

Roma, 9 aprile 2024

Governare con i numeri:  
ricerca, elaborazione e presentazione dei dati

# LA PRIVACY BY DESIGN E BY DEFAULT NELL'ACQUISIZIONE DELLE FONTI AMMINISTRATIVE

# Indice della presentazione

---

- I dati personali
- Il trattamento dei dati personali
- I diversi ruoli nel trattamento di dati personali
- Data protection by default and by design
- Rischio inerente al trattamento
- Adempimenti da parte del titolare
- Valutazione d'impatto della protezione dei dati (DPIA)
- I flussi di dati gestiti dalla Direzione Centrale della Raccolta Dati
- Gli interventi del Garante su SIM
- Le osservazioni del Garante su SIM (Parere sul Piano Generale del Censimento)
- Le indicazioni del Garante per l'evoluzione di SIM
- La soluzione individuata dall'Istat per l'evoluzione di SIM
- La pseudonimizzazione secondo la logica dei domini specifici di integrazione
- I domini specifici di integrazione
- I domini specifici di integrazione dinamici
- La classificazione dei dati

# I dati personali

---

Qualsiasi **informazione che identifichi o renda identificabile, direttamente o indirettamente, una persona fisica** e che possa fornire informazioni sulle sue caratteristiche, le sue abitudini, il suo stile di vita, le sue relazioni personali, il suo stato di salute, la sua situazione economica, ecc.

Particolarmente rilevanti:

- **I dati che permettono l'identificazione diretta** - come i dati anagrafici (ad esempio: nome e cognome), le immagini, ecc. - e i **dati che permettono l'identificazione indiretta**, come un numero di identificazione (ad esempio: il codice fiscale, l'indirizzo IP, il numero di targa);
- **I dati rientranti in particolari categorie (GDPR, art. 9)**: si tratta dei dati c.d. "sensibili", cioè quelli che rivelano l'origine razziale o etnica, le convinzioni religiose, filosofiche, le opinioni politiche, l'appartenenza sindacale, i dati relativi alla salute o alla vita e all'orientamento sessuale, i dati genetici e biometrici;
- **I dati relativi a condanne penali e reati (GDPR, art.10)** si tratta dei dati c.d. "giudiziari", cioè quelli che possono rivelare l'esistenza di determinati provvedimenti giudiziari soggetti ad iscrizione nel casellario giudiziale (ad esempio, i provvedimenti penali di condanna definitivi, la liberazione condizionale, il divieto od obbligo di soggiorno, le misure alternative alla detenzione) o la qualità di imputato o di indagato.

# Il trattamento dei dati personali

---

Il trattamento di dati personali viene effettuato nel rispetto del **Regolamento (UE) 679/2016 (GDPR)**

Il **Decreto Legislativo 101/2018** adegua la normativa nazionale alle disposizioni del GDPR

**Qualsiasi operazione o insieme di operazioni, compiute con o senza l'ausilio di processi automatizzati e applicate a dati personali o insiemi di dati personali**

Ad esempio: la raccolta, la registrazione, l'organizzazione, la strutturazione, la conservazione, l'adattamento o la modifica, l'estrazione, la consultazione, l'uso, la comunicazione mediante trasmissione, diffusione o qualsiasi altra forma di messa a disposizione, il raffronto o l'interconnessione, la limitazione, la cancellazione o la distruzione (GDPR, art. 4, par. 1, punto 2)

I soggetti che procedono al trattamento di dati personali devono adottare **misure tecniche e organizzative per garantire un livello di sicurezza adeguato ai rischi per gli interessati**, tenendo conto dello stato dell'arte e dei costi di attuazione, nonché della natura, dell'oggetto, del contesto e delle finalità del trattamento, come anche del rischio di varia probabilità e gravità per i diritti e le libertà delle persone fisiche (GDPR, art. 32)

# I diversi ruoli nel trattamento di dati personali

---

**Interessato** (GDPR, art. 4, par. 1, p. 1) è la persona fisica alla quale si riferiscono i dati personali

**Titolare** (GDPR, art. 4, par. 1, p. 7) è la persona fisica, l'autorità pubblica, l'impresa, l'ente pubblico o privato, l'associazione, ecc., che adotta le decisioni sugli scopi e sulle modalità del trattamento

**Responsabile** (GDPR, art. 4, par. 1, p. 8) è la persona fisica o giuridica alla quale il titolare richiede di eseguire, **per suo conto**, specifici e definiti compiti di gestione e controllo del trattamento dei dati

Il DL 101/2018 introduce due ulteriori ruoli che sono individuati dal titolare (o dal responsabile) **nell'ambito del proprio assetto organizzativo e che operano sotto la propria autorità:**

**Soggetto designato** è la persona fisica a cui vengono attribuiti specifici compiti e funzioni connessi al trattamento di dati personali

**Incaricato del trattamento** è la persona fisica autorizzata al trattamento dei dati personali

# Data protection by default and by design

---

**Configurare il trattamento prevedendo fin dall'inizio le garanzie indispensabili** «al fine di soddisfare i requisiti» del GDPR e tutelare i diritti degli interessati, tenendo conto del **contesto complessivo** ove il trattamento si colloca e dei **rischi per i diritti e le libertà degli interessati**

Tutto questo deve avvenire **a monte**, prima di procedere al trattamento dei dati vero e proprio («sia al momento di determinare i mezzi del trattamento sia all'atto del trattamento stesso») e richiede, pertanto, un'analisi preventiva e un impegno applicativo da parte dei titolari che devono sostanziarsi in una serie di **attività specifiche e dimostrabili**

**Responsabilizzazione (accountability):** viene affidato ai titolari il compito di decidere autonomamente le modalità, le garanzie e i limiti del trattamento dei dati personali, tramite l'adozione di **comportamenti proattivi tali da dimostrare la concreta adozione di misure** per assicurare l'applicazione del GDPR

# Rischio inerente al trattamento

---

## Rischio di impatti negativi sulle libertà e i diritti degli interessati

Tali impatti dovranno essere analizzati attraverso un **apposito processo di valutazione** tenendo conto dei **rischi noti o evidenziabili** e delle **misure tecniche e organizzative** che il titolare ritiene di dover adottare per **mitigare tali rischi**

**Il titolare decide in autonomia, in base all'esito della valutazione**, se iniziare il trattamento (avendo adottato le misure idonee a mitigare sufficientemente il rischio) ovvero consultare il Garante Privacy per ottenere indicazioni su come gestire il rischio residuale

**Il Garante non ha il compito di "autorizzare" il trattamento**, bensì di **indicare le misure ulteriori** eventualmente da implementare a cura del titolare

# Adempimenti da parte del titolare

---

## ○ Registro dei trattamenti

- indispensabile per ogni valutazione e analisi del rischio
- deve essere esibito su richiesta al Garante

## ○ Misure di sicurezza

- devono garantire un livello di sicurezza adeguato al rischio
- la lista di misure (GDPR, art. 32, par. 1) è aperta e non esaustiva
- non sussistono obblighi generalizzati di adozione di misure "minime" di sicurezza

## ○ Notifica delle violazioni di dati personali

- entro 72 ore e comunque "senza ingiustificato ritardo", ma soltanto se il titolare ritiene probabile che da tale violazione derivino rischi per i diritti e le libertà degli interessati
- se la probabilità è elevata, si dovranno informare della violazione anche gli interessati

## ○ Designazione di un "responsabile della protezione dati"

- facilitare l'attuazione del regolamento da parte del titolare
- sensibilizzazione e formazione del personale



# Valutazione d'impatto della protezione dei dati (DPIA)

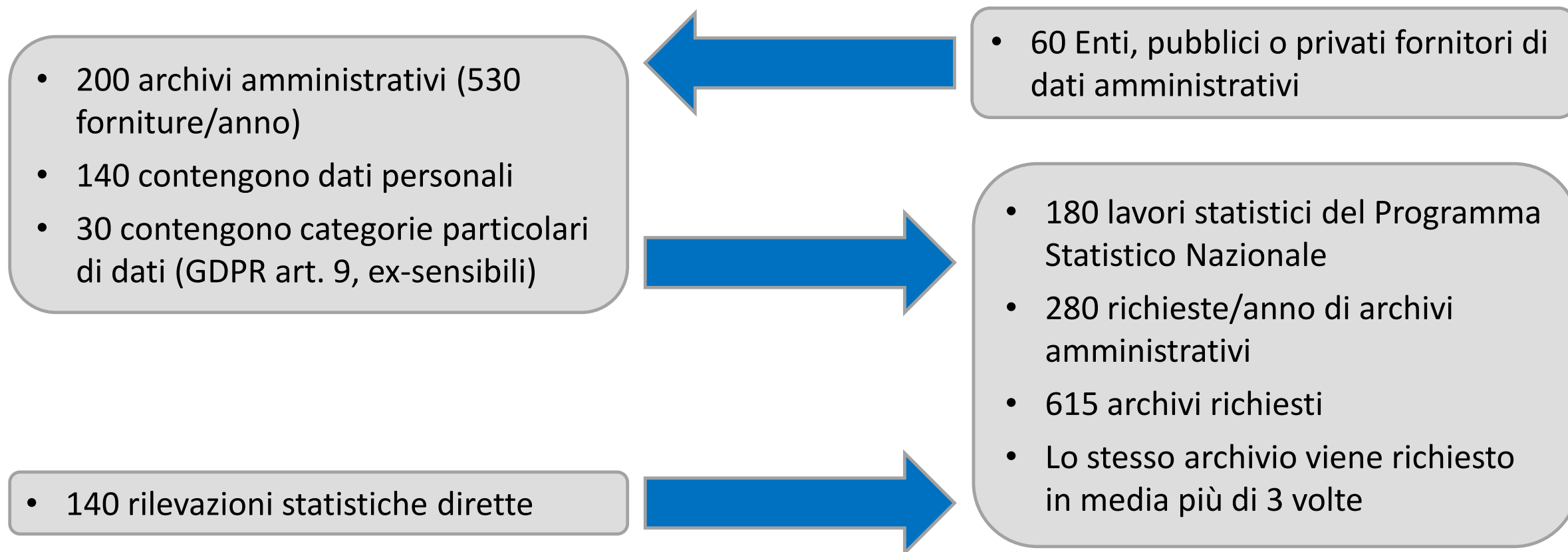
---

Quando un trattamento comporta un **rischio elevato** per i diritti e le libertà delle persone interessate il Regolamento **obbliga i titolari** a svolgere una valutazione di impatto prima di darvi inizio, consultando il Garante in caso le misure tecniche e organizzative da loro stessi individuate per mitigare l'impatto del trattamento non siano ritenute sufficienti, ovvero **quando il rischio residuale per i diritti e le libertà degli interessati resti elevato**

Rischio elevato:

- monitoraggio sistematico dei comportamenti degli interessati
- gran numero dei soggetti interessati
- trattamento di dati sensibili
- combinazione di questi e altri fattori

# I flussi di dati gestiti dalla Direzione Centrale della Raccolta Dati



Gli archivi con dati personali comprensivi di dati identificativi vengono integrati nel SIM - Sistema Integrato di Microdati da fonti amministrative e statistiche – attribuendo alle unità un codice anonimo (pseudonimo), univoco nel Sistema e stabile nel tempo

# Gli interventi del Garante su SIM 1

Provvedimento	Osservazioni
n. 324 del 26.06.2014 (PSN 2014)	«con riferimento al Sistema di integrazione logico-fisica di microdati amministrativi e statistici-SIM (identificato con il codice IST-02270) è stato ben descritto, nella sezione del prospetto identificativo recante le informazioni strutturali del lavoro statistico, che "il processo di integrazione ha come principale risultato: 1. l'identificazione di ogni oggetto (famiglia; individuo; unità economiche; loro relazioni) in fonti diverse con un numero ID univoco e stabile nel tempo che favorisce l'utilizzo, per ulteriori studi, dei dati individuali privi degli identificativi diretti conservando le potenzialità informative derivanti dal processo di integrazione. 2. La definizione, per ogni oggetto, delle relazioni logiche e fisiche (nel tempo e nello spazio) tra le informazioni disponibili nelle fonti diverse".
n. 566 del 29.10.2015 (PSN 2016) n. 87 del 02.03.17 (PSN 2017)	«... i trattamenti di dati personali previsti dall'integrazione del Sim ... , IST-02270 Sistema di integrazione logico-fisica di microdati amministrativi e statistici ..., potranno essere avviati solo a conclusione dell'apposita verifica preliminare ai sensi dell'art. 17 del Codice»

# Gli interventi del Garante su SIM 2

Provvedimento	Osservazioni
n. 271 del 09.05.2018 (PSN 2018)	<p>«...il lavoro IST-02270-Sistema di integrazione logico-fisica di microdati amministrativi (SIM), sul quale, come noto, in considerazione delle già rilevate criticità per i diritti e le libertà degli interessati, nel 2017, l'Autorità ha avviato, presso l'Istat, un apposito approfondimento istruttorio, tutt'ora in corso, finalizzato a verificare, in via preliminare, i rischi specifici derivanti dai trattamenti di dati personali ivi previsti, subordinandone l'effettivo utilizzo all'esito di tale verifica.</p> <p>Al riguardo, giova ribadire che nel SIM, contenente dati identificativi diretti, attraverso nuove acquisizioni e integrazioni di dati provenienti da fonti amministrative e statistiche, in costante incremento e in prospettiva diacronica, si determina una vera e propria schedatura permanente di ogni individuo, nel tempo e nello spazio, con gravi rischi per i diritti e le libertà degli interessati.»</p>
n. 10 del 23.01.2020 (Piano Generale del Censimento)	<p>«adottare idonee tecniche di pseudonimizzazione per garantire l'effettività dei principi di minimizzazione e di limitazione della conservazione»</p>

# Le osservazioni del Garante su SIM (Parere sul Piano Generale del Censimento)

---

- Il Garante riconosce l'**importanza dei dati di fonte amministrativa e della loro integrazione** per lo svolgimento dei lavori statistici in carico all'Istat, in quanto essi migliorano la rilevanza e la qualità dei risultati ottenuti e permettono di ridurre l'onere statistico sui rispondenti.
- Riconosce inoltre che **una gestione unitaria e organica** della raccolta, codifica e conservazione dei dati, rende **più razionale il funzionamento delle attività statistiche dell'Istituto**, anche nell'ottica dell'applicazione dei **principi di esattezza e accuratezza del dato**.
- Sottolinea alcune specifiche criticità in relazione alla effettiva attuazione del principio di **minimizzazione dei dati** trattati e di **limitazione della conservazione**.
  - L'assegnazione di un **codice invariante nelle diverse basi dati** non è in grado di offrire quella flessibilità necessaria a selezionare, di volta in volta, le **informazioni effettivamente pertinenti rispetto alla specifica finalità statistica perseguita**.
  - Il mantenimento di tale **codice univoco nel tempo** impedisce di **differenziare i tempi di conservazione dei dati in relazione alle diverse finalità statistiche perseguite**.

# Le indicazioni del Garante per l'evoluzione di SIM

---

- Ai fini della conformità al principio di minimizzazione dei dati occorre:
  - Assegnare **diversi codici pseudonimi in diverse basi di dati**, ciascuno con una validità limitata alla **specificata finalità perseguita**, in modo di realizzare un **disaccoppiamento logico** tra le stesse
  - Realizzare una **struttura gerarchica degli pseudonimi**, in modo di avere la possibilità di **ricongiungere i diversi pseudonimi** al medesimo interessato per conseguire una **nuova finalità**
- Ai fini della conformità al principio di limitazione della conservazione occorre:
  - Definire uno **specifico periodo di validità degli pseudonimi**, allo scadere del quale si può, in ragione delle esigenze statistiche, provvedere alla loro **rigenerazione** o alla **cancellazione dei codici e dei dati ad essi associati**

# La soluzione individuata dall'Istat per l'evoluzione di SIM

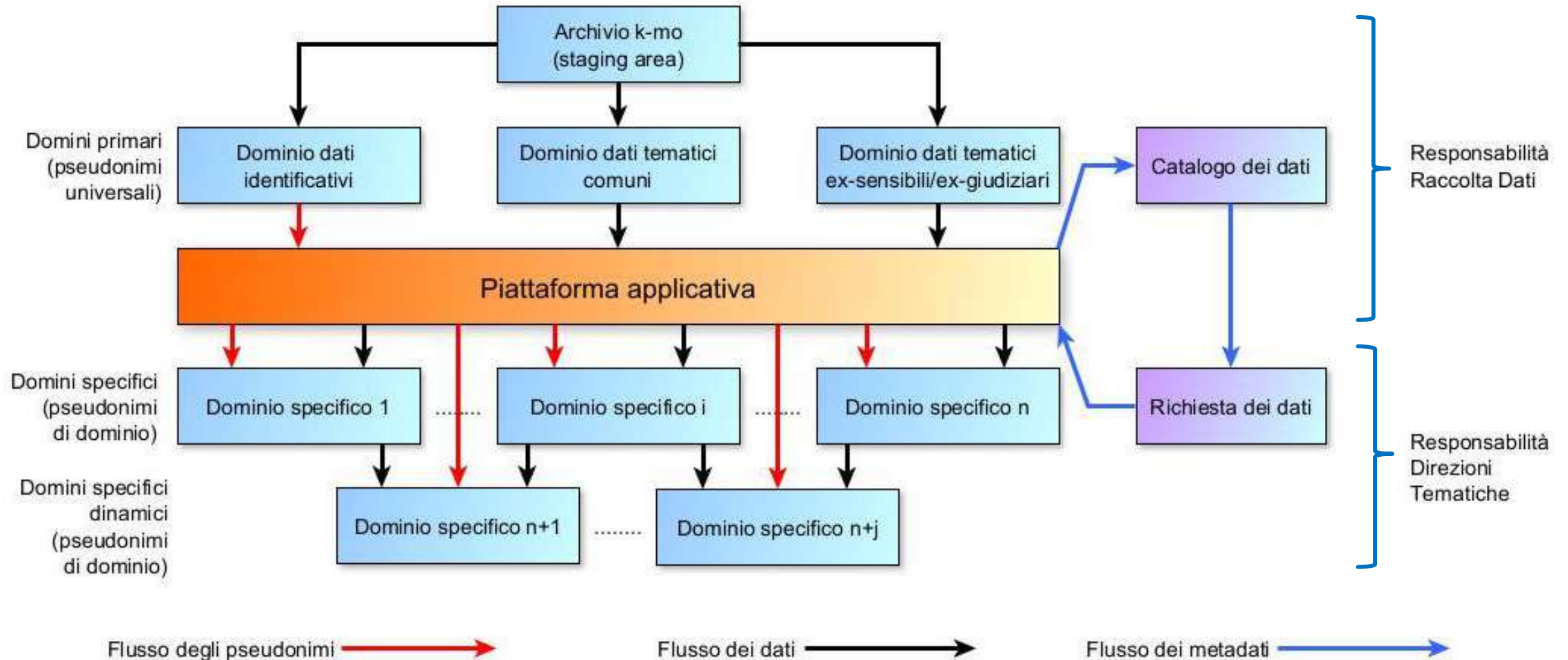
---

Le indicazioni del Garante sono state "tradotte" in **macro-requisiti** progettuali:

- Implementazione della **pseudonimizzazione secondo la logica dei domini specifici di integrazione**: in un dominio devono essere disponibili, per il tempo necessario e dotati di una specifica codifica degli pseudonimi, solo i dati pertinenti con il conseguimento di una determinata finalità statistica
- Introduzione di **nuove fasi** nel processo di trattamento dei dati:
  - **classificazione dei dati**, per indirizzare il trattamento a seconda delle diverse tipologie di archivi, di unità e di variabili e per la costruzione di un **Catalogo dei Dati** in cui selezionare i contenuti informativi pertinenti con le finalità statistiche da conseguire nei domini
  - **rilascio dei dati nei domini** specifici di integrazione
  - **rilascio di dati re-integrati nei domini dinamici** (output intermedi o finali di diversi domini specifici di integrazione, in particolare i Registri Statistici)
- **Adeguamento dei flussi documentali e maggiore integrazione con i flussi di trattamento dati**
- Implementazione di **indicatori e report** che consentano ai direttori «Tematici», in fase di richiesta dei dati, di **valutare il trade-off tra il rischio potenziale per gli interessati, la minimizzazione dei dati e la limitazione della conservazione** rispetto alle finalità statistiche



# La pseudonimizzazione secondo la logica dei domini specifici di integrazione





# I domini specifici di integrazione

---

Uno dominio specifico di integrazione viene generato dalla **piattaforma applicativa**, gestita dal Servizio DCRD/RDG, a seguito di **formale richiesta** da parte di una Direzione tematica, che ne diventa responsabile.

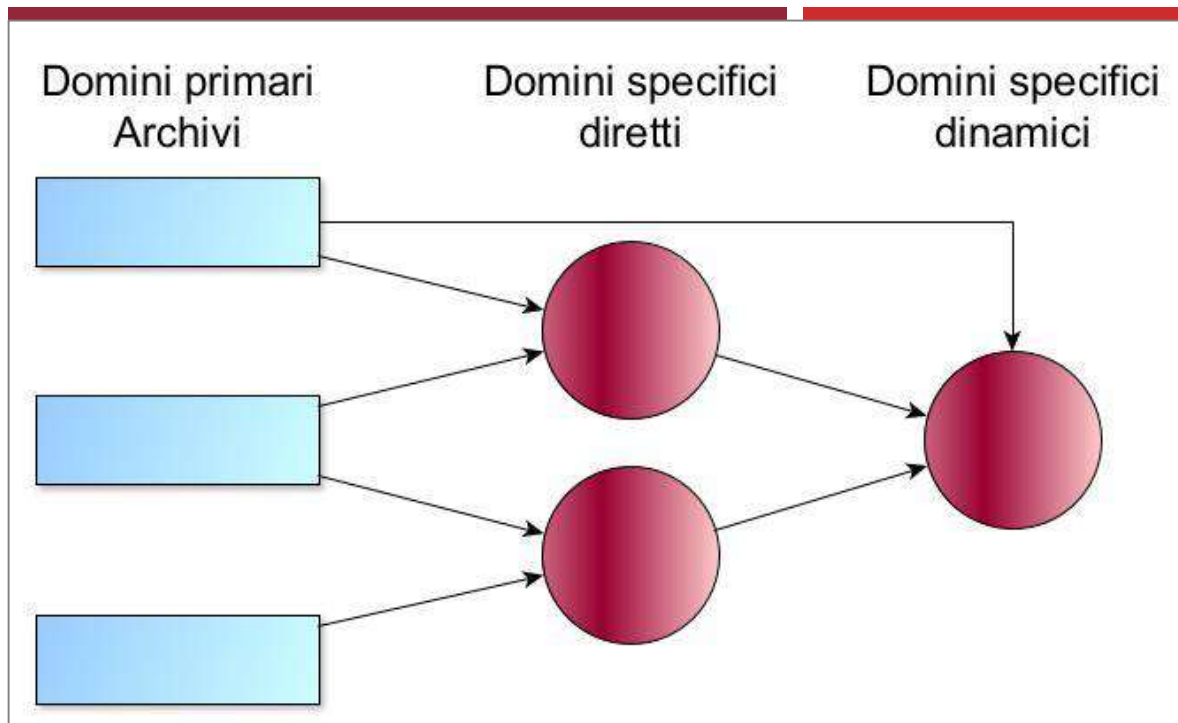
Gli elementi essenziali della richiesta sono:

- La finalità statistica associata al dominio
- I dati necessari, selezionati nel Catalogo dei Dati a livello di dataset, variabili e unità
- La durata di conservazione dei dati, necessaria per la realizzazione della finalità statistica
- Il personale incaricato del trattamento dei dati

Ogni unità inclusa in un dominio specifico di integrazione è caratterizzata da una **specifico codifica degli pseudonimi** che non consente l'integrazione con i dati contenuti in altri domini.

Al termine della **durata di conservazione**, definita nella richiesta, il dominio cessa di esistere, così come la sua specifica codifica degli pseudonimi.

# I domini specifici di integrazione dinamici



L'integrazione dei dati contenuti in diversi domini specifici è realizzata esclusivamente nell'ambito della piattaforma applicativa gestita dal Servizio DCRD/RDG ed è consentita dalla struttura gerarchica degli pseudonimi.

Le strutture responsabili della produzione di output intermedi o finali (in particolare i Registri Statistici), nel proprio dominio, mettono i dataset a disposizione della piattaforma e forniscono alla fase di classificazione i relativi metadati da inserire nel Catalogo dei Dati.

Le strutture tematiche interessate a tali dataset possono selezionarli, al pari delle fonti primarie, tramite nuove richieste, associate a nuove finalità statistiche, da realizzarsi in nuovi domini specifici di integrazione (domini dinamici), incrementando la qualità e l'efficienza dei processi di produzione statistica.

La piattaforma applicativa rende disponibili ai domini dinamici una specifico codifica degli pseudonimi.

# La classificazione dei dati

## Prima dell'acquisizione dei dati

- **Screening sintetico a livello di archivio:** *dati aggregati, micro-dati con o senza identificativi, tipologie di unità presenti*
- **Screening analitico a livello di singola variabile:** *identificativa, comune, sensibile, giudiziaria o di altro tipo correlato a un rischio specifico per gli interessati*
- Consente di **automatizzare la fase di caricamento dei dati** nei domini primari
- Fornisce gli elementi utili a **definire il rischio associato** ai trattamenti nei domini primari
- Permette di stabilire le **tipologie di trattamenti** da adottare allo scopo di mitigare i rischi:
  - Variabili identificative → pseudonimizzazione
  - Sensibili/giudiziarie → cifratura

## Prima del rilascio dei dati nei domini

- **Classificazione analitica delle unità** tramite link ai Registri Statistici: *persone fisiche, imprese senza personalità giuridica, imprese con personalità giuridica, istituzioni pubbliche, istituzioni non profit*
- **Descrizione analitica degli oggetti presenti nei domini primari** in termini di unità e variabili contenute
- Consente agli utenti tematici di **selezionare i dati necessari** per la specifica finalità statistica e di **automatizzare la fase di rilascio** dei dati nei domini specifici di integrazione
- Fornisce gli elementi utili a **definire il rischio associato** ai trattamenti nei domini specifici di integrazione

# Grazie

SILVANO VITALETTI | [vitalett@istat.it](mailto:vitalett@istat.it)