

## Allegato 1a all'avviso

### **OGGETTO: Servizi di data management e integration volti alla implementazione e integrazione del Sistema Informativo “Le scuole di INDIRE” – Descrizione servizio**

Struttura di Ricerca n. 10 “Modelli e metodologie per la documentazione, l’analisi e la lettura dei principali fenomeni del sistema scolastico” CUP B59C21000400001  
Agenzia Erasmus +: CUP- B59H21000000007

#### **PREMESSA**

Dal 2019 la struttura di ricerca “Modelli e metodologie per l’analisi, la lettura e la documentazione dei principali fenomeni del sistema scolastico” è impegnata nella definizione del sistema informativo “Le Scuole di INDIRE” che fornisce una rappresentazione complessiva geolocalizzata delle scuole con cui l’Istituto interagisce per varie finalità (sperimentazione di soluzioni innovative, formazione, monitoraggio, video-documentazione ecc.).

Una parte fondamentale di queste interazioni riguarda la dimensione dell’internazionalizzazione, ovvero la partecipazione delle scuole a progetti di mobilità, formazione, collaborazione, confronto internazionale, anche oltre l’ambito europeo, gestiti dall’Agenzia Erasmus.

Questa rappresentazione complessiva consente:

1. di esplicitare e mettere a fattore comune un’informazione che altrimenti rimarrebbe circoscritta e sottesa alle singole esperienze di ricerca;
2. di restituire una visione d’insieme che può essere funzionale a comprendere meglio nel loro complesso i processi di supporto al miglioramento e all’innovazione che INDIRE promuove nel sistema scolastico nazionale;
3. di costruire una profilazione più ricca delle scuole con cui INDIRE si relaziona e di individuare cluster rappresentativi dei modelli comportamentali che evidenziano i diversi processi di trasformazione in atto.
4. di sviluppare un Osservatorio a carattere permanente su alcuni temi strategici della scuola (ad esempio DDI, educazione civica) anche in relazione a temi sociali (inclusione, educazione di genere, educazione al rispetto, contrasto al bullismo al cyberbullismo e al razzismo, studio e analisi di situazioni di povertà educativa)
5. di sviluppare reti di relazioni tra le “Scuole di INDIRE” attraverso il web semantico per affinità di temi, interessi, progetti, partenariati.

6. di descrivere e studiare le variabili indipendenti, contestuali e intervenienti che caratterizzano i comportamenti identificati al fine di individuare possibili indicatori di processi di trasformazione.

Attraverso un processo incrementale di integrazione dei dati quali-quantitativi, strutturati e non strutturati, che sono conservati nelle diverse banche dati progettuali dell'Ente, la ricerca ha sviluppato una risorsa informativa funzionale allo stesso tempo per la documentazione e lo studio dei processi di trasformazione, miglioramento, innovazione in atto nelle scuole, che sono sostenuti dalle attività di ricerca di Indire e dell'Agenzia Erasmus. Gli interventi di potenziamento tecnologico di data integration consentono di estrarre valore informativo dai dati depositati nei database dell'Ente per elaborare, a partire da questi, nuova conoscenza da mettere a disposizione dell'utente sia esso la comunità educativa, quella scientifica e lo stesso policy maker.

Ad oggi il progetto ha visto il censimento di 37 progetti promossi e gestiti da INDIRE a partire dall'all'a.s. 2015/16 e l'integrazione dei rispettivi DB progettuali in un unico datawarehouse. Questo ha permesso di identificare un numero complessivo di 6.106 scuole statali (corrispondente al 69,6% della rete scolastica nazionale) con cui l'Istituto ha avuto interazioni negli ultimi anni.

La crescente complessità che caratterizza il settore educativo rafforza l'esigenza di osservare i fenomeni della scuola da prospettive sempre più ampie e articolate, attingendo a un patrimonio informativo eterogeneo e in continua evoluzione, che ne consenta lo studio anche in relazione alle interconnessioni con il contesto di riferimento.

In particolare per quanto riguarda la partecipazione delle scuole al Programma Erasmus + sarà rilevante comprendere la distribuzione territoriale delle scuole coinvolte, la capacità di presentare progettualità che vengano approvate per arrivare a identificare scuole particolarmente virtuose che potrebbero avere la funzione di "ambasciatrici" dell'innovazione e dell'internazionalizzazione sul territorio.

Ai fini di quanto sopra, INDIRE intende avvalersi di un servizio specialistico esterno per potenziare la capacità del sistema informativo integrando i dati disponibili con linked data provenienti da fonti informative esterne all'ente.

Rileviamo infatti che non sono presenti o disponibili all'interno dell'istituto professionalità, risorse e strumenti che possano realizzare quanto di seguito specificato.

#### **OGGETTO DEL SERVIZIO**

Nello specifico, l'oggetto del servizio da affidare a ditta esterna dovrà comprendere le seguenti attività:

##### **1) *Supportare l'arricchimento dell'archivio con le informazioni pubblicate sul Portale unico dei Dati della Scuola***

Il primo arricchimento riguarda i dati open pubblicati su Portale unico dei Dati della Scuola in particolare relativamente alle seguenti quattro aree tematiche:

- Studenti
- Personale
- Sistema Nazionale di Valutazione
- Edilizia scolastica

Si tratta di operare un'estrazione dei dati indicati per l'intera rete scolastica nazionale e di predisporre l'integrazione nel datawarehouse denominato LE SCUOLE DI INDIRE già implementato dalla struttura di ricerca "Modelli e metodologie per l'analisi, la lettura e la documentazione dei principali fenomeni del sistema scolastico" è impegnata nella definizione del sistema informativo "Le Scuole di INDIRE".

## **2) Supportare l'arricchimento dell'archivio LE SCUOLE DI INDIRE con informazioni statistiche di contesto e approccio alla costruzione di Layer informativi di riferimento**

Si prevede uno step iniziale con l'obiettivo di integrare le principali informazioni di contesto demografico e socio economico, in particolare in relazione alle seguenti aree tematiche:

- Popolazione: caratteristiche della popolazione residente, occupazione, mobilità,
- Servizi: luoghi di interesse culturale, di intrattenimento, poli fieristici, etc
- Economia: caratteristiche delle imprese, ricchezza delle famiglie

Per le informazioni a livello di sezione di censimento si ipotizza un arricchimento che vede per ogni scuola/polo:

- Dato sulla sezione di pertinenza
- Dato sull'insieme di sezioni che ricadano in un raggio di gravitazione intorno alla scuola/plesso.

Questo approccio permetterà:

- ✓ Di creare nuove dimensioni informative "INDIRE" su tematiche specifiche;
- ✓ Analizzare più dimensioni congiuntamente appartenenti a diverse dimensioni di analisi; in particolare selezionando un punto nel territorio (es. scuola/plesso) dovrebbe essere possibile estrarre tutte le informazioni ad esso collegate relativi ad altri layer informativi di interesse;
- ✓ Sviluppare modelli statistici territoriali su concentrazioni/ diffusi di fenomeni sul territorio

Si tratta di operare un'estrazione dei dati indicati per l'intera rete scolastica nazionale e di predisporre l'integrazione nel datawarehouse LE SCUOLE DI INDIRE già implementato.

## **3) Arricchimento dell'archivio con i dati della rete nazionale delle scuole paritarie**

Con riferimento all'elenco delle scuole paritarie reperibile nella sezione opendata del MIUR all'indirizzo <https://dati.istruzione.it/opendata/opendata/catalogo/elements1/?area=Scuole>, si richiedono le seguenti operazioni

- Sistemizzazione delle informazioni disponibili in un database MySql con un modello E/R che documenti la persistenza di Istituti di Riferimento e Sedi di erogazione del servizio (Plessi) dall'AS 2015/2016 all'AS 2020/2021;
- La geolocalizzazione delle sedi di erogazione del servizio e degli istituti di riferimento;

- L'integrazione delle informazioni relative alla popolazione scolastica in ciascun anno di osservazione e più in generale di tutte le informazioni reperibili nel catalogo opendata del MIUR <https://dati.istruzione.it/opendata/opendata/catalogo/#Scuola> (es. Indirizzi per la scuola secondaria di II grado);
- Integrazione di questo database con tutte le informazioni ritenute rilevanti come illustrato nel precedente punto 2.

Si tratta di operare un'estrazione dei dati indicati per l'intera rete scolastica nazionale delle scuole paritarie, di integrarla con i dati richiesti, se mancanti, e di predisporre l'integrazione nel datawarehouse LE SCUOLE DI INDIRE già implementato.

#### **4) *Profilazione docente INDIRE***

In riferimento all'anagrafica dei docenti che collaborano a vario titolo con Indire (partecipano alle sperimentazioni/ alle formazioni previste dai progetti INDIRE) presente nei vari DB progettuali e nella 'global anagrafe', si richiede un'analisi della struttura attuale del DB anagrafica docenti per la standardizzazione dei dati in entrata e individuazione record doppi o duplicati.

Si richiede inoltre arricchimento e integrazione con dati mancanti (istituzione scolastica di afferenza, altri dati legati all'anagrafica professionale).

Si tratta di effettuare un'operazione di data quality sulle estrazioni fornite e di completarle con i dati mancanti, laddove necessario, quindi predisporre l'integrazione nel datawarehouse LE SCUOLE DI INDIRE già implementato.

Si specifica che per ciascuno dei punti indicati, le attività richieste comprendono, oltre alla realizzazione del data merge e data integration richiesto, anche un servizio di consulenza complessivo che riguarda sia l'impostazione che l'espletamento delle diverse azioni di configurazione, integrazione e implementazione dell'archivio.