

La migrazione dei dati da ASFi a SIASFi: problemi e soluzioni

Diana Spano

Col termine “migrazione” si intende il passaggio dei dati da un sistema di partenza ad un altro, in cui le modalità di memorizzazione e le strutture dati non necessariamente rispecchiano quelle originariamente prodotte per la raccolta degli stessi.

Le differenti caratteristiche strutturali dei sistemi coinvolti rendono necessario stabilire “precise regole” che consentano di conservare tutto l’insieme dei dati originali.

Durante le fasi di migrazione è possibile, e talvolta necessario, integrare i dati di partenza attraverso un loro arricchimento mediante informazioni aggiuntive, normalmente dedotte dai primi mediante l’ausilio di meccanismi indicati da specialisti del tipo di informazione da processare (nel nostro caso, l’esperto è sempre un archivista).

Nel caso specifico, l’attività di migrazione ha avuto come obiettivo il recupero, nel sistema SIASFi, dei dati contenuti nel database ASFi.

Il recupero dei dati è avvenuto interamente via software utilizzando il linguaggio PL-SQL di Oracle. Il recupero dei dati è stato completo, anzi, in alcuni casi (come nel recupero del soggetto versante nel complesso archivistico) sono state recuperate in maniera semiautomatica relazioni tra le informazioni non esistenti nel sistema di origine ma molto chiare nella mente degli archivisti specialisti nel conoscere i dati presenti nel sistema ASFi e il modello dei dati del SIASFi.

L’operazione complessiva non è stata particolarmente complessa grazie alla omogeneità dei DBMS coinvolti (entrambi Oracle) e all’alta compatibilità tra i rispettivi schemi.

Sono state recuperate le informazioni relative a:

- Complessi archivistici
- Soggetti produttori
- Soggetti conservatori
- Contesti politici
- Strumenti di ricerca
- Bibliografia

- Complilatori
- Relazioni tra le informazioni recuperate, elencate nei punti precedenti
- Relazioni non modellate e quindi non presenti nel sistema d'origine (ASFi) ma modellate nello schema di destinazione (SIASFi). Il loro recupero è stato possibile grazie alla conoscenza che gli archivisti hanno dei dati.

Le difficoltà incontrate nel processo di migrazione, sono legate alla risoluzione dei conflitti tra i due schemi coinvolti:

- Conflitti sulla struttura delle tabelle, in particolare, conflitti sui vincoli delle tabelle.

Sono stati risolti in modo semiautomatico con il coinvolgimento degli archivisti in quanto conoscitori del dominio. In questa categoria rientrano i problemi derivanti dalla valorizzazione dei campi obbligatori in SIASFi ma non in ASFi. Un esempio è rappresentato dalla migrazione dei soggetti produttori: in SIASFi i soggetti produttori sono partizionati in 3 insiemi disgiunti a seconda della loro tipologia. La tipologia è una informazione obbligatoria e non è previsto un valore di default per tale campo, nel senso che, non è possibile inserire nel database un soggetto produttore se non se ne indica la tipologia di appartenenza. Al contrario, in ASFi, il vincolo di obbligatorietà non era previsto e questo ha reso impossibile una migrazione automatica di 2945 soggetti produttori aventi la tipologia non valorizzata.

Per poter completare il recupero di questi dati in modo rapido, è stato necessario estrarli dal database di origine e sottoporli agli archivisti in modo che potessero indicare, per ogni record, l'informazione mancante e procedere poi all'inserimento degli stessi in SIASFi con l'ausilio di un programma specifico.

Per individuare meglio il problema, bisogna sapere che le tre tipologie ammesse in SIASFi sono: ente, persona e famiglia. Quando un soggetto produttore viene inserito nel database, le informazioni non vengono inserite in un'unica tabella ma in due tabelle distinte. La prima è la tabella dei soggetti produttori che contiene le informazioni comuni a tutti i soggetti produttori, la seconda è una tabella specifica che viene determinata in automatico dal sistema in base alla tipologia specificata al momento dell'inserimento e ha lo scopo di memorizzare le informazioni che caratterizzano la specifica tipologia di soggetto produttore.

Il fatto che un soggetto produttore debba necessariamente essere inserito in una tabella specifica in base alla sua tipologia non ha permesso, in questo contesto, l'utilizzo di un valore di default per poter effettuare la migrazione dei dati in modo automatico.

SOGGETTI	
Ch.pr.	PKENTE
Ch.pr.	PKSOGG
I1	CODICE
I2	TIPOSOGGETTO
	DENOMINAZIONE
	FONTIDENOMINAZIONE
	NOTIZIESTORICHE
	FONTI

ASFi: tabella Soggetti

SPROD_SOGGETTOPRODUTTORE	
Ch.pr.	PKIDSPRODSOGGETTO
	TIPOSOGGETTO

SPROD_ENTE	
Ch.pr.	PKIDSPRODENTE
	DENOMINAZIONEAUTORITA
	LUOGO
	NOTESTORICOGIURIDICHE
	FONTI
	CODICEANAGRAFE

SPROD_FAMIGLIA	
Ch.pr.	PKIDSPRODFAMIGLIA
	DENOMINAZIONE
	TITOLI
	DATAESTREMOMOTO
	SECOLOESTREMOMOTO
	SPECIFICAESTREMOMOTO
	CODIFICAESTREMOMOTO
	VALIDITAESTREMOMOTO
	DATAESTREMORECENTE
	SECOLOESTREMORECENTE
	SPECIFICAESTREMORECENTE
	CODIFICAESTREMORECENTE
	VALIDITAESTREMORECENTE
	VERSIONEIMMAGINEALBEROGEN
	VERSIONETESTOALBEROGEN
	NOTE
	FONTI
	CODICEANAGRAFE

SPROD_PERSONA	
Ch.pr.	PKIDSPRODPERSONA
	COGNOME
	NOME
	PATRONIMICO
	PROVENIENZA
	SOPRANNOME
	LUOGONASCITA
	LUOGOMORTE
	NOTEBIOGRAFICHE
	FONTI
	CODICEANAGRAFE
	VALIDITADATANASCITA
	VALIDITADATAMORTE
	DATAESTREMOMOTO
	DATAESTREMORECENTE
	CODIFICAESTREMOMOTO
	CODIFICAESTREMORECENTE
	VALIDITAESTREMOMOTO
	VALIDITAESTREMORECENTE
	SECOLOESTREMOMOTO
	SECOLOESTREMORECENTE
	SPECIFICAESTREMOMOTO
	SPECIFICAESTREMORECENTE

SIASFi: Soggetti produttori Enti, Persone e Famiglie

- Conflitti relativi alla valorizzazione dei campi obbligatori in SIASFi con inserimento di valori di default nel caso in cui il dato non fosse presente in ASFi.

Questo problema è stato riscontrato in vari casi. Uno per tutti è rappresentato dalla migrazione dei complessi archivistici, dove, per alcuni record in ASFi, non era stata valorizzata la denominazione in quanto non era stato imposto su questo campo il vincolo di obbligatorietà. In

SIASFi invece tale vincolo esiste e di conseguenza va rispettato. I dati sono stati recuperati inserendo nelle denominazioni non valorizzate la stringa 'NON_SPECIFICATO', lasciando agli archivisti il compito di inserire le denominazioni corrette nella fase di bonifica dei dati successiva alla migrazione.

- Conflitti relativi a dati aventi come origine una tabella e come destinazione più tabelle.

Un esempio è dato dalla migrazione dei compilatori. Nella tabella ASFi SYS_AA_SCHEDE, insieme alle informazioni sull'intervento fatto dal compilatore è presente anche una codifica del nome della scheda sulla quale si è intervenuti. Nel SIASFi invece, esiste una tabella di relazione tra il compilatore e ogni possibile scheda oggetto di compilazione.

In base a questo attributo (TIPOSCHEDE) è stato possibile smistare i dati contenuti nella tabella SYS_AA_SCHEDE nelle opportune tabelle del SIASFi.

L'oggetto al quale fa riferimento la scheda, viene stabilito in base al valore dell'attributo TIPOSCHEDE nel seguente modo:

TIPOSCHEDE = 'AG', la scheda è in relazione con il Complesso archivistico

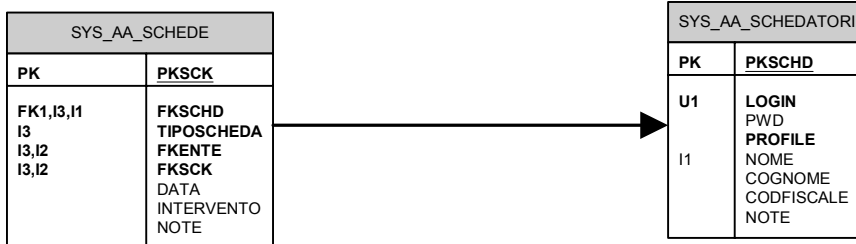
TIPOSCHEDE = 'SG', la scheda è in relazione con il Soggetto produttore

TIPOSCHEDE = 'CT', la scheda è in relazione con il Contesto politico

TIPOSCHEDE = 'AM', la scheda è in relazione con gli Ambiti territoriali

TIPOSCHEDE = 'ST', la scheda è in relazione con gli Strumenti di corredo

TIPOSCHEDE = 'BB', la scheda è in relazione con la Bibliografia

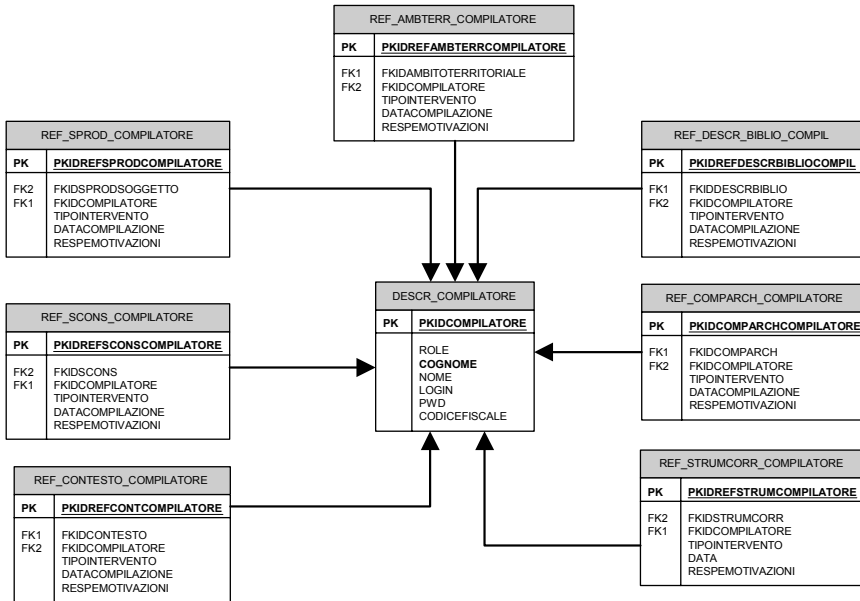


Compilatori (ASFi)

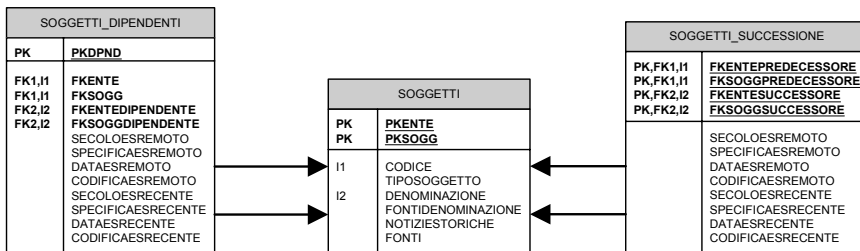
- Conflitti relativi al recupero delle informazioni modellate con l'utilizzo di più di una tabella in ASFi e modellate in una tabella in SIASFi.

Anche in questo caso, un esempio è dato delle relazioni tra i soggetti produttori.

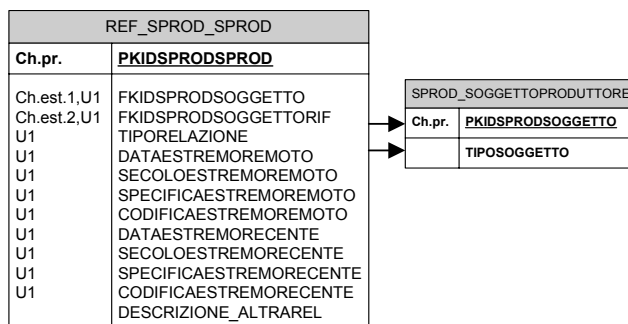
In alcuni casi, come per i Profili istituzionali, sono stati recuperati dati e relazioni non modellate in ASFi. I dati provengono dalle tabelle SOGGETTI e SOGGETTI_COLLEGATI, in quanto uno degli enti in ASFi è in realtà un



SIASFi: Sottoschema Compilatori

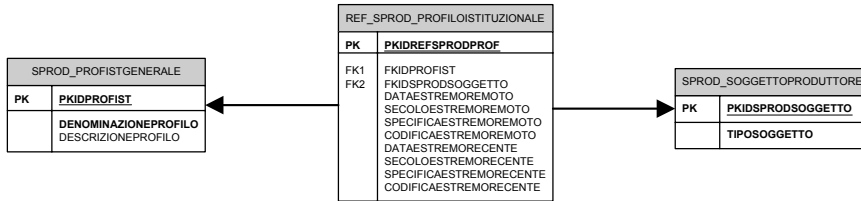


ASFi: Sottoschema per la modellazione delle relazioni tra soggetti produttori



SIASFi: Sottoschema per la modellazione delle relazioni tra soggetti produttori

profilo istituzionale. Operativamente, il profilo istituzionale contenuto nella tabella dei soggetti non viene migrato insieme agli altri enti ma viene inserito nella tabella `SPROD_PROFISTGENERALE`. Poiché otto soggetti produttori della tabella `soggetti` possono essere relazionati col profilo istituzionale creato, vengono create otto relazioni nella tabella `REF_SPROD_PROFILOISTITUZIONALE` che collega il profilo istituzionale con i produttori individuati.



Dopo la migrazione, dietro richiesta degli archivisti che lavorano alla bonifica dei dati, è stata effettuata una serie di interventi sui dati utilizzando SQL per velocizzare alcune delle operazioni di bonifica. Gli archivisti hanno specificato il tipo di intervento e gli esatti valori da assegnare ai dati in SIASFI. Un esempio è dato dalla modifica dei valori a vocabolario controllato presenti nel database.

Come detto precedentemente, una volta stabilite le corrispondenze tra i nomi delle strutture dati nei due database coinvolti nel processo di migrazione (ASF_i e SIASF_i), il recupero dei dati non è stato particolarmente complicato in quanto, limitatamente alle informazioni gestite in entrambi i sistemi, non vi era una grossissima differenza nelle modellizzazioni delle stesse.

