

**Indagine conoscitiva su
“Misure di contrasto all’evasione fiscale, sicurezza delle banche
dati dell’anagrafe tributaria e tutela della riservatezza dei dati
dei contribuenti”**

Audizione dell’Istituto nazionale di statistica

Dott. Massimo Fedeli

**Direttore del Dipartimento per lo sviluppo di metodi e tecnologie per la
produzione e diffusione dell’informazione statistica**

Dott. Stefano Menghinello

**Direttore del Dipartimento per le statistiche economiche, ambientali e
conti nazionali**

Indice

Introduzione	4
1. L'economia non osservata nei Conti nazionali: aspetti metodologici, principali evidenze e il contributo alla stima dell'evasione fiscale e contributiva	4
2. La strategia dell'Istat per la produzione statistica multi-fonte: <i>data governance</i> e protezione dei dati	11
Allegato statistico	16

Introduzione

Questa memoria è divisa in due sezioni.

La prima descrive l'impianto concettuale e metodologico alla base della stima dell'economia non osservata all'interno del Sistema dei Conti nazionali; fornisce una breve analisi dei dati più recenti sull'economia sommersa e illegale – relativi all'anno 2023 – e il loro andamento dal 2011; richiama il contributo dell'Istat alla stima dell'evasione fiscale e contributiva nelle diverse sedi istituzionali.

La seconda parte entra nel dettaglio della strategia dell'Istituto in tema di gestione e produzione delle statistiche nonché di rafforzamento della *data governance* e descrive l'approccio alla tutela dei dati personali in conformità al Regolamento europeo 2016/679 (GDPR – *General Data Protection Regulation*).

In generale, è opportuno ricordare che lo scopo istituzionale dell'Istat è servire la collettività attraverso la produzione, lo sviluppo e la diffusione di informazioni statistiche ufficiali, favorendo i processi decisionali di istituzioni pubbliche, operatori economici e cittadini. I dati trattati, a prescindere dalla fonte (rilevazione diretta, dato amministrativo o nuove fonti di dati), sono tutelati dal segreto statistico¹ e possono essere utilizzati esclusivamente per fini statistici; la loro diffusione, inoltre, è effettuata in forma aggregata, secondo modalità che non consentono l'identificazione del soggetto al quale si riferiscono.

1. L'economia non osservata nei Conti nazionali: aspetti metodologici, principali evidenze e il contributo alla stima dell'evasione fiscale e contributiva

Le componenti dell'economia non osservata e l'impianto metodologico

I Conti nazionali si fondano su schemi di misurazione che includono anche l'"Economia non osservata" (NOE, dall'inglese *Non-Observed Economy*), ovvero l'insieme delle attività economiche che, per diverse ragioni, sfuggono alla rilevazione statistica diretta.

L'inclusione dell'economia non osservata nel perimetro di misurazione dei Conti nazionali consente di garantire il rispetto del principio di esaustività nella rappresentazione dei flussi economici, in conformità con quanto previsto dai manuali internazionali SNA (System of National Accounts) e SEC (Sistema europeo

¹ Si veda l'art. 9 del [d. lgs. n. 322/89](https://www.gazzettaufficiale.it/eli/2017/05/23/01322899/01/pt).

dei conti nazionali e regionali)². Questo approccio, sottoposto al controllo delle autorità statistiche europee, assicura la comparabilità a livello internazionale dei dati relativi al Prodotto interno lordo (Pil) e al Reddito nazionale lordo (Rnl).

La NOE è costituita, a livello definitorio, da quattro componenti:

- 1) Il *sommerso economico* include tutte le attività volontariamente celate alle autorità fiscali, previdenziali e statistiche; è costituito dall'ammontare di valore aggiunto non osservato derivante dalle dichiarazioni mendaci riguardanti il fatturato e/o i costi (sotto-dichiarazione del valore aggiunto) o dall'utilizzo di input di lavoro non regolare.
- 2) L'*economia illegale* include le attività che producono beni e servizi illegali o che, pur riguardando beni e servizi legali, sono svolte senza adeguata autorizzazione o titolo; seppure il sistema dei Conti nazionali dovrebbe registrare, all'interno dei confini della produzione, tutte le attività illegali caratterizzate dal mutuo consenso fra i contraenti³, sulla base delle indicazioni fornite da Eurostat sono state incluse nel sistema dei Conti il traffico di droga, i servizi di prostituzione e il contrabbando di tabacco.
- 3) Il *sommerso statistico* include tutte le attività non osservate per motivi riferibili alle inefficienze informative che caratterizzano le basi di dati (errori campionari e non) o per errori di copertura negli archivi.
- 4) L'*economia informale* include tutte le attività produttive svolte in contesti poco o per nulla organizzati, basati su rapporti di lavoro non regolati da contratti formali, ma definiti nell'ambito di relazioni personali o familiari.

L'attuale impianto metodologico per la stima della NOE è stato introdotto in occasione della revisione generale dei Conti nazionali connessa all'adozione del SEC 2010. In questo contesto, l'Istat ha investito importanti risorse nello sviluppo di un progetto organico volto all'aggiornamento e all'ulteriore potenziamento delle fonti informative e dei metodi di misurazione dell'economia non osservata, anche attraverso un confronto sistematico con i principali stakeholder istituzionali⁴,

² L'applicazione del SEC 2010 è stabilita dal Regolamento (Ue) del Parlamento europeo e del Consiglio, n. 549/2013, che definisce il Sistema europeo dei conti nazionali e regionali dell'Unione europea. A livello globale, il riferimento metodologico è rappresentato dal System of National Accounts (SNA) delle Nazioni Unite, la cui versione più recente risale al 2008.

³ Stimare alcune attività dell'economia illegale non equivale a misurare il fatturato o la ricchezza delle organizzazioni criminali, sia perché l'analisi è limitata a un sotto-insieme di attività, sia perché non si prendono in considerazione le attività legali possedute da soggetti criminali.

⁴ L'Istat ha costituito una Commissione incaricata di sviluppare e valutare nuove metodologie per la stima del sommerso economico. Oltre ai ricercatori dell'Istat, ne hanno fatto parte esponenti del mondo accademico e di importanti enti pubblici, tra cui l'Agenzia delle Entrate, l'Agenzia delle Dogane, il Ministero del Lavoro e la Banca d'Italia, con funzioni di *advisor* e *peer reviewer*. I lavori della Commissione hanno guidato e supportato la revisione dei metodi di stima esistenti e la definizione di nuove strategie per misurare alcune specifiche componenti dell'economia non osservata, quali la sotto-dichiarazione del valore aggiunto e il ricorso al lavoro non regolare, nonché un nuovo approccio alla misurazione dell'input di lavoro non regolare.

rafforzando così una tradizione già consolidata e la propria posizione di riferimento nel panorama internazionale.

La metodologia adottata dall'Istat per la stima della NOE si fonda su due approcci tra loro interconnessi: da un lato, lo sviluppo del Sistema Integrato dei Registri (SIR) e la disponibilità di dati amministrativi a livello micro, relativi sia agli individui sia alle unità economiche, hanno consentito un significativo ampliamento della base informativa su imprese e mercato del lavoro; dall'altro, tale disponibilità informativa ha reso possibile lo sviluppo di stime micro-fondate basate su modelli comportamentali e di profilo, migliorando la coerenza tra i metodi di stima e le caratteristiche dei fenomeni da misurare e di conseguenza l'affidabilità complessiva delle stime.

In Italia, la misurazione dell'economia non osservata è limitata alle componenti del sommerso economico e dell'economia illegale. L'ampliamento della copertura informativa, conseguito mediante il ricorso estensivo alle fonti amministrative per l'integrazione dei Registri statistici, ha reso possibile l'individuazione pressoché esaustiva delle attività economiche regolari svolte da unità produttive caratterizzate da assetti organizzativi minimi⁵, determinando la sostanziale trascurabilità dell'economia informale; inoltre, la significativa riduzione del ricorso a dati di tipo campionario nelle stime e, conseguentemente, la virtuale eliminazione dell'errore statistico ha reso marginale l'incidenza del sommerso attribuibile a carenze delle basi informative.

I dettagli sulle stime delle diverse componenti del sommerso economico e dell'economia illegale sono descritti nella Nota metodologica allegata al Comunicato Stampa su "L'economia non osservata nei Conti nazionali", diffusa lo scorso 17 ottobre.⁶ L'Istituto è naturalmente disponibile a fornire ulteriori chiarimenti riguardo le metodologie di stima.

Le principali evidenze

L'Istat diffonde con regolarità le stime dell'economia non osservata, sia a livello nazionale⁷ sia territoriale⁸, coerenti con i Conti nazionali pubblicati ogni anno a settembre, collocandosi tra i pochi Paesi europei che rendono disponibili tali informazioni con continuità. Le stime, diffuse a ottobre di ogni anno, sono relative a due anni precedenti quello di pubblicazione, poiché solo con questo ritardo temporale si rende disponibile la base informativa microfondata completa necessaria a implementare la metodologia illustrata.⁹

Sulla base dei Conti nazionali pubblicati a settembre, nel 2023 il valore aggiunto generato dall'economia non osservata si è attestato a 217,5 miliardi di euro, con

⁵ Il Sistema dei Registri include il registro ASIA-unità giuridiche che include tutte le unità economiche e le interrelazioni esistenti tra di esse (circa 14 milioni di operatori economici).

⁶ Si veda il Report "[L'economia non osservata nei Conti nazionali – Anni 2020-2023](#)".

⁷ Si veda la nota 6.

⁸ Si veda il Report "[Conti economici territoriali - 2022-2024](#)".

⁹ Ad esempio, a ottobre 2025 sono state diffuse le informazioni aggiornate relative all'anno 2023.

un'incidenza sul Pil pari al 10,2%; tale incidenza risulta inferiore di 0,6 punti percentuali rispetto al 2019 (10,8%) e di 1,6 punti rispetto al 2011 (11,8%), primo anno disponibile della serie storica. In particolare, dopo aver raggiunto il 12,5% del Pil nel 2014, l'incidenza è scesa gradualmente fino a toccare il 10% nel 2020; nel triennio successivo, si è mantenuta su livelli prossimi a tale quota, evidenziando una lieve ripresa.¹⁰

Le componenti dell'economia sommersa presentano, in termini di incidenza sul Pil, un'evoluzione nel tempo sostanzialmente analoga all'aggregato generale; la presenza di intensità differenti nella dinamica delle singole componenti si riflette però in una diversa evoluzione del loro contributo all'economia non osservata. Si osserva, in particolare, un ridimensionamento del peso del valore aggiunto generato dall'impiego di lavoro irregolare, la cui incidenza scende al 35,5% nel 2023 dal 38,4% del 2014; in parallelo, aumenta il peso relativo della sotto-dichiarazione del valore aggiunto, che si attesta al 49,7% nel 2023 (dal 47,6% del 2014); più limitata rimane, infine, la rilevanza delle altre componenti – mance, fitti non dichiarati e integrazione domanda-offerta – la cui incidenza si riduce al 5,6% nel 2023 (6,0% nel 2014). Il peso delle attività illegali risulta invece in aumento nello scorso decennio, passando dall'8,2% del 2011 al 10,5% del 2020, per poi attestarsi al 9,2% nel 2023.

Le nostre stime non consentono di valutare direttamente l'effetto prodotto da cambiamenti nei quadri normativi; l'evoluzione appena descritta appare tuttavia coerente con le misure avviate nel tempo dalle politiche di contrasto all'evasione (quali ad esempio l'introduzione e il successivo rafforzamento della fatturazione elettronica e la connessione telematica dei registratori fiscali), cui si sono associati i cambiamenti strutturali indotti dall'emergenza sanitaria. In particolare, se da un lato le politiche pubbliche sembrano aver contribuito a ridurre l'impatto dell'economia sommersa, dall'altro la pandemia e le relative misure di mitigazione degli effetti economici avversi hanno consolidato tale riduzione, accelerando i processi in atto; le conseguenze dell'emergenza sanitaria, infatti, hanno determinato non solo una temporanea contrazione delle attività irregolari, ma anche una diffusione di comportamenti più virtuosi da parte degli operatori, con un rafforzamento dei processi di regolarizzazione del lavoro e una più ampia diffusione dei pagamenti elettronici, incentivati sia dalle restrizioni sanitarie sia dalle politiche pubbliche.

Nel seguito si riporta una breve sintesi dell'andamento recente delle diverse componenti dell'economia non osservata.

Nel 2023, il valore aggiunto generato dalla sotto-dichiarazione è aumentato del 6,6% rispetto all'anno precedente, raggiungendo i 108,2 miliardi di euro (il 5% del Pil); quello derivante dall'impiego di lavoro irregolare ha segnato un incremento più sostenuto (+11,3%), attestandosi a 77,2 miliardi (3,6% del Pil). Le altre componenti del sommerso (mance e fitti non dichiarati) hanno evidenziato una crescita più contenuta (+3,8%), per un valore complessivo di 12,2 miliardi (0,6% del Pil). Nel complesso, il sommerso economico ammonta a 197,6 miliardi (+8,2%). Le attività

¹⁰ La Tavola 1 dell'Allegato statistico dà conto delle principali evidenze descritte in questa sezione.

illegali hanno generato nel 2023 un valore aggiunto pari a 20 miliardi (0,9% del Pil), in crescita dell'1% sul 2022. Tra queste, il traffico di stupefacenti rappresenta la componente prevalente (15,3 miliardi), seguito dai servizi di prostituzione (4,1 miliardi); il contrabbando di tabacco mantiene invece un ruolo più marginale.

Dal punto di vista settoriale, il sommerso economico presenta un'incidenza più elevata negli Altri servizi alle persone (il 32,4% del valore aggiunto del comparto nel 2023), nel Commercio, trasporti, alloggio e ristorazione (18,8%) e nelle Costruzioni (16,5%); valori più contenuti si osservano nei comparti industriali e negli altri servizi alle imprese.

Nel 2023 le unità di lavoro a tempo pieno (Ula) in condizione di non regolarità ammontano a 3 milioni e 132 mila, in aumento del 4,9% rispetto all'anno precedente. Il tasso di irregolarità si attesta al 12,7%, in lieve crescita dopo un quinquennio di riduzione, per effetto di una dinamica del lavoro non regolare più sostenuta rispetto a quella del lavoro regolare. L'incidenza del lavoro non regolare risulta più rilevante nel terziario (13,9%), in particolare negli Altri servizi alle persone (40,5%), e in Agricoltura, silvicoltura e pesca (17,6%).

L'incidenza dell'economia non osservata sul valore aggiunto evidenzia una marcata eterogeneità sul territorio nazionale. Nel 2023, tale incidenza risulta superiore alla media nazionale nelle ripartizioni del Mezzogiorno (16,5%) e del Centro (11,8%). Al contrario, nelle aree del Nord-est e del Nord-ovest si osservano valori inferiori alla media, pari rispettivamente al 9,3% e all'8,9%. In tutte le ripartizioni territoriali prevale la componente della sotto-dichiarazione, con un'incidenza massima nel Mezzogiorno (7,6%) e minima nel Nord-ovest (4,5%). Anche il lavoro irregolare presenta una maggiore rilevanza nel Mezzogiorno (6,5%), si attesta su valori prossimi alla media nazionale nel Centro (4%) e risulta più contenuto nel Nord-est (3,1%) e nel Nord-ovest (3%). A livello regionale, l'incidenza dell'economia non osservata varia dal 19% della Calabria al 7,4% della Provincia autonoma di Bolzano/Bozen. La sotto-dichiarazione raggiunge i valori più elevati in Puglia (8,3%), Sardegna e Marche (7,7%), mentre risulta più contenuta nelle Province autonome di Bolzano (2,9%) e Trento (3,5%) e in Lombardia (4,2%). Il valore aggiunto generato da lavoro irregolare è particolarmente rilevante in Calabria (8,3%), Campania (7%) e Sicilia (6,4%), mentre assume valori più ridotti in Lombardia (2,8%), nella Provincia autonoma di Bolzano (2,9%) e in Veneto (3%). Le attività illegali e le altre componenti dell'economia non osservata incidono come detto in misura più limitata, con valori compresi tra l'1,2% (Lombardia e Veneto) e il 3,2% (Calabria).

Il contributo dell'Istat alla stima dell'evasione fiscale e contributiva nelle diverse sedi istituzionali

- ✓ *La Commissione per la redazione della Relazione sull'economia non osservata e sull'evasione fiscale e contributiva*

La "Relazione sull'economia non osservata e sull'evasione fiscale e contributiva", predisposta annualmente dalla Commissione come disposto dalla legge di contabilità

e finanza pubblica¹¹, recepisce e valuta le stime sull'economia non osservata effettuate dall'Istat sulla base della normativa comunitaria che regola la compilazione dei Conti economici nazionali. In particolare, ne è parte integrante la stima del sommerso economico elaborata dall'Istat a livello nazionale e territoriale.¹²

Nell'ambito della Relazione l'Istat elabora anche la stima dell'evasione contributiva associata al lavoro dipendente irregolare, basata sulle stime dei Conti nazionali relative all'input di lavoro non regolare e alle corrispondenti retribuzioni. Le retribuzioni irregolari costituiscono la base imponibile potenziale per la quantificazione dei contributi evasi, sia per la componente a carico dei datori di lavoro sia per quella a carico dei lavoratori. Per la componente datoriale si utilizzano aliquote contributive implicite desunte dal Sistema informativo Frame-SBS; per quella a carico dei lavoratori si applicano le aliquote di legge, corrette mediante un fattore di aggiustamento volto a ricondurre la retribuzione alla base imponibile contributiva.

L'evasione contributiva è stimata secondo due ipotesi alternative:

- i. ipotesi massima, che assume retribuzioni dei lavoratori irregolari analoghe a quelle dei lavoratori regolari;
- ii. ipotesi minima, che assume l'invarianza delle retribuzioni dei lavoratori irregolari rispetto ai valori definiti nell'ambito dei Conti nazionali.

Le due ipotesi definiscono un intervallo di valori plausibili.

Le stime sono ottenute nell'ipotesi di invarianza dell'input di lavoro (modello statico), senza considerare effetti comportamentali derivanti dall'emersione.

Nel 2023, il valore dei contributi evasi evidenzia un incremento del 10% nell'ipotesi massima e del 9,9% nell'ipotesi minima, confermando una marcata ripresa rispetto ai livelli particolarmente contenuti del biennio 2020-2021 e riallineandosi su valori prossimi a quelli osservati nel periodo pre-pandemico. L'ammontare complessivo dell'evasione contributiva si attesta a 12,7 miliardi di euro nell'ipotesi massima (+1,2

¹¹ La legge 31 dicembre 2009, n. 196 (Legge di contabilità e finanza pubblica), all'art. 10-bis.1. ("Monitoraggio dell'evasione fiscale e contributiva"), comma 1, ha previsto che, contestualmente alla Nota di aggiornamento del Documento di economia e finanza 2016, sia presentato un "Rapporto sui risultati conseguiti in materia di misure di contrasto all'evasione fiscale e contributiva" per la cui redazione il Governo si avvale della Relazione sull'economia non osservata e sull'evasione fiscale e contributiva, predisposta da una Commissione istituita con decreto del Ministro dell'Economia e delle Finanze. La Commissione è composta da quindici esperti nelle materie economiche, statistiche, fiscali, lavoristiche o giuridico-finanziarie, di cui un rappresentante della Presidenza del Consiglio dei ministri, quattro rappresentanti del Ministero dell'Economia e delle Finanze, due rappresentanti dell'Istat, un rappresentante del Ministero del lavoro e delle politiche sociali, un rappresentante dell'Inps, un rappresentante dell'Associazione nazionale dei comuni italiani, un rappresentante della Conferenza delle regioni e delle province autonome di Trento e di Bolzano, un rappresentante della Banca d'Italia e tre professori universitari.

¹² Si veda la "[Relazione sull'economia non osservata e sull'evasione fiscale e contributiva Anno 2025](#)" (capitolo II) e "[Aggiornamenti per gli anni 2019-2023 a seguito della revisione dei Conti nazionali apportata dall'Istat](#)".

miliardi rispetto al 2022) e a 9,2 miliardi nell'ipotesi minima (+0,8 miliardi rispetto al 2022).

Ai fini delle stime del *tax gap* elaborate dall'Agenzia delle Entrate per la Relazione sull'economia non osservata e sull'evasione fiscale e contributiva, l'Istat fornisce un'articolata base informativa, costituita da dati di contabilità nazionale in serie storica, con un livello di dettaglio per settore di attività economica e settore istituzionale superiore a quello di diffusione. Tali dati sono forniti due volte l'anno in occasione delle diffusioni dei Conti economici nazionali. Per la stima del *tax gap* relativo all'Irap e alle imposte dirette, l'Istat mette a disposizione dell'Agenzia delle Entrate basi dati a livello micro, definite nell'ambito di una Convenzione sottoscritta tra le due istituzioni.¹³

✓ *Il Comitato nazionale per la prevenzione e il contrasto del lavoro sommerso*

Il Comitato costituisce lo strumento di coordinamento delle politiche di contrasto al lavoro sommerso previste dal Piano Nazionale di Ripresa e Resilienza (PNRR). È istituito presso il Ministero del lavoro e delle politiche sociali ed è composto da rappresentanti delle principali amministrazioni centrali e degli enti coinvolti nel contrasto al lavoro sommerso, tra cui l'Istat. La partecipazione dell'Istat è finalizzata al supporto statistico e metodologico per la produzione degli indicatori e per il monitoraggio del fenomeno.

Il Comitato svolge funzioni di:

- coordinamento e monitoraggio dell'attuazione del Piano nazionale per la lotta al lavoro sommerso;
- raccordo tra i soggetti istituzionali coinvolti;
- costruzione, validazione e certificazione degli indicatori PNRR (M5C1-10 e M5C1-11);
- definizione degli interventi per i cicli di programmazione successivi.

L'Istat contribuisce alle attività del Comitato attraverso:

- la produzione di statistiche ufficiali sul lavoro non regolare nell'ambito della Contabilità nazionale;
- il supporto metodologico alla costruzione degli indicatori previsti dal PNRR;
- la partecipazione allo sviluppo del Portale Nazionale del Lavoro Sommerso (PNS) e del patrimonio informativo integrato;
- il contributo alle attività di analisi e valutazione delle politiche;
- la partecipazione ai gruppi di lavoro tematici, incluso il Tavolo caporalato, fornendo supporto analitico e contributi alla relazione annuale.

¹³ Convenzione tra l'Agenzia delle Entrate e l'Istituto nazionale di statistica per la fruizione dei servizi di cooperazione informatica.

Per quanto riguarda l'indicatore macroeconomico coincidente del lavoro sommerso, le analisi condotte dal gruppo di lavoro Istat-Banca d'Italia evidenziano l'impossibilità di produrre, nei tempi richiesti, una stima anticipata completa dell'incidenza del lavoro irregolare coerente con i Conti nazionali, a causa della mancata disponibilità tempestiva delle fonti informative necessarie. È tuttavia possibile stimare indicatori parziali, quali la mancata copertura di secondo livello (COP2), che misura la quota di individui con posizione lavorativa dichiarata non riscontrata negli archivi amministrativi. Tale indicatore presenta limiti connessi sia alla copertura delle fonti sia alla qualità degli abbinamenti informativi.

In conclusione, il contributo dell'Istat è essenziale ai fini della disponibilità di basi informative integrate, della definizione degli indicatori e della valutazione degli interventi, assicurando coerenza metodologica e qualità statistica al sistema di monitoraggio.

2. La strategia dell'Istat per la produzione statistica multi-fonte: *data governance* e protezione dei dati

Negli ultimi anni l'Istat ha avviato una trasformazione profonda del proprio modello di produzione dei dati orientata all'adozione di un approccio multi-fonte, centrato in particolare sui Registri statistici; quest'evoluzione risponde alla necessità di far fronte a sfide che accomunano tutti gli Istituti nazionali di statistica, quali il progressivo calo dei tassi di risposta alle indagini e i costi elevati delle rilevazioni campionarie su larga scala; il nuovo approccio ha inoltre consentito la riduzione del carico statistico sui rispondenti, come ad esempio recentemente concretizzato nell'ambito delle indagini economiche congiunturali attraverso la collaborazione con l'Agenzia delle Entrate.¹⁴

In questo contesto l'attenzione si è spostata sempre più verso l'individuazione e l'utilizzo di fonti complementari e in particolare sull'uso dei dati amministrativi, oltre all'esplorazione di nuove fonti come i big data. L'obiettivo è costruire un sistema informativo più efficiente, sostenibile e capace di integrare informazioni provenienti da diverse fonti.

A partire dal 2016, l'Istat ha adottato un nuovo modello di produzione statistica basato sull'integrazione sistematica di dati di diversa natura, combinando dati amministrativi¹⁵ con quelli da indagine. Tale cambiamento ha comportato profonde innovazioni metodologiche, tecnologiche e organizzative, accompagnate da una revisione sostanziale dei processi tradizionali. In particolare, metodi di *data integration* e modelli di stima per piccole aree sono diventati strumenti centrali per combinare in modo efficace le informazioni; anche metodologie consolidate, quali

¹⁴ Si veda il Comunicato "[Statistiche con fatturazione elettronica per la riduzione degli adempimenti](#)", diffuso lo scorso 18 marzo.

¹⁵ Si fa presente che l'acquisizione da parte dell'Istat delle fonti amministrative, tra cui rientrano i dati fiscali, è programmata nell'ambito del Piano Statistico Nazionale (PSN) in cui sono previste le specifiche finalità statistiche connesse all'utilizzo e integrazione di tali dati.

stima, calibrazione e trattamento degli errori di misura, hanno richiesto un adattamento per la loro applicazione in un contesto multi-fonte.

Il Sistema Integrato dei Registri (SIR) costituisce l'infrastruttura portante di questo nuovo sistema di produzione. Esso è costruito a livello di microdato attraverso un'integrazione di archivi amministrativi e dati da indagine. Il processo di costruzione dei Registri prevede l'impiego di tecniche di *record linkage*, procedure di controllo e correzione dei dati e metodi di imputazione dei valori mancanti. Il Sistema consente di disporre di un patrimonio informativo ampio e dettagliato sull'intera popolazione, sulle unità economiche e su quelle territoriali e si articola in diversi Registri di Base e Tematici, tra i quali quelli sulla popolazione, sulle unità economiche, sul lavoro, sull'istruzione e i percorsi formativi, sui redditi e sulla disabilità.

Ogni anno l'insieme dei dati aggiornati dei Registri configura la parte preponderante del patrimonio informativo attraverso cui l'Istat ottempera agli obblighi normativi nazionali ed europei in merito ai censimenti. Con la chiusura della tornata censuaria del 2010-2011 l'Istat ha, infatti, avviato la nuova stagione dei censimenti permanenti della popolazione e delle unità economiche (inclusa l'Agricoltura), formalmente avviati con la legge di bilancio 2018. I censimenti permanenti non coinvolgono direttamente tutti i cittadini e le unità economiche ma di volta in volta un loro campione appositamente progettato. La restituzione al Paese dei dati ottenuti è però di tipo censuario: un risultato possibile attraverso l'integrazione tra le fonti amministrative e le indagini.

La crescente integrazione di fonti eterogenee rende centrale il tema della *data governance*, in particolare per quanto riguarda la gestione dei metadati e delle semantiche, e indispensabile il ricorso a nuovi approcci rigorosi alla tutela dei dati personali in conformità al GDPR.

Una parte rilevante del lavoro di integrazione riguarda la comprensione, il controllo e l'armonizzazione dei metadati, ossia dell'informazione dei dati nella sua interezza. L'Istat sta sviluppando un sistema di metadati basato su un glossario comune, con l'obiettivo di garantire coerenza semantica tra i diversi processi di produzione statistica e documentare in modo ancora più trasparente l'origine e le trasformazioni dei dati. In questo contesto, la dimensione semantica diventa cruciale: integrare fonti diverse significa assicurare che concetti, definizioni e classificazioni siano allineati e interpretabili in modo univoco. Questo viene realizzato utilizzando linguaggi standard propri del web semantico, ossia le ontologie, che consentono di realizzare integrazione e interoperabilità semantica dei dati elaborabili direttamente dalle macchine.

Come osservato, nell'ambito della strategia dell'Istituto un tema centrale riguarda la politica sul trattamento dei dati personali. In questo contesto è stato sviluppato, tenendo conto degli indirizzi forniti dal Garante per la Privacy, un algoritmo di pseudonimizzazione che consente di trattare in modo adeguato le tipologie di dati in base alla loro caratteristica. Tale algoritmo è pienamente integrato all'interno di un sistema in grado di governare i dati fornendo ai ricercatori dell'Istituto la possibilità di trattarli in modo sicuro.

Tale sistema per la Gestione dei Microdati Amministrativi e statistici, denominato SIGMA, permette di svolgere i trattamenti statistici di dati personali all'interno di Domini Specifici di Integrazione (DSI).

Il DSI è l'ambiente di elaborazione in cui sono presenti solo i dati necessari per conseguire una specifica finalità statistica, per il tempo strettamente necessario e a cui hanno accesso esclusivamente le persone opportunamente autorizzate. La realizzazione di SIGMA è stata perseguita dall'Istat per ottemperare alle disposizioni dell'Autorità Garante per la protezione dei dati personali. I dati sono caricati, organizzati in strutture tabellari e sottoposti a procedure di pseudonimizzazione che si suddividono nelle seguenti fasi:

1. separazione delle variabili identificative dalle altre variabili tematiche e dai dati di cui agli art. 9 e 10 del GDPR;
2. individuazione delle unità statistiche di base mediante procedure di record linkage di tipo deterministico;
3. apposizione del codice pseudonimo alle unità statistiche: ogni unità viene "registrata" e resa individuabile univocamente mediante il codice pseudonimo.

Lo pseudonimo consente di non utilizzare i dati identificativi diretti mantenendo la completa possibilità di eseguire le analisi statistiche necessarie per la finalità da perseguire; la pseudonimizzazione delle variabili è specifica per ogni singolo DSI.

Il sistema SIGMA, inoltre, gestisce il tracciamento e monitoraggio di tutte le richieste di dati; in questo modo si garantisce la *data governance* in modalità *privacy by design e by default*.

Il funzionamento del sistema è basato su un catalogo dei dati che contiene tutti i metadati relativi alle fonti statistiche: archivi amministrativi e prodotti statistici. Ogni variabile è classificata sulla base del GDPR e caratterizzata da attributi che ne guidano il rilascio in modo controllato limitandone la visibilità, ove necessario.

Questo complesso sistema, originato dal provvedimento del Garante per la protezione dei dati personali nel 2020¹⁶, ha ottenuto nel 2025 l'approvazione conclusiva da parte del Garante. In particolare *"Il Garante privacy ha dato l'ok alle misure tecniche e organizzative adottate dell'Istituto nazionale di statistica (Istat) per garantire la tutela dei dati personali trattati e l'effettiva applicazione dei principi di protezione dei dati. Tali soluzioni, che hanno determinato un importante cambiamento nella gestione e utilizzazione delle banche dati da parte dell'Istituto, nascono da due provvedimenti prescrittivi del Garante del 2020 e del 2023. Tra le misure allora ingiunte, la necessità di dotarsi di un sistema di pseudonimizzazione dei dati robusto, avanzato e adeguato a tutelare l'enorme mole di dati trattati. L'Istat è infatti il principale depositario del patrimonio informativo del Paese e può acquisire dati personali da tutte le banche dati pubbliche nonché da soggetti privati oltre che dai social media. Attualmente l'Istituto, in ottemperanza agli obblighi di privacy by*

¹⁶ Si veda il [Provvedimento](#) del 23 gennaio 2020.

design e by default, a seguito degli interventi dell’Autorità, ha adottato misure di pseudonimizzazione adeguate a garantire la tutela dei dati trattati e un’effettiva applicazione dei principi di protezione dei dati personali, in particolare quelli di minimizzazione, di limitazione della finalità e della conservazione nei trattamenti necessari per la produzione della statistica ufficiale. [...] Mediante la classificazione e mappatura delle fonti e dei dati personali acquisiti, la definizione degli ambiti di utilizzo e l’implementazione di una codifica gerarchica degli pseudonimi, ciascuno con una validità limitata nel tempo e rispetto alla specifica finalità perseguita, l’Istat ha invece potuto mettere in atto un sistema di governance dei dati conforme al Regolamento. La rilevanza delle misure adottate rende oggi l’Istat – a parere dell’Autorità – punto di riferimento per le altre amministrazioni pubbliche che gestiscono grandi banche dati a fini statistici.”¹⁷

Il sistema di pseudonimizzazione sopra descritto ha anche come conseguenza il rafforzamento delle misure di sicurezza informatica, che sono comunque certificate conformi allo standard ISO/IEC 27001 fin dal 2017.

Relativamente alle richieste di rettifica, opposizione e cancellazione, queste sono gestite dal Responsabile della Protezione dei Dati (RPD) dell’ente in conformità al GDPR, tenendo conto che l’Istat opera sulla base di un compito di interesse pubblico.

Anche i dati provenienti dai modelli fiscali sono trattati tramite il sistema SIGMA. Questi sono oggetto, però, di un pre-trattamento al fine di organizzare le informazioni pseudonimizzate nelle unità che costituiscono il modello, ad esempio nei quadri o sezioni, al fine di mantenere i legami tra tali unità senza dover utilizzare identificativi diretti. L’organizzazione che viene ottenuta dal pre-trattamento è unica per tutti i processi statistici che li dovranno elaborare e questo costituisce un grande vantaggio data la complessità delle fonti fiscali.

Analoga applicazione del sistema si ha con le banche dati acquisite tramite la Piattaforma Digitale Nazionale Dati (PDND).

L’Istat in questo contesto è impegnato in un progetto europeo che ha l’obiettivo di sviluppare tecniche di produzione di dati sintetici con l’obiettivo di gestire dati critici in tutta sicurezza e soprattutto in grado di generare dataset per l’addestramento di algoritmi di IA privi di *bias*. Parallelamente, l’Istituto sta rafforzando il proprio ruolo anche sul piano esterno, promuovendo l’interoperabilità semantica tra le Pubbliche Amministrazioni (PA) e assumendo il ruolo di coordinamento nella definizione e armonizzazione della semantica dei dati. Questo implica una collaborazione stretta con enti produttori di dati amministrativi, al fine di rendere i loro archivi pienamente integrabili nel sistema statistico e anche tra sistemi di dati delle PA. In tal modo, la *governance* dei dati si estende oltre i confini dell’Istituto e diventa un elemento chiave per il funzionamento dell’intero ecosistema informativo pubblico. In questa prospettiva l’Istat ha messo a disposizione delle PA la sua esperienza di *governance* dei dati nel progetto di realizzazione del Catalogo Nazionale Dati.

¹⁷ Si veda la [NEWSLETTER](#) dell’8 maggio 2025.

Guardando alle prospettive future, la strategia dell'Istat prevede un ulteriore ampliamento del patrimonio informativo utilizzato, attraverso l'inclusione di fonti non tradizionali, quali i big data, e lo sviluppo di soluzioni tecnologiche avanzate a supporto dei processi di produzione statistica. Nei processi produttivi sono già utilizzati gli scanner data per la produzione dell'indice dei prezzi al consumo; risultano invece in fase di sperimentazione ulteriori fonti, tra cui dati satellitari e ortofoto, dati AIS (*Automatic Identification System*), oltre a strumenti evoluti di raccolta dati come le *smart survey*.

L'utilizzo dei big data rappresenta un'opportunità rilevante per ampliare la capacità descrittiva del sistema statistico, grazie alla loro elevata granularità, tempestività e copertura. Tuttavia, queste fonti presentano caratteristiche profondamente diverse rispetto ai dati tradizionali e di conseguenza il loro impiego richiede un adattamento delle metodologie e delle infrastrutture. In questo contesto, emerge la necessità di innovazioni tecnologiche in grado di gestire volumi elevati di dati, garantendo al contempo efficienza, scalabilità e sicurezza. Ciò implica lo sviluppo e l'adozione di architetture informatiche avanzate, capaci di supportare l'acquisizione, l'archiviazione e l'elaborazione distribuita dei dati. Si rendono inoltre necessari strumenti evoluti per il trattamento automatico dei dati; in questa prospettiva, sono in corso attività di ricerca e sviluppo finalizzate all'applicazione di tecniche di intelligenza artificiale e *machine learning*, con l'obiettivo di migliorare l'efficienza e la qualità dei processi statistici, in particolare nelle fasi di integrazione, classificazione, imputazione e analisi dei dati.

In conclusione, la crescente integrazione di fonti eterogenee comporta nuove sfide, che rendono necessario definire e adottare standard e criteri metodologici per la valutazione, la gestione e la comunicazione dell'affidabilità, qualità e sicurezza dei dati e delle statistiche prodotte, in un contesto caratterizzato da elevata complessità, varietà e dinamicità delle fonti disponibili.

Allegato statistico

Tavola 1 - Economia non osservata (Non-Observed Economy, NOE) e sue componenti. Anni 2011-2023

(valori correnti in milioni di euro e incidenza percentuale sul Pil)

Componenti	Valori correnti in milioni di euro												
	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Economia sommersa	179.232	179.953	180.806	187.597	183.369	181.065	186.558	182.011	175.408	149.900	165.510	182.642	197.570
- da Sottodichiarazione	93.265	97.096	97.923	97.056	92.625	94.604	97.846	94.037	88.978	79.287	90.467	101.510	108.171
- da Lavoro irregolare	70.314	70.943	71.312	78.259	76.856	74.649	76.699	74.977	74.206	60.047	65.509	69.359	77.174
- Altro	15.653	11.913	11.570	12.282	13.887	11.812	12.013	12.997	12.224	10.566	9.534	11.773	12.225
Attività illegali	16.105	17.346	17.650	16.464	17.233	18.078	18.896	19.368	19.610	17.556	18.531	19.774	19.972
NOE	195.338	197.299	198.456	204.061	200.602	199.142	205.455	201.378	195.017	167.456	184.041	202.416	217.541
Valore aggiunto	1.480.015	1.455.351	1.445.574	1.453.768	1.479.704	1.515.429	1.552.004	1.582.453	1.604.270	1.496.322	1.644.016	1.793.607	1.924.798
PIL	1.657.362	1.632.899	1.621.261	1.635.871	1.663.278	1.704.857	1.744.493	1.777.744	1.804.067	1.670.012	1.842.507	1.998.073	2.142.602

Componenti	Incidenza percentuale sul Pil												
	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Economia sommersa	10,8	11,0	11,2	11,5	11,0	10,6	10,7	10,2	9,7	9,0	9,0	9,1	9,2
- da Sottodichiarazione	5,6	5,9	6,0	5,9	5,6	5,5	5,6	5,3	4,9	4,7	4,9	5,1	5,0
- da Lavoro irregolare	4,2	4,3	4,4	4,8	4,6	4,4	4,4	4,2	4,1	3,6	3,6	3,5	3,6
- Altro	1	1	1	1	1	1	1	1	1	1	1	1	1
Attività illegali	1,0	1,1	1,1	1,0	1,0	1,1	1,1	1,1	1,1	1,1	1,0	1,0	0,9
NOE	11,8	12,1	12,2	12,5	12,1	11,7	11,8	11,3	10,8	10,0	10,0	10,1	10,2

Fonte: Istat, Conti Nazionali