



**PROCEDURA PER IL MONITORAGGIO
DEL PROGRAMMA OPERATIVO NAZIONALE
INIZIATIVA OCCUPAZIONE GIOVANI
ALLEGATO 2
SICUREZZA E QUALITÀ DATI IN DB PAPL**



*Ministero del Lavoro
e delle Politiche Sociali*

Direzione Generale per le Politiche Attive, i
Servizi per il Lavoro e la Formazione



SOMMARIO

Introduzione	3
1. Flussi di alimentazione	4
2. Standard	4
3. Semantica dei dati	4
4. Validazione flussi e contenuto informativo	4



Introduzione

La Base dati delle Politiche Attive e Passive del Lavoro (BDPAPL), raccoglie i dati relativi ai giovani che hanno aderito al piano Garanzia Giovani (tramite la Scheda Anagrafica Professionale), il loro stato all'interno del progetto Garanzia Giovani (tramite le informazioni sull'Adesione, Stato Adesione), così come i dati relativi alle politiche, servizi e misure ad essi erogate (sempre tramite la Scheda Anagrafica Professionale, Flusso CO incentivate, etc.).

1. Flussi di alimentazione

La BDPAPL viene alimentata dai SIL Regionali, tramite cooperazione applicativa, o tramite il portale ClicLavoro Nazionale, tramite accesso puntuale dell'operatore, che inviano le informazioni al Nodo di Coordinamento Nazionale, che si occupa di validarli e centralizzarli: la costante e corretta alimentazione delle banche dati da parte dei SIL regionali è fondamentale al fine di garantire un corretto monitoraggio.

2. Standard

La correttezza e coerenza sintattica dei dati inviati al Nodo di Coordinamento Nazionale viene garantita col rispetto delle *Regole di Trasmissione* alle quali sono obbligati ad attenersi tutti i nodi periferici durante l'invio dei dati al Nodo di Coordinamento nazionale.

Tali regole sono formalizzate in *standard* tecnologici e comprendono : Accordi di Servizio, Standard Tecnici e Modalità tecniche tra il nodo centrale e i nodi periferici (cfr. Documentazione allegata relativa agli standard tecnici). Ciò formalizza il rispetto di standard determinati in relazione ai modelli ed alla struttura prevista per i vari moduli di informazione (SAP, Adesione, Stato Adesione, etc).

3. Semantica dei dati

Per quanto riguarda i valori ammessi per le varie informazioni all'interno dei moduli, sono stati definiti e condivisi delle *Classificazioni Standard*: per ogni parametro all'interno della SAP, dell'Adesione, etc, sono ammessi solo codici riconducibili alle classificazioni standard.

4. Validazione flussi e contenuto informativo

Per ogni tipo di comunicazione/modello dati al livello del Nodo di Coordinamento Nazionale sono stati implementati una serie di controlli e verifiche automatiche che permettono di stabilire se l'informazione inviata dal nodo periferico è coerente con gli standard e le regole, altrimenti viene rigettata e viene inviata alla Regione (o comunque al nodo periferico) mittente una notifica con la segnalazione del motivo del rigetto/scarto.

Le informazioni raccolte all'interno della BDPAPL attraverso l'esecuzione giornaliera di flussi e procedure automatizzate vengono archiviate e gestite all'interno di un Data

Warehouse (che al livello più alto alimenta direttamente il Sistema di Monitoraggio di Garanzia Giovani).

All'interno del Datawarehouse le informazioni della BDPAPL vengono aggiornate, integrate, elaborate, normalizzate e consolidate al fine di ottenere un sistema che garantisca accuratezza, pertinenza, rilevanza, trasparenza e tempestività dei dati.

In particolare il Datawarehouse del BDPAPL:

- Contiene dati coerenti e "ripuliti" da eventuali errori nei dati alla fonte, preservando la qualità stessa dei dati utilizzati per l'analisi ; è garantita la rispondenza dei dati alle norme e agli Standard definiti ;
- I dati esposti possono essere ricombinati e separati rispetto alle esigenze di analisi, è possibilità di produrre indicatori a differenti livelli di aggregazione (che, poi il sistema di raccolta dei dati è decentrato, coinvolgendo tutte le regioni, ha implicato l'omogeneità delle classificazioni adottate);
- E' possibile produrre indicatori per scopi differenti (individuazione delle criticità e delle anomalie, monitoraggio performance di sistema);
- Viene garantita la sicurezza dei dati: è possibile consultare i dati tramite la Dashboard di Monitoraggio alla quale si può accedere tramite utenze e password di dominio ministeriali. il Data Warehouse è un sistema in sola lettura, gli utenti non eseguono azioni di aggiunta, modifica o eliminazione dei dati.
- L'accesso ai dati è garantito con bassi tempi di attesa fra l'interrogazione dei dati e l'output di risultati.

Nel datawarehouse i dati sono presenti al massimo dettaglio e storicizzati raggiungendo dimensioni poco compatibili con l'interrogazione diretta attraverso i diversi tool di analisi a disposizione: sono stati perciò generati aggregati di dati (datamart) ai quali accedono direttamente le Dashboard di Monitoraggio. Per rispondere a richieste di date di dettaglio o puntuali

E' all'interno del DataWarehouse quindi che vengono effettuati una gran parte di controlli semantici sulla coerenza, 'pulizia' e affidabilità dei dati ricevuti che non possono essere effettuati automaticamente a livello di nodo nazionale:

- sia perché possono emergere dal confronto tra le stesse informazioni provenienti da diverse fonti;

- sia perché possono emergere dal confronto tra i dati diversi ricevuti in modo asincrono e quindi le verifiche possono essere effettuate in un momento successivo alla ricezione;

Ad esempio a livello di NCN ci sono controlli che non permettono di ricevere adesioni relative a soggetti con età non in fascia Garanzia Giovani al momento dell'adesione stessa, o inserire all'interno della sezione 6 della una politica attiva non prevista; ma è a livello di DataWarehouse che viene verificata la coerenza tra: le informazioni associate al Giovane all'interno della SAP; le informazioni sullo stato Adesione del Giovane; le informazioni sull'eventuale Presa in Carico del Giovane (e quindi del calcolo del profiling).

Gli indicatori definiti per il Sistema di Monitoraggio utilizzano solo dati corrispondenti ad informazioni coerenti con il flusso previsto e la normativa (ad esempio tra i Presi in Carico vengono conteggiati solo i giovani per i quali è presente la politica A02 all'interno della Sezione 6 della SAP e per i quali è stato effettuato il calcolo del Profiling coerentemente con la A02); gli altri dati vengono gestiti come scarti ed anomalie da verificare ed eventualmente sanare con l'intervento delle Regioni: il Data Warehouse è infatti un sistema in sola lettura, gli utenti non eseguono azioni di aggiunta, modifica o eliminazione dei dati.

La qualità del Data Warehouse può mettere in luce problematiche nascoste all'interno dei processi aziendali, tipicamente l'assenza di dati relativamente ad una particolare attività dovuta all'opzionalità di compilazione dei dati (causa principale dell'assenza di dati) può evidenziare, proprio grazie al Data Warehouse, la necessità di verificare a livello di processo direttamente con le Regioni la possibilità/necessità di rendere obbligatorie informazioni prima opzionali piuttosto che gestire in modo diverso alcune informazioni (es. Addendum al vademecum).