sul problema della valutazione della qualità complessiva del SIR con riferimento sia alla dimensione della tempestività, sia a quella della coerenza dell'intero sistema di statistiche prodotte integrando dati derivati dalle fonti amministrative (attraverso i registri) e dalle rilevazioni statistiche dirette.

Il principale impegno futuro è di sfruttare l'esperienza fin qui accumulata per migliorare la documentazione, sistematizzare e generalizzare gli strumenti sviluppati e inquadrarli in una strategia complessiva che tenga conto delle informazioni sulla qualità attinenti agli archivi amministrativi acquisiti attraverso l'interoperabilità tra sistemi di gestione della qualità.

## **7 Il miglioramento della trasparenza all'Istat**

Alcune delle raccomandazioni dell'ultima Peer Review di Eurostat nel 2015 hanno interessato il tema della trasparenza, aumentando l'offerta informativa sulla qualità dei processi e dei prodotti statistici verso gli utenti esterni e migliorandone l'accessibilità.

In particolare la raccomandazione 19 richiedeva di stabilire *“una politica delle revisioni con norme seguenti e procedure trasparenti e pubblicarlo sul proprio sito web, compreso un calendario di revisione per informare gli utenti di quando sono previste pubblicazioni di revisioni di risultati statistici.”* mentre la raccomandazione 20 richiedeva di *“espandere i propri sistemi per riportare indicatori di qualità per gli utenti in maniera regolare”*.

### *7.1 Avanzamenti delle politiche di revisione*

Rispetto a quanto comunicato nell'audizione precedente (marzo 2017) le pratiche dell'Istat in risposta alla raccomandazione n. 19 scaturita dall'ultima peer review da parte di Eurostat, relative alla politica e all'analisi delle revisioni degli indicatori congiunturali al fine di garantire una migliore aderenza ai corrispondenti principi del codice delle statistiche europee, sono state espletate con la pubblicazione all'interno

del sito “congiuntura” dell’Istat<sup>11</sup>, di un’apposita sezione, disponibile dal giugno 2017, che comprende:

- ✓ Una pagina introduttiva che oltre a presentare i concetti e le definizioni principali, descrive la politica di revisione adottata dall’Istituto ed elenca tutte le informazioni sulle revisioni disponibili sul sito;
- ✓ L’elenco di tutti gli indicatori congiunturali soggetti a revisione, raggruppati per processo statistico che li produce. Per ognuno di questi processi viene fornita una scheda standardizzata che descrive la politica di revisione adottata nello specifico, elenca i motivi delle revisioni (sia ordinarie che straordinarie), e fornisce il calendario del ciclo completo delle revisioni ordinarie. Quindi, per ogni processo, è presente l’elenco degli indicatori statistici, per ciascuno dei quali è disponibile:
  - ✓ Un file excel con i triangoli delle serie dei vintage (tale organizzazione dei dati è definita “real time data set”);
  - ✓ Una scheda in formato excel con gli indicatori di qualità sulle revisioni, secondo lo standard dell’OECD e includendo gli indicatori standard Eurostat, quali Mean Revision (RM), Mean Average Revision (MAR), Relative Mean Average Revision (RMAR);
  - ✓ Una scheda di revisione riepilogativa che riporta un grafico dinamico con la rappresentazione delle revisioni relative alla differenza tra la seconda stima e la prima stima pubblicata e una tabella riassuntiva degli indicatori.

È importante sottolineare che le funzioni sviluppate hanno un elevato grado di automazione, permettendo l’aggiornamento delle informazioni in contemporanea con altri sistemi di diffusione dei dati. I linkage degli indicatori sono infatti accessibili anche attraverso il datawarehouse di Istituto, I.Stat, mentre i metadati che descrivono la politica di qualità sono accessibili anche da SIQual, Sistema Informativo Istat per la qualità dei processi statistici<sup>12</sup>.

## 7.2 I Quality report e le Schede standard di qualità

In risposta alla raccomandazione 20, in sede di Peer Review l’Istat si è impegnato prima a diffondere a livello nazionale i quality report predisposti per Eurostat e già diffusi da Eurostat stesso e successivamente a incrementare gli indicatori di qualità disponibili per gli utenti, incrementare la documentazione sulla qualità e migliorare l’accessibilità delle informazioni sulla qualità sul sito istituzionale.

I quality report compilati per Eurostat, e quindi in lingua Inglese, sono diffusi a partire da dicembre 2016 attraverso il sistema SIQual<sup>13</sup> organizzati per tematica e associati ai

---

<sup>11</sup> <https://www4.istat.it/it/congiuntura/revisioni>

<sup>12</sup> <http://siqual.istat.it>

<sup>13</sup> <http://siqual.istat.it/SIQual/docQualityReport.do?ric=0>

processi statistici che generano le statistiche cui i quality report si riferiscono. Mentre sul sito di Eurostat è sempre disponibile solo il report più recente per ogni statistica, in SIQual sono disponibili anche le edizioni precedenti.

Le richieste di maggiore trasparenza, incremento delle informazioni e degli indicatori di qualità a disposizione degli utenti esterni e miglioramento dell'accessibilità di tali informazioni sono pervenute all'Istat sia in forma di raccomandazione della Peer Review sia dalla precedente audizione della COGIS.

In risposta a tali richieste l'Istat ha sviluppato le Schede standard di qualità, pubblicando sul sito web un primo insieme di 75 schede, in italiano, per altrettanti processi statistici.

Tali schede seguono un template standardizzato e includono, oltre agli opportuni metadati, un set di indicatori di qualità conformi agli standard Europei e già disponibili nella versione interna di SIQual ma mai diffusi in maniera sistematica dall'Istat, come il tasso di mancata risposta totale, o il tasso di imputazione.

Il template usato nelle schede, riportato in figura 1, è basato sullo standard Europeo SIMS<sup>14</sup>.

Esso include informazioni descrittive di base sui dati (metadati concettuali) e sul processo di produzione (metadati metodologici) delle statistiche, riporta la politica dell'Istituto in vari ambiti, dalla riservatezza alla diffusione alla qualità, e poi approfondisce le dimensioni della qualità dell'output, come definite dalla legge statistica europea, riportando per l'accuratezza anche le azioni di controllo della qualità adottate nel processo statistico per ridurre gli errori non campionari e, come già menzionato, i valori assunti per gli indicatori di qualità. La descrizione e la formula di calcolo degli indicatori sono illustrati in un apposito approfondimento metodologico<sup>15</sup>. Tramite la sezione sull' "Accessibilità e chiarezza" è anche possibile raggiungere direttamente i comunicati o le pubblicazioni che riportano le statistiche a cui la scheda si riferisce.

Tabella 2 - Template Standard

Nome del processo	
Anno di riferimento	
Contatti	Accuratezza e attendibilità
Istituto	Campionamento
Struttura responsabile	Errori non campionari
Responsabile del processo	Errore di copertura
Informazioni sui dati	Tasso di sovra-copertura
Descrizione	Errore di mancata risposta
Unità di analisi	Tasso di mancata risposta totale
Riferimento territoriale	Errore di misurazione

<sup>14</sup> <http://ec.europa.eu/eurostat/web/quality/quality-reporting>

<sup>15</sup> <https://www.istat.it/it/files/2018/06/approfondimento-metodologico.pdf>



Periodicità delle stime	Tasso di imputazione
Periodo di disponibilità delle stime	Errori di trattamento
Anno base	Politica dell'Istituto sulle revisioni
Descrizione del processo	Politica del processo sulle revisioni
Fonte dei dati	Indicatori di qualità sulle revisioni
Periodicità dell'acquisizione dei dati	Tempestività e puntualità
Tecnica di acquisizione dei dati	Tempestività
Trattamento	Puntualità
Validazione degli aggregati	Coerenza e comparabilità
Mandato istituzionale	Lunghezza della serie storica comparabile
Tutela della riservatezza	Coerenza
Politica dell'Istituto sulla riservatezza	Accessibilità e Chiarezza
Trattamento dei dati per la tutela la riservatezza	Comunicati stampa
Diffusione	Pubblicazioni
Politica dell'Istituto sulla diffusione	Database online
Calendario delle diffusioni	Accesso ai microdati
Periodicità del rilascio delle stime	Documentazione metodologica
Gestione della qualità	Documentazione sulla qualità
Politica dell'Istituto sulla qualità	
Valutazione della qualità del processo	
Pertinenza	
Esigenze degli utenti	
Soddisfazione degli utenti	

Non tutti i campi sono applicabili a tutti i processi, quelli non applicabili non sono visualizzati in diffusione.

Le schede sono state compilate automaticamente a partire dalla documentazione presente in SIQual, garantendo quindi la coerenza delle informazioni, capitalizzando quanto già disponibile internamente.

Oltre alla definizione dei contenuti delle schede, particolare cura è stata riservata a disegnare come queste dovessero essere diffuse nel sito istituzionale, al fine di migliorare anche l'accessibilità delle informazioni sulla qualità, così come richiesto dalla raccomandazione della Peer Review. È stata quindi sviluppata una sezione dedicata<sup>16</sup> sul sito web Istat e sono stati introdotti diversi punti di accesso alle schede stesse, ad esempio da ogni sezione tematica del sito tramite il tab "Classificazioni e metadati".<sup>17</sup>

## 8 Miglioramenti nelle politiche di accesso ai microdati

La crescente disponibilità di basi di microdati risultanti da indagini e fonti amministrative rappresenta un patrimonio informativo per la cui gestione è necessario attivare politiche che ne agevolino l'accesso e adottare strumenti in grado di tutelare la *privacy* dei

<sup>16</sup> <https://www.istat.it/it/metodi-e-strumenti/strumenti-per-la-qualit%C3%A0/schede-standard-di-qualit%C3%A0>

<sup>17</sup> Per un esempio si faccia riferimento alla scheda dell'Indagine trimestrale su posti vacanti e ore lavorate <https://www.istat.it/it/archivio/217230>

soggetti interessati. Al perseguimento di questi obiettivi concorrono sia un'appropriatezza disciplinaria giuridica sia efficaci soluzioni tecnologiche.

### *8.1 Conservazione e accessibilità dei microdati*

L'Istat, cogliendo le indicazioni provenienti dall'Unione europea, si è impegnato assiduamente, nel rispetto delle norme vigenti<sup>18</sup>, ad ampliare quanto più possibile l'accesso ai microdati d'indagine di cui è titolare. Fra il 2002 e l'Ottobre 2018 (più precisamente l'8 ottobre 2018) i file di microdati archiviati nell'Archivio dei microdati validati (ARMIDA) sono 14.681.

Come già ricordato nelle precedenti audizioni, accanto alla tradizionale offerta di File standard, già da alcuni anni l'Istat diffonde i File per la ricerca (MFR)<sup>19</sup>, realizzati specificamente per soddisfare esigenze conoscitive da parte dei ricercatori di università o enti di ricerca e, dal 2013, ha avviato la produzione di File a uso pubblico (PUF), collezioni di dati elementari scaricabili liberamente e gratuitamente dal sito istituzionale, prima attraverso i cosiddetti "file micro.STAT" poi con i "file a uso pubblico storici", grazie ai quali anche un pubblico "generalista" può accedere ai microdati delle rilevazioni di maggiore interesse per la collettività (Forze di lavoro, Uso dell'ICT nelle imprese, Aspetti della vita quotidiana, Spese delle famiglie ecc.).

Alla produzione di questi file si associa inoltre la possibilità per ricercatori e studiosi di condurre in autonomia analisi statistiche sui microdati prodotti dall'Istat, grazie all'accesso al Laboratorio ADELE. L'output finale, risultato delle elaborazioni condotte dai ricercatori, viene attentamente valutato, sotto il profilo della riservatezza statistica, dagli esperti del Laboratorio e rilasciato a condizione che soddisfi delle precise "Regole per il rilascio dei risultati".

Nel 2017 le richieste di file per la ricerca sono state 254, quasi il doppio rispetto all'anno precedente (131), mentre sono 42 gli accessi al Laboratorio ADELE (in aumento di circa il 14 per cento rispetto al 2016). Scende invece il numero delle richieste sia dei file standard (che passano da 306 a 187) in quanto non più prodotti dall'Istituto e progressivamente sostituiti dai file MFR, sia dei file micro.STAT, che calano in un anno del 19 per cento ma segnano una ripresa più marcata nei primi mesi del 2018.

---

<sup>18</sup> Art. 10, comma 2, del d.lgs. 322/1989 e Art. 7, Allegato A3 del 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, relativo a "Comunicazione a soggetti non facenti parte del Sistema statistico nazionale".

<sup>19</sup> L'Istat li diffonde dal 2010 in conformità all'art. 6 del Regolamento (CE) n. 831/2002 che ha introdotto l'accesso a "microdati anonimi" ottenuti da alcune indagini, poi abrogato dal successivo Regolamento (UE) n. 557/2013, che ha definito l'accesso ai dati riservati destinati a fini scientifici.

## 8.2 *L'informazione sui microdati*

Il recente sviluppo e lancio del nuovo sito web ha favorito l'accesso alle informazioni sui microdati. La scelta di inserire direttamente in *home page* una sezione sui microdati agevola gli utenti nel conoscere e reperire le informazioni disponibili: la funzione di ricerca avanzata – basata sia sul tema di riferimento sia sulla tipologia di documento – consente di navigare tra i diversi tipi di file (file standard, file MFR, file a uso pubblico) per individuare le soluzioni di maggiore interesse.

Le innovazioni più importanti, scaturite da un proficuo rapporto con gli utilizzatori del sito, hanno riguardato in particolare il Laboratorio ADELE relativamente alla possibilità che viene offerta ai ricercatori di consultare liberamente e in dettaglio, attraverso le pagine web, tutti i metadati relativi a ciascuna delle indagini dell'Istituto. In questo modo, gli studiosi – ancor prima di chiedere l'accesso al Laboratorio per condurre analisi sui file disponibili – possono conoscere le informazioni riguardanti non solo le indagini cui sono interessati (una sorta di catalogo online) ma anche le tecniche di rilevazione, le variabili, le classificazioni, i tracciati record.

Un'apposita sezione del sito guida inoltre i ricercatori nella predisposizione online della proposta di progetto di ricerca e nella compilazione della domanda di accesso ai file di microdati, innovazione importante che ha consentito di ridurre drasticamente i tempi di risposta agli utenti.

## 8.3 *Il Laboratorio ADELE e l'accesso da remoto*

Il mondo della ricerca scientifica manifesta da tempo l'esigenza di condurre analisi più dettagliate sul settore delle imprese, al fine di condurre studi longitudinali e disporre di basi di dati che integrino informazioni da più fonti.

L'Istat già dal 2011 ha reso disponibili presso il Laboratorio ADELE i primi file integrati di microdati di impresa, modalità di diffusione che si è andata progressivamente arricchendo nel tempo, per quanto ancora insufficiente a soddisfare le esigenze conoscitive della comunità scientifica.

Dal 2016 a oggi sono stati messi a disposizione del Laboratorio ADELE sette nuovi file risultanti dall'integrazione di dati provenienti da più rilevazioni, mentre cinque sono stati aggiornamenti di edizioni precedenti, a conferma di come l'investimento fatto in questa direzione stia progressivamente entrando a regime.

L'ampliarsi dell'interesse verso questo tipo di analisi ha spinto l'Istituto a cercare nuove soluzioni, che consentano ai ricercatori di effettuare direttamente link tra uno o più file.

Ad esempio, per i microdati sulle imprese, il progetto, attualmente in fase di sviluppo, prevede che il codice statistico che Istat attribuisce a ciascuna impresa (abituamente denominato codice ASIA) sia sostituito da un codice criptato generato al momento

dell'accesso al Laboratorio da parte dell'utente autorizzato. Il codice criptato è tale per cui non sarà possibile risalire da questo al codice ASIA<sup>20</sup> e viceversa: l'eventuale conoscenza pregressa di uno o più codici ASIA da parte del ricercatore che accede ad ADELE non offrirebbe quindi alcuna informazione aggiuntiva, potenzialmente utile a fini di intrusione.

#### *8.4 Le Linee guida accesso ai microdati*

In conformità a quanto previsto all'art. 5-ter del decreto legislativo 14 marzo 2013, n. 33, sono state predisposte e adottate dal Comstat le "Linee guida per l'accesso a fini scientifici ai dati elementari del Sistema statistico nazionale (Sistan)", attraverso la predisposizione della direttiva n. 3/2018, indirizzata agli uffici di statistica di cui all'art. 3 del decreto legislativo n. 322 del 1989. Tale direttiva, in conformità alle disposizioni vigenti, per gli altri uffici di statistica di cui agli articoli 4 e 5 del medesimo decreto assume il valore di atto di indirizzo.

Le linee guida – approvate dal Garante per la protezione dei dati personali con atto n. 388 del 21 giugno 2018 e attualmente in attesa dell'assenso da parte del Ministro vigilante, ai sensi dell'art. 3 del DPR n. 166 del 2010 saranno definitivamente emanate dal Comstat nella prima riunione utile – stabiliscono le condizioni in base alle quali le istituzioni di ricerca, nonché gli Enti pubblici o privati in possesso dei requisiti richiesti, possono accedere ai dati elementari di titolarità degli Enti Sistan, purché privi di ogni riferimento che permetta l'identificazione diretta delle unità statistiche. In particolare, il Comstat, in collaborazione con l'Istat, ha definito dettagliatamente i criteri e i requisiti soggettivi e oggettivi che permettono ai ricercatori di accedere ai dati elementari prodotti dagli enti e uffici del Sistan, uniformando la disciplina nazionale a quanto già regolamentato dalla normativa europea in materia di accesso ai dati riservati destinati a fini scientifici (Reg. (UE) n. 557/2013 della Commissione).

Inoltre, è stata introdotta la possibilità per i ricercatori interessati di accedere ai microdati anche da remoto, ossia presso altri laboratori accreditati diversi dal laboratorio ADELE che si trova nella sede centrale dell'Istat e da quelli dislocati negli uffici delle sedi territoriali Istat. E questo sempre nel pieno rispetto delle disposizioni che tutelano la riservatezza dei dati. Altra novità è la possibilità per gli Enti pubblici e privati di ottenere il riconoscimento di "ente di ricerca", condizione necessaria per l'accesso ai microdati, laddove tra gli scopi istituzionali dell'ente vi sia riferimento ad attività di ricerca e all'interno dello stesso vi siano strutture che svolgano tali attività. Ciò permette

---

<sup>20</sup> A rigore anche il codice ASIA non è un identificativo diretto, ma per le sue caratteristiche non offre sufficienti garanzie contro il rischio di intrusione. Il codice ASIA viene infatti utilizzato regolarmente per le comunicazioni tra Istat e imprese. Considerato che in nessun momento le imprese sono state informate della necessità di non renderlo noto e considerata anche la sua persistenza nel tempo, il codice ASIA risulta vulnerabile sotto il profilo della riservatezza, pur non essendo un identificativo diretto.

un ampliamento sia dei destinatari dei microdati prodotti dal Sistan, sia dell'offerta di informazione statistica grazie all'introduzione di un ulteriore canale di diffusione che favorisce una maggiore accessibilità ai microdati d'indagine, rispondendo così alla sempre crescente richiesta di dati elementari da parte della comunità scientifica.

L'applicazione di queste modalità è subordinata all'entrata in vigore delle nuove Linee guida per l'accesso a fini scientifici ai dati elementari del Sistan. È esclusa la conservazione sine die e la condivisione con persone non esplicitamente individuate nel modulo di accesso.

In particolare, per quel che riguarda l'accesso remoto sicuro via Internet le Linee guida prevedono che ciascun ente Sistan, che intende mettere a disposizione questa modalità di fruizione, possa condurre una sperimentazione per valutare l'infrastruttura tecnologica e le misure per la sicurezza dei dati più adeguate, nonché l'onere organizzativo e finanziario che l'accesso da remoto comporta, attraverso l'individuazione e la collaborazione con uno o più soggetti scelti in base alle loro caratteristiche scientifiche, tecnologiche e logistiche, l'utenza potenziale del laboratorio di cui dispone e le misure messe in atto per la sicurezza dei dati. L'Istat ha da individuato come possibili partner per la sperimentazione il Dipartimento Metodi e Modelli per l'Economia, il Territorio e la Finanza (MEMOTEF) della Sapienza Università di Roma e il Dipartimento Economia e statistica di Banca d'Italia.

La sperimentazione dovrà riguardare tanto gli aspetti informatici relativi alla sicurezza dei dati (configurazione della rete e dei pc, sistemi di autenticazione ecc.), quanto quelli organizzativi e procedurali (presidio del *safe centre*, gestione delle richieste, valutazione e rilascio degli output ecc.).

Si tratta comunque di un percorso molto complesso. Il Decreto legislativo, se da un lato consente indiscutibili miglioramenti sul fronte della fruibilità dei dati, dall'altro introduce talune complessità gestionali e organizzative già presenti nell'ordinamento europeo, introdotte allo scopo di bilanciare le accresciute possibilità di utilizzo dei dati con più stringenti garanzie.

Le nuove norme prevedono infatti una fase di preliminare riconoscimento delle istituzioni di appartenenza dei ricercatori, sia per l'accesso al Laboratorio, sia per le richieste di file MFR, sia per la stipula dei protocolli di ricerca. Ai fini di tale riconoscimento è richiesto che le istituzioni, oltre ad attestare, come avviene già oggi, la propria attività di ricerca, sottoscrivano, attraverso un soggetto abilitato a rappresentarle all'esterno, un impegno di riservatezza esteso a tutti i ricercatori che hanno accesso ai dati elementari, nel quale siano specificate le condizioni di utilizzo dei dati, gli obblighi dei ricercatori e i provvedimenti previsti in caso di violazione degli impegni assunti, nonché le misure adottate per tutelare la riservatezza dei dati.

Nel caso dell'Istat, per dare completa attuazione al Dlgs sarà quindi necessaria una revisione complessiva del percorso autorizzatorio per l'accesso ai dati e – di conseguenza – una revisione della modulistica, tanto per la richiesta dei file MFR quanto per l'accesso al Laboratorio ADELE. Ancor più gravoso può essere l'iter per gli Uffici di statistica del Sistan che intendono rendere accessibili i propri microdati.

Si segnala, inoltre, che le Linee guida, nell'ambito della sezione dedicata ai file MFR, prevedono la possibilità per la Commissione per la garanzia dell'informazione statistica di richiedere agli enti del Sistan produttori di questo tipo di file la documentazione relativa alle misure di protezione adottate, misure che – sulla base dei protocolli consolidati in materia di riservatezza – non vengono rese pubbliche, al fine di non indebolire la protezione statistica dei dati.