

# **SIQual: il sistema informativo sulla qualità per gli utenti esterni**

*A cura di Giovanna Brancato, Riccardo Carbini e Concetta Pellegrini*

*ISTAT – Istituto Nazionale di Statistica  
Servizio Metodologie, Tecnologie e Software per la produzione dell'Informazione  
statistica (MTS)*

Attribuzioni:

*I capitoli 1 e 2 sono da attribuire a G. Brancato*

*Il capitolo 3 è da attribuire a C. Pellegrini*

*Il capitolo 4 è da attribuire a R. Carbini*

*Il capitolo 5 è da attribuire ai tre autori*

## Riassunto

In questo documento si riporta l'attività di un progetto di studio e realizzazione di un sistema di navigazione sui metadati e sulla qualità dei processi produttivi statistici dell'Istituto descrivendo le principali tappe che hanno portato allo sviluppo del sistema SIQual.

SIQual è un sistema di navigazione di metadati di processo e di documentazione sulla qualità orientato agli utenti del sito web dell'Istat. L'alimentazione del sistema deriva dal sistema di documentazione delle indagini, SIDI, in uso presso l'Istituto. Nel progettare il sistema SIQual particolare attenzione è stata posta verso l'orientamento agli utenti, la facilità di navigazione, il rispetto degli standard grafici dell'Istituto, un architettura funzionale con i principali browser a disposizione degli utenti e l'ottimizzazione dei tempi di risposta del sistema.

Gli obiettivi iniziali del gruppo di lavoro, di offrire agli utenti del portale le tavole dell'Annuario Statistico Italiano corredate dalle informazioni sui processi produttivi che hanno generato i dati, sono stati rivisti e nel corso dei lavori del gruppo si è optato per una impostazione differente, ma più idonea a tale sistema, e cioè quella di metterlo in relazione con altri sistemi informativi di dati presenti sul sito dell'Istituto.

Il sistema SIQual è uno strumento utile per gli utenti esterni che siano interessati a comprendere gli aspetti più salienti del processo produttivo statistico che genera i dati, nonché gli aspetti sulla qualità. La documentazione sulla qualità è presente sia sotto forma di attività di prevenzione, controllo in corso d'opera e valutazione dell'errore non campionario, sia attraverso la produzione di documenti, con un diverso grado di standardizzazione e approfondimento, sulla qualità (note metodologiche, schede standard sulla qualità, altri documenti sulla qualità).

Il sistema è uno strumento che contribuisce al miglioramento della qualità perché risponde principalmente alle esigenze di trasparenza nei confronti degli utenti.

## Capitolo 1

---

### Introduzione

La trasparenza nei confronti degli utenti è una delle dimensioni della qualità. Questa si può attuare a diversi livelli (nel prodotto, nel processo, nella diffusione). Laddove esistano dei metadati rilevanti per gli utenti sull'informazione statistica prodotta, sulle procedure utilizzate e sugli strumenti messi in atto per la prevenzione, il controllo in corso d'opera e la valutazione della qualità, questi dovrebbero essere messi a loro disposizione. L'obiettivo è quello di rendere gli utilizzatori dei dati maggiormente consapevoli delle qualità e delle limitazioni dei dati stessi, per consentirne un uso corretto e quindi migliorarne anche la rilevanza.

Queste considerazioni, hanno portato a studiare le modalità per attuare un maggiore sfruttamento del patrimonio informativo in possesso dell'Istituto, sia di processo che sulla qualità, che costituiscono un insieme di informazioni utili a corredo della produzione statistica.

Per tale motivo, al lavoro di informatizzazione dell'acquisizione delle note metodologiche dell'Annuario Statistico Italiano - note che sono rivolte agli utenti esterni - e alla loro integrazione con il Sistema Informativo di Documentazione delle Indagini, denominato SIDI, si è fatto seguire un progetto più ampio con l'obiettivo primario di offrire agli utenti esterni un insieme maggiore e più organico di informazioni.

L'obiettivo iniziale del lavoro era quello di portare sul sito Internet un sottoinsieme dei metadati e le informazioni sulla qualità contenuti nel sistema SIDI insieme ai dati rappresentati dalle tavole contenute nelle pubblicazioni trasversali, a partire dall'Annuario Statistico Italiano (ASI). A tal fine, erano state identificate le seguenti attività:

- acquisizione delle tavole ASI;
- selezione e adeguamento delle informazioni di SIDI in informazioni adatte all'utente esterno e conseguente organizzazione del database;
- progettazione dell'interfaccia del sistema;
- sviluppo informatico.

Nel corso degli avanzamenti del lavoro, in considerazione di altre attività parallele che si sono venute a sviluppare e, in particolare la creazione di un sistema di acquisizione delle tavole dell'Annuario, denominato DB-OUT da una parte e l'estensione del sistema informativo generalizzato di diffusione dei dati statistici dell'Istat, denominato ISTAR, dall'altra, gli obiettivi del lavoro sono stati riadattati al mutato contesto. Considerando l'obiettivo di diffusione sul sito web del sistema SIQual, è stato deciso che questo dovesse avere come unico legame ai dati il link ai sistemi informativi di dati presenti nel portale, e non alle tavole dell'Annuario.

Nel presente documento si riporta l'attività sviluppata per arrivare a realizzare il sistema SIQual e in particolare verranno riassunte le analisi che hanno condotto a determinate scelte nel sistema (capitolo 2), gli aspetti di progettazione e sviluppo informatico (capitolo 3) e l'esperienza del sistema di acquisizione delle tavole (capitolo 4). Infine vengono tratte le conclusioni e date indicazioni per la prosecuzione del lavoro (capitolo 5).

Parte delle conclusioni tratte in questo documento, derivano dalle attività di un gruppo di lavoro interno all'Istituto.

## Capitolo 2

---

### Presupposti e scelte adottate nel sistema SIQual

Una prima considerazione nella progettazione del sistema ha riguardato l'utenza a cui il sistema stesso si rivolge. Si può ritenere che l'utenza del sito web dell'istituto sia un'utenza con un certo livello di esperienza e conoscenza dei dati statistici e dei processi che portano alla loro formazione. A tale proposito si è ritenuto che tale utenza potesse essere interessata agli aspetti di qualità sottostanti i processi produttivi dei dati e i dati stessi. Nello stesso tempo, la diffusione di informazioni riguardanti la qualità non può prescindere da quella dei metadati sul processo produttivo. Considerando il patrimonio informativo sui metadati e sulla qualità rappresentato da SIDI, era però necessario procedere ad un'analisi della idoneità di tale contenuto.

Inoltre, anche per le funzionalità che dovevano essere messe a disposizione degli utenti esterni in questo nuovo sistema informativo, ci si è basati su quanto già sviluppato per gli utenti interni. Infatti è da tempo disponibile sulla Intranet per gli utenti interni all'Istat un sistema di navigazione, denominato SIDI-TOP, che oltre ai metadati sul contenuto informativo e sul processo, documenta anche indicatori standard di qualità. A partire dalla analisi di questo sistema, e considerando la diversa utenza e i diversi motivi tecnici che saranno illustrati nel paragrafo 3, che imponevano comunque cambiamenti della base dati, è stato sviluppato SIQual.

## **2.1 Contenuto informativo del sistema SIQual**

L'oggetto principale del sistema è rappresentato, così come in SIDI, dai processi produttivi statistici dell'Istituto, identificabili in rilevazioni ed elaborazioni, così come definite dal Comstat.

Relativamente al contenuto da fornire agli utenti esterni, si è valutato che:

- non tutte le informazioni presenti nel sistema SIDI dovessero essere incluse in SIQual
- alcuni tesauri di SIDI, che si presentavano molto tecnici, dovessero essere "semplificati"

Un tesoro è una lista di voci standard di supporto alla documentazione dei concetti o delle varie operazioni svolte durante il processo. Per esempio, in SIDI, per documentare che una indagine è di tipo postale, esiste una voce specifica utilizzabile da tutte le indagini postali. Questo meccanismo permette di non avere una documentazione disomogenea e soggettiva, ma di avere una documentazione standard e precisa. Bisogna aggiungere che il tesoro non è uno strumento rigido, ma è ampliabile con voci nuove qualora sia necessario, confermate sempre attraverso un sistema di validazione che assicuri che il tesoro sia completo e non ridondante.

In SIQual, non sono state caricate le informazioni di SIDI maggiormente di servizio (direzione, servizio e unità operativa). Relativamente alle unità di analisi, che in SIDI sono classificate per tipologia (unità principali, di sommario, aggregate e di subset), si è ritenuto che tale classificazione fosse troppo dettagliata per gli utenti esterni, e quindi di scarso interesse, pertanto non è stata mantenuta, riportando il solo elenco delle unità di analisi senza distinzione per tipologia.

La descrizione del disegno di campionamento è rimasta invariata rispetto a SIDI. C'è da dire però che, rispetto ad una versione iniziale del sistema di documentazione, nelle ultime release si è puntato molto, per le principali caratteristiche del sistema, all'integrazione delle voci standard di tesoro con delle note esplicative di carattere descrittivo.

Per quanto riguarda la documentazione del processo produttivo statistico, si è colta l'occasione per ristrutturare il tesoro di SIDI delle operazioni, inserendo le fasi e rivedendo alcune voci per renderle più comprensibili ad utenti non necessariamente esperti di processi produttivi statistici. Sono stati compiuti degli sforzi per riorganizzare le voci secondo la logica temporale delle operazioni che si svolgono nei tipici processi produttivi statistici, tenendo conto delle loro peculiarità, se indagini di tipo diretto, amministrativo oppure se elaborazioni.

Così come è nella logica del SIDI, per ogni operazione dettagliata possono essere documentate le attività messe in atto per la qualità, dette "azioni di controllo". Rispetto al tesoro presente in SIDI, uno sforzo è stato fatto per organizzarle in modo esplicito secondo la logica - comunque sottostante - di attività di prevenzione, controllo in corso d'opera e valutazione dell'errore.

Tali cambiamenti apportati nei tesauri hanno implicato la creazione di un livello aggiuntivo nel tesoro stesso. Per evitare un incremento della sua complessità, dove possibile, si è cercato di compensare levando dei livelli della gerarchia.

In SIDI, le informazioni sulle operazioni e azioni di controllo sono corredate da informazioni aggiuntive (attributi dell'operazione, enti e agenti dell'operazione) che, dato il livello troppo elevato di dettaglio, non sono state riportate su SIQual.

Per quanto riguarda quelli che in SIDI sono chiamati "modelli di rilevazione" e "supporti di diffusione", ci si è limitati semplicemente a chiamarli "questionari" e "pubblicazioni" nel sistema SIQual, mentre la tipologia di informazione è rimasta invariata.

E' stato deciso di includere in SIQual un'area legata alle rilevazioni o alle elaborazioni di documenti sulla qualità, mentre si è ritenuto di non includere l'ambiente di interrogazione degli indicatori standard di qualità per le rilevazioni, presente nel sistema SIDI.

Questa area di documenti sulla qualità è organizzata in tre tipologie di documentazione:

- note metodologiche sintetiche
- le schede standard di qualità
- documenti approfonditi sulla qualità.

Le note metodologiche sintetiche corrispondono come schema alle note metodologiche dell'ASI. Infatti, se fino ad oggi, questa documentazione sintetica sugli aspetti salienti del processo e della qualità è stata predisposta per gli utenti esterni a corredo delle tavole incluse nell'Annuario Statistico Italiano, in futuro tale documentazione dovrà rappresentare uno standard minimo disponibile per tutte le indagini, da poter inserire in tutte le pubblicazioni di tipo trasversale o in quelle tematiche che integrano dati provenienti da più fonti.

Le schede standard di qualità rappresentano una documentazione di tipo qualitativo standard e strutturata sulla qualità. Per le rilevazioni, il modello di riferimento adottato e da seguire nella predisposizione di tali schede è quello che si basa sulle dimensioni della qualità definite in sede Eurostat (rilevanza, accuratezza, tempestività e puntualità, accessibilità e trasparenza, confrontabilità, coerenza). Per le elaborazioni, non è ancora stato formalizzato un modello di riferimento, ma l'orientamento è quello di assumere come modello la documentazione sulla qualità del Fondo Monetario Internazionale (DQAF, Data Quality Assessment Framework).

Infine, i documenti approfonditi sulla qualità rappresentano una forma di documentazione non standard che i responsabili di indagine possono predisporre per documentare particolari aspetti attinenti alla qualità.

Così come in SIDI, esiste poi un'area di documentazione più generale dove gli utenti possono accedere a documentazione di tipo metodologico, operativo, documenti Eurostat, regolamenti europei e all'elenco di tutti i questionari disponibili on line (per le indagini con più questionari in linea, solo il più recente viene mostrato in questo elenco). Relativamente ai documenti di tipo metodologico e operativo, si tratta di un sottoinsieme di documenti già disponibili sul sito dell'Istat in altre pagine, che hanno un carattere di interesse per la qualità e che è quindi coerente diffondere nell'ambito del sistema informativo sulla qualità SIQual.

Infine, è stata inserita una area di *download* dove l'utente può scaricare i tesauri del sistema. I singoli tesauri sono comunque accessibili mediante link, da ogni parte del sistema dove vengono citati.

## **2.2 Completezza e validità delle informazioni presenti nel sistema**

Per evitare che l'utente del sistema SIQual si trovi ad interrogare processi scarsamente documentati, sono stati stabiliti dei requisiti minimi di documentazione per le rilevazioni e le elaborazioni presenti nel sistema. Infatti vengono caricate nel database del sistema le sole rilevazioni che hanno documentato almeno il fenomeno, le unità di analisi e almeno una operazione e una azione di controllo. E' da sottolineare che raramente si incontrano indagini con solo una operazione, in quanto, una volta che il referente di indagine documenta il processo in SIDI, lo fa in modo più completo.

Relativamente al processo di validazione delle informazioni immesse, questo si sviluppa principalmente in due passi. Bisogna considerare che i metadati di processo documentati in SIDI si basano su una serie di tesauri, costruiti e mantenuti a livello centralizzato, le cui voci sono pertanto state validate. I tesauri non sono però rigidi ma consentono ai referenti di indagine di inserire nuove voci qualora non trovino in quelle già presenti delle descrizioni idonee a documentare la loro rilevazione o la loro elaborazione. Queste voci vengono quindi validate a livello centralizzato per assicurare la opportunità dell'inserimento o meno. Nel sistema di diffusione SIQual, solo le voci validate sono messe a disposizione degli utenti, pertanto ciò che non è stato considerato valido da parte del gruppo che gestisce il sistema SIDI non viene caricato nel data base di SIQual. Un ulteriore processo di validazione delle informazioni immesse, soprattutto per quelle note descrittive che sono messe a corredo delle voci standard, è stato effettuato in relazione alla Direttiva trasversale dell'Istituto del 2005 sul completamento del popolamento di SIDI. In questa occasione, ai referenti di direzione sono stati forniti i report con tutte le informazioni immesse per le indagini della loro direzione, al fine di ottenere un ulteriore controllo su quanto documentato.

Per quanto riguarda le elaborazioni, queste risultano documentate e validate nel momento della predisposizione delle note metodologiche dell'Annuario Statistico Italiano, in quanto in tale occasione viene predisposta questa documentazione direttamente per gli utenti esterni.

## **2.3 Modalità di interrogazione, navigazione e altre funzioni del sistema**

Relativamente alle funzionalità di navigazione del sistema le principali scelte che sono state valutate hanno riguardato:

- le modalità di accesso alle informazioni relative ai processi produttivi statistici;
- le soluzioni grafiche da adottare;

Per stabilire in che modo l'utente potesse accedere alle informazioni relative ai processi produttivi statistici, si è effettuata una analisi delle modalità di accesso disponibili in SIDI-TOP. Queste, come era ovvio, sono

risultate maggiormente idonee ad utenti esperti ed interni all'istituto, pertanto si è ritenuto di enfatizzare una selezione dei processi di tipo guidato e tematico (Scelta guidata). Altre modalità di selezione di un processo da interrogare relativamente ai metadati e alla qualità sono: attraverso l'elenco completo dei processi oppure con una ricerca (semplice o avanzata) per parola chiave.

La scelta guidata avviene attraverso un albero di selezione formato da: area, settore, argomento e fenomeno. Le aree i settori e gli argomenti sono quelli definiti dal Piano Statistico Nazionale.

L'utente che selezioni un area ottiene automaticamente l'elenco dei processi (rilevazioni ed elaborazioni) definiti per quella area; se seleziona un settore all'interno dell'area, questo elenco si restringe ai processi specifici di quel settore, e così via per gli argomenti e i fenomeni. L'utente è aiutato graficamente a capire per quale caratteristica (area, settore ...) sta interrogando il sistema in quanto questa, presente nel titolo della porzione di schermo di selezione, diventa in grassetto. Inoltre l'albero si apre seguendo quelle che sono gli standard di tutti i software in distribuzione, un piccolo quadrato con un *più* o con un *meno* all'interno indica che l'albero può essere ulteriormente espanso oppure no.

Il tesoro da cui effettuare le scelte è costruito in modo dinamico e assicura che non vi siano richieste che danno luogo a casi vuoti; in altre parole qualsiasi richiesta porta sempre ad un processo come risposta. La schermata rappresentativa di questa funzionalità è illustrata nel capitolo 3, Figura 3.2.

Un'altra modalità di accesso è attraverso l'elenco completo dei processi. Qui viene data la possibilità di evidenziare in neretto quei processi che contengano nel titolo una certa parola o radice ed esiste la possibilità di restringere l'elenco completo applicando dei filtri su aspetti del processo di carattere più generale. I filtri predisposti riguardano:

- la periodicità
- il tipo di processo (rilevazione o elaborazione).

Per le rilevazioni ulteriori filtri sono:

- il tipo di rilevazione (diretta, amministrativa, mista)
- il tipo di disegno (totale, campionario).

Se la rilevazione è diretta o mista, un ulteriore filtro riguarda la

- tecnica di rilevazione (faccia a faccia, telefonica, postale, non postale per autocompilazione, telematica).

Se invece il processo è di tipo elaborazione, il solo filtro aggiuntivo riguarda:

- fonte dei dati (Istat, non Istat, mista).

La terza modalità di accesso ai dati è la ricerca semplice e avanzata. Nella ricerca semplice è sufficiente digitare una parola affinché il sistema fornisca l'elenco dei processi per i quali la parola è stata trovata in almeno uno dei seguenti campi: nome processo, area, settore, argomento e fenomeno, unità di analisi, nome della pubblicazione, dati on-line. L'utente può scegliere di limitare la ricerca ad un sottoinsieme dei precedenti campi. Nella ricerca avanzata è possibile ricercare per la presenza nei suddetti campi di: più parole, frasi o una qualunque parola di un elenco di parole. Le modalità di evidenziazione di parole o radici di parole e di filtro così come definite per l'elenco completo sono quindi applicabili al risultato della ricerca sia semplice che avanzata.

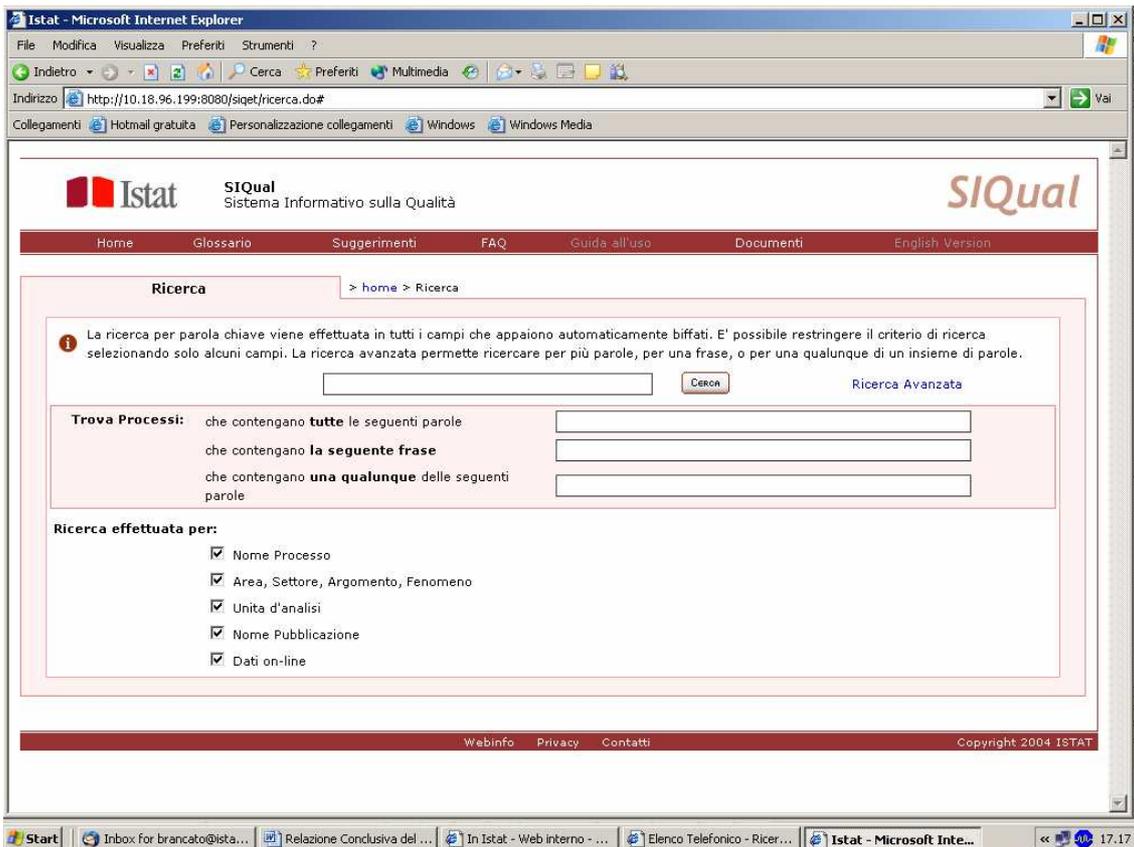


Figura 2.1. Ricerca del processo per parola chiave

Una volta che l'utente ha identificato, attraverso una delle modalità di accesso precedentemente descritte, il processo di cui vuole interrogare i metadati o la qualità, si trova di fronte ad una pagina di presentazione dei metadati, dove sono raggruppate le caratteristiche più salienti per il processo e può via via navigare nelle pagine specifiche per le rilevazioni e le elaborazioni dove trova ulteriori informazioni.

Nel sistema sono inoltre disponibili:

- un glossario di termini utilizzati all'interno del sistema
- un questionario sul gradimento on-line per acquisire informazioni sugli utenti
- una pagina dei quesiti più frequenti (faq)
- report di stampa

Il glossario è stato predisposto in modo da poter essere organizzato in sottosezioni specifiche (per esempio termini specifici per i processi di tipo economico, ...).

Il questionario sul gradimento in linea rileva informazioni su:

- motivo della navigazione;
- valutazione del sistema su una scala di gradimento da 1 a 7, dove 1 rappresenta la massima insoddisfazione e 7 il massimo gradimento, relativamente ad una serie di caratteristiche quali: la quantità e la qualità delle informazioni messe a disposizione, la chiarezza del linguaggio utilizzato, la facilità nel trovare le informazioni, la funzionalità e la grafica del sistema;
- il profilo dell'utente: età, sesso, stato civile, nazionalità, titolo di studio, professione;
- un area di commenti libera.

Considerando che a priori un'analisi strutturata sulle esigenze conoscitive degli utenti nell'ambito della qualità non è stata fatta, era importante per un sistema come SIQual, avere delle informazioni di ritorno da parte degli utenti per cercare di migliorare continuamente l'offerta di informazione. Per una visualizzazione del tipo di questionario on-line si faccia riferimento a alla Figura 3.3 del capitolo 3.

La pagina dei quesiti più frequenti dovrebbe via via ampliarsi sulla base dei commenti che verranno inviati da parte degli utenti. Al momento del primo rilascio, sono stati inseriti solo alcuni quesiti sulle definizioni di base.

Infine, una funzione di report di stampa, permette di stampare o salvare su file in formato *pdf* tutte o parte delle informazioni del sistema.

## 2.4 Soluzioni grafiche adottate e struttura

Considerando che il sistema informativo sulla qualità è stato progettato per essere inserito nel portale dell'Istituto, particolarmente rilevante era l'aspetto grafico sia per aderire a quelli che erano gli standard del portale, sia per offrire agli utenti un sistema di navigazione gradevole e intuitivo.

Per quanto riguarda gli aspetti di gradevolezza e facilità di comprensione, la struttura generale comprende delle note esplicative per ogni funzione e dei link al glossario per ogni voce del sistema.

Per la rappresentazione delle operazioni e azioni di controllo delle rilevazioni e delle elaborazioni si è pensato ad una soluzione grafica che rappresentasse il flusso delle operazioni nella loro sequenza temporale. Questa viene riportata nella Figura 2.2 che segue. Quindi, cliccando su ciascuna fase, si ottiene il dettaglio delle operazioni e delle azioni di controllo della qualità svolte all'interno della fase. Le azioni della qualità sono evidenziate in rosso e la tipologia (preventive, in corso d'opera, valutative) in grassetto.

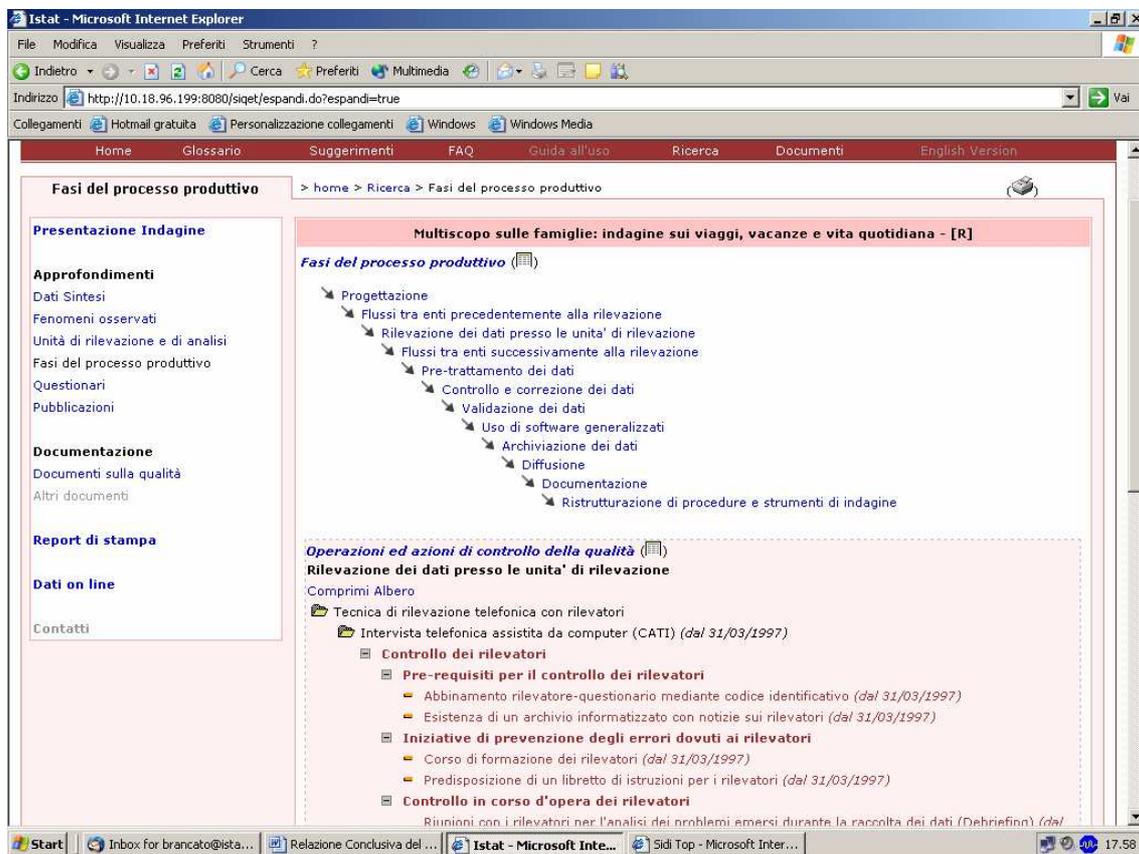


Figura 2.2. Documentazione delle fasi, operazioni e azioni di controllo di un processo produttivo statistico

Le sequenze tra enti precedentemente e successivamente la raccolta dei dati sono state inserite nel flusso e la loro rappresentazione fa comprendere proprio la sequenza delle operazioni e gli enti che vi sono coinvolti.

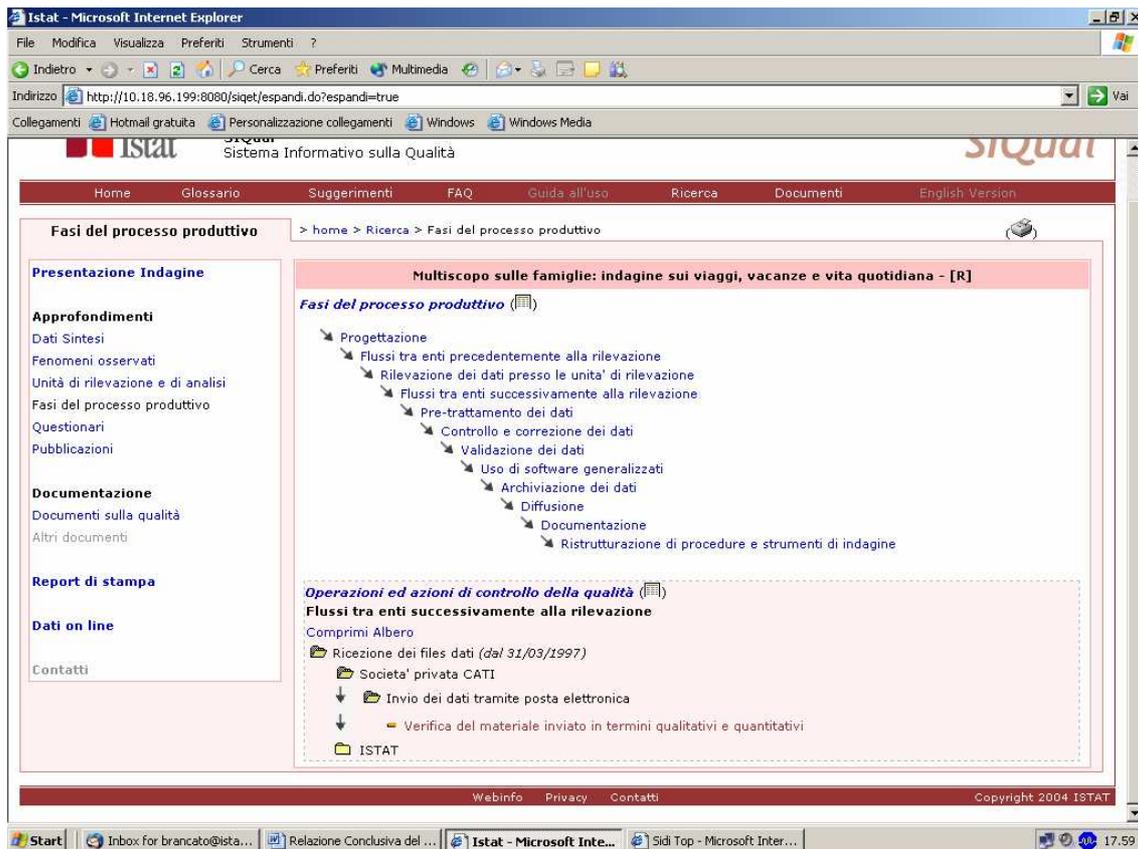


Figura 2.3. Documentazione delle sequenze operazioni preliminari o successive alla fase di rilevazione

### Capitolo 3. Aspetti di progettazione e sviluppo informatico

Nella progettazione del sistema SIQual si è tenuto conto di diversi fattori che permettessero di realizzare un sistema che, oltre che orientato ad un'utenza esterna, rispettasse gli standard definiti dai responsabili del portale di Istituto. L'obiettivo di questo studio è stato quello di individuare soluzioni dal punto di vista tecnologico, applicativo ed organizzativo ed a tal fine si sono analizzati diversi aspetti prima di procedere alla realizzazione del sistema.

#### 3.1 Architettura logico-fisica

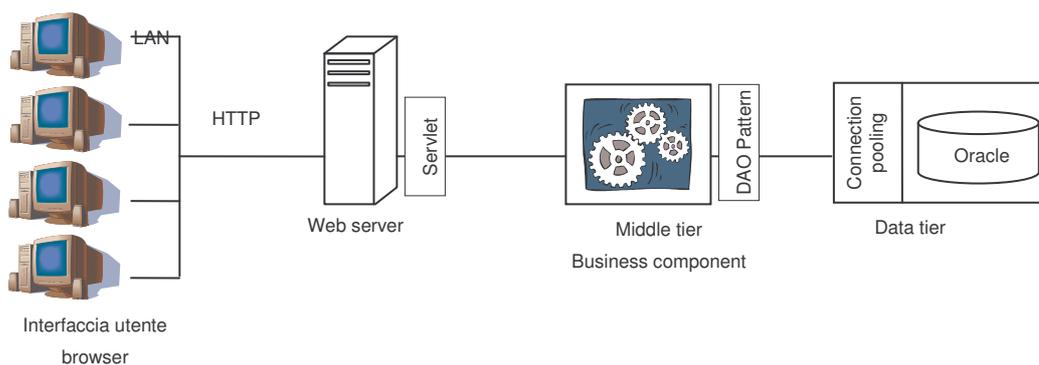
La soluzione architetturale adottata è quella a 3 livelli, di tipo thin client che prevede l'utilizzo di Oracle RDBMS (Relational DataBase Management System) come database server, di Apache/Tomcat e della tecnologia JSP (Java Server Pages) come application server e di un web browser come interfaccia utente.

L'applicazione web è stata concepita per funzionare con i principali browser, le pagine fornite al browser dall'applicazione sono prive di qualsiasi logica di controllo o funzionale, il back-end possiede tutta la logica dell'applicazione la quale risiede in normali classi java utilizzate dalle jsp, preposte a presentare le elaborazioni a seguito di istanziazione delle prime (le stesse classi di business logic accedono al terzo livello dell'architettura, il Db Oracle, via Jdbc).

L'installazione e messa in esercizio del sistema dal punto di vista tecnico è molto semplice, infatti, l'applicazione viene consegnata sotto forma di .war ed è sufficiente inserire questo file nella cartella webapps di Tomcat ed effettuare il deploy.

SIQual è un'applicazione Java Web-based costruita su uno schema architetturale riferito al paradigma MVC (Model-View-Controller). Questo modello, in base ai principi della programmazione Object-Oriented, tende a separare in componenti i sistemi per i quali l'interazione uomo-macchina assume un ruolo molto importante. Il principio fondamentale alla base del modello MVC, infatti, è la netta separazione fra le strutture dati dell'applicazione (Model), la loro rappresentazione (View) e l'interazione con l'utilizzatore (Controller). Tutto ciò al fine di rendere più semplici sia l'implementazione e la manutenzione della procedura sia le future evoluzioni.

Nello schema seguente è sintetizzata la suddivisione dell'applicazione in base al modello MVC.



Sul server Web è stato installato il JSP Servlet Engine Tomcat con il compito di controllare le richieste degli utenti, richiedere le elaborazioni ed i dati ai componenti di business ed inviare le risposte ai client.

I componenti di business risiedono su un livello intermedio fra web server e data server. In questi componenti è implementata la logica di business dell'applicazione.

Sul data server, infine, è disponibile l'RDBMS Oracle dove sono memorizzate le informazioni persistenti.

In particolare il modello MVC è stato implementato come segue:

- Il livello Model attraverso le JavaBean.
- Il livello View attraverso lo standard JSP/JSTL, le librerie standard JSTL fornite dal gruppo jakarta.apache.com, sotto licenza GNU
- Il livello Controller attraverso il framework applicativo **Struts** 1.1 fornito dal gruppo jakarta.apache.com, sotto licenza GNU.

La particolarità di Struts risiede nella possibilità di definire mediante un file in formato XML, "struts-config.xml", quale è il flusso logico della navigazione del sito e quali modelli vanno caricati e se vanno validati o meno.

In questa sorta di "file di istruzioni" del flusso sono descritte le classi che implementano:

- i Form Bean: Java Bean destinati ad accogliere e validare i parametri della request generalmente provenienti da forms HTML
- le Action: contenenti singole parti della Business Logic dell'applicazione che agiscono "per delega" del controller principale per mezzo dell'evento scatenato dall'utente attraverso il click del mouse.

Inoltre sono anche definiti i flussi in uscita da ciascuna Action ai quali viene attribuito un identificativo logico testuale.

### 3.2 Architettura dati

L'RDBMS utilizzato dall'applicazione per la memorizzazione dei dati è Oracle versione 8.1.7 e risiederà in esercizio sul server ISTAT esposto RODI.

La struttura architetturale è diversa da quella attuata per il SIDI perché si estraggono soltanto le informazioni di carattere qualitativo e valide e si è reso necessario denormalizzare la base dati per ottimizzare i principali criteri di selezione presenti nel sistema.

Come descritto nel capitolo 2 si è fatto in modo che nella funzione relativa alla ricerca tematica (Scelta guidata) le richieste effettuate dall'utente portassero sempre alla selezione di almeno un processo. Per far ciò si è creata una tabella gerarchica categorizzando i concetti Area Settori Argomenti e Fenomeni tramite le relazioni con le indagini.

Per realizzare la presentazione dei processi produttivi e delle azioni di controllo a loro collegate si è reso necessario creare una tabella che contenesse la gerarchia delle fasi, operazioni ed azioni di controllo in forma denormalizzata (si veda paragrafo 2.4, Figura 2.2).

Molto importante per un sistema di navigazione che tratta metainformazione è avere un potente motore di ricerca che permetta di ricercare parole all'interno del sistema sia su tabelle di dati che all'interno di documenti. A tal fine è stata effettuata una sperimentazione che ha permesso di ottimizzare l'utilizzo del tool di *Oracle InterMedia*.

*Oracle interMedia* è un sistema di gestione dei dati multimediali molto potente che fornisce funzionalità per ricerche testuali avanzate (ricerche tematica, catalogazione automatica, etc.) e tecniche per il confronto di immagini, video e suoni.

Il modulo utilizzato all'interno del sistema è *Oracle Text* per la gestione dei dati testuali, che attraverso tecniche di indicizzazione e interrogazione riesce a reperire le informazioni di interesse. Per l'aggiornamento dei dati, attraverso l'informazione di SIDI, sono state realizzate delle procedure di ETL in PL/SQL che girano in batch con cadenza settimanale.

### 3.3 Standard grafici

Nella realizzazione del sistema si è tenuto conto degli standard definiti dai responsabili del portale, a tal fine:

- ci si è orientati su colori non contraddittori con quelli del sito Istat e con tonalità sobrie
- lo sfondo delle pagine è bianco
- non sono state utilizzate applets e codice JavaScript
- sono state utilizzate con molta parsimonia immagini e la maggior parte corrispondono con quelle già presenti sul sito Istat
- sono stati utilizzati dei fogli di stili (CSS) dove per i testi sono stati definiti i font adottati per le diverse categorie di caratteri:

Tipo	Font	Dimensione	Colore
Titoli	Verdana	2 grassetto	nero
Sottotitoli e Testo	Verdana	1 normale	nero
link	Verdana	1 normale	blu

L'interfaccia del sistema SIQual è di tipo *GUI (Graphical User Interfaces)* per Web Application conforme allo standard ISO.

Il sistema SIQual è stato progettato, dal punto di vista dell'interfaccia grafica, intorno a delle aree raggruppate per caratteristiche funzionali, per rendere la navigazione agevole all'utente. Esse si possono distinguere in:

- Home Page
- Pagine di consultazione
- Pagine interattive
- Pagine di esito.

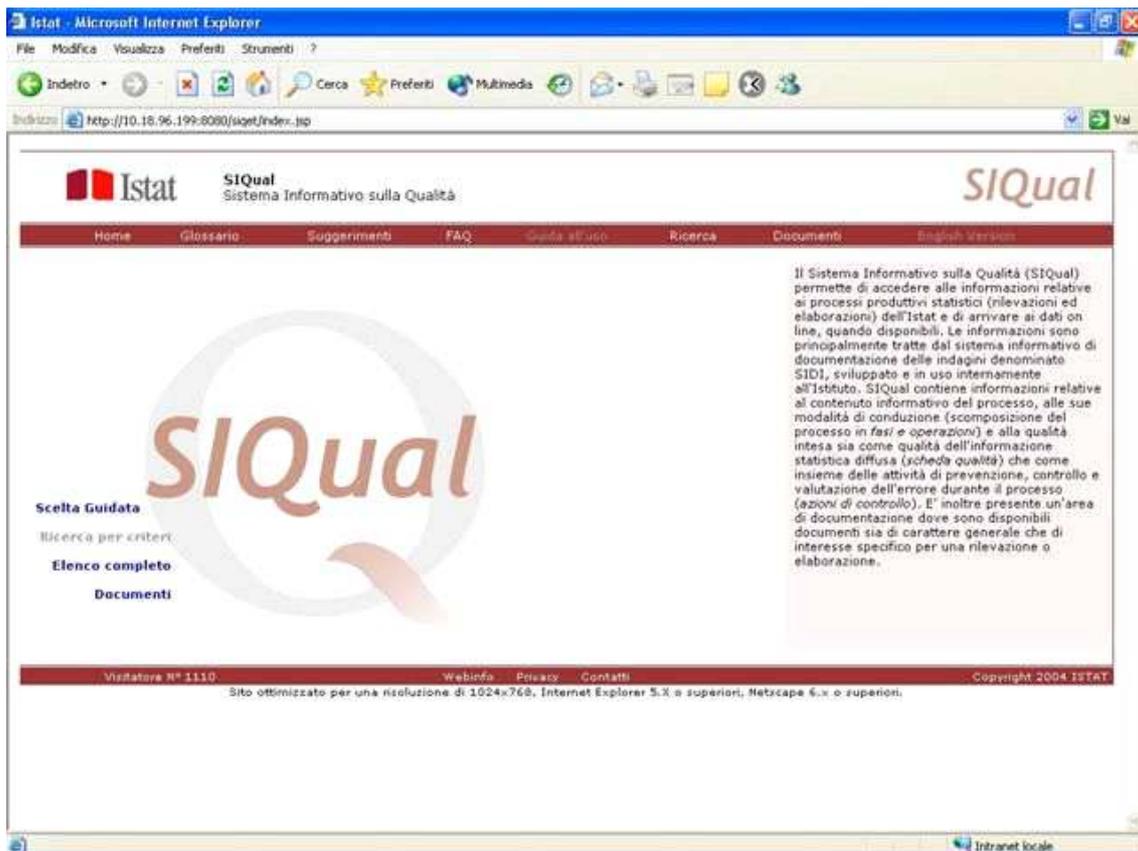
#### *Home Page*

L'home page (Figura 3.1) presenta già buona parte delle interazioni possibili tra l'utente ed il sistema SIQual. Le differenze con le altre pagine sono date dalle particolari funzionalità di questa pagina che è la via di accesso principale al sito.

Quelli che seguono sono i principali aspetti attorno a cui è stata progettata l'home page:

- **Fornire identità:** per assolvere a questa funzione è stato inserito il logo di SIQual in evidenza al centro dell'home page e rappresenta un invito al nuovo utente e un facile riconoscimento per l'utente che abbia già visitato le pagine del sistema.
- **Orientamento e navigazione:** molta attenzione è stata fatta nel fornire un chiaro e semplice orientamento e criterio di navigazione all'utente. La navigazione nel sistema è stata divisa in chiari e diversi menù di navigazione:
  - il menù principale in testa,
  - il menù della consultazione dei dati al centro accanto al logo SIQual
  - il menù secondario o delle avvertenze.

Nel corpo della pagina è stato previsto uno spazio a destra dove l'utente viene informato su cosa offre il sistema tramite una breve introduzione.



**Figura 3.1** Home page del sistema

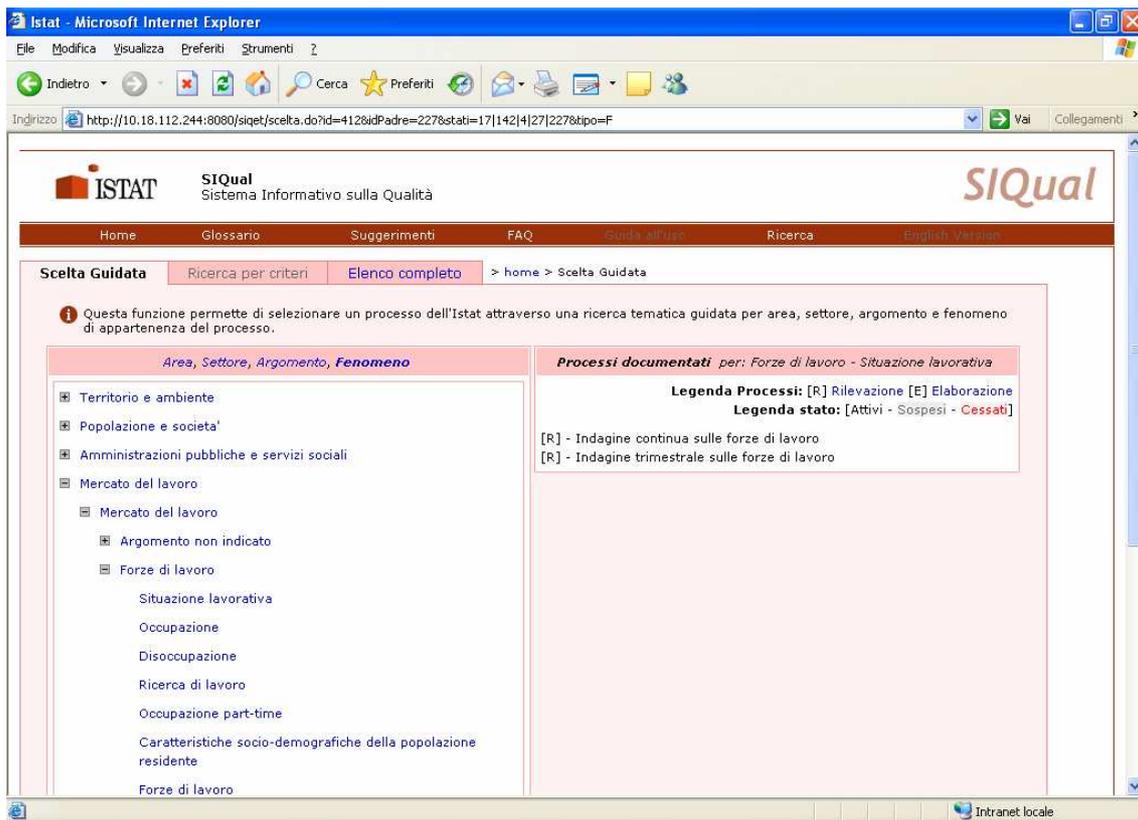
In tutti i casi sopra riportati i menù sono costituiti da termini che corrispondono a dei link che indirizzano l'utente alla pagina corrispondente alla funzione scelta.

Ai rispettivi link o elementi del menù è legata una "Action", che viene intercettato dal server attraverso la tecnologia JSP del web server Tomcat ed il Plug-in Jakarta Struts della Apache Software Foundation. Intercettata la richiesta dell'utente il server, dopo aver effettuato le elaborazioni ed estratto le informazioni dal Database per la richiesta specifica, caricherà la pagina opportuna.

### ***Pagine di consultazione***

Le pagine di consultazione sono estremamente dinamiche e mutano considerevolmente al variare del processo in esame. In queste pagine compaiono oggetti più complessi (appartenenti allo standard della navigazione internet) come combo box, check box, radio button con i quali è possibile operare per applicare dei filtri sugli oggetti che si stanno consultando.

Ma soprattutto gli alberi che nella tipica visualizzazione rappresentano un oggetto potente per la navigazione in strutture di catalogazione. Questi oggetti sono estremamente interattivi poiché possono essere penetrati dall'utente per poter scendere nel dettaglio dell'informazione ricercata e selezionare tra quelle presenti ciò che interessa. Gli alberi sono stati utilizzati nel menù della consultazione dei dati nella scelta guidata (Figura 3.2).



**Figura 3.2 Albero della scelta guidata**

Le pagine di consultazione reperiscono le loro informazioni accedendo al Database e visualizzando solo i dati che soddisfano le esigenze dell'utente.

### ***Pagine interattive***

Permettono all'utente di agire direttamente col sistema. Esse contengono, al loro interno, anche le aree di testo per poter inserire dati in input che vengono memorizzati nel DB (come ad esempio la pagina dei suggerimenti, Figura 3.3) oppure aree di testo digitate per effettuare delle ricerche attraverso il motore di ricerca presente nel sistema (Figura 2.1).

### ***Pagine di esito***

Costituiscono le pagine dove vengono forniti all'utente i messaggi del corretto invio di dati al sistema o le informative a causa di un malfunzionamento da parte del server (Figura 3.4).

Nelle schermate ci sono ulteriori indicazioni inserite in aiuto all'utente sia per interpretare correttamente le informazioni che il sistema offre, come già descritto nel capitolo 2, e sia per orientarsi nella navigazione, a tal fine è stato inserito un riepilogo linkabile sul percorso effettuato. Esso permette all'utente, tramite i link, di tornare alle pagine precedentemente visitate.

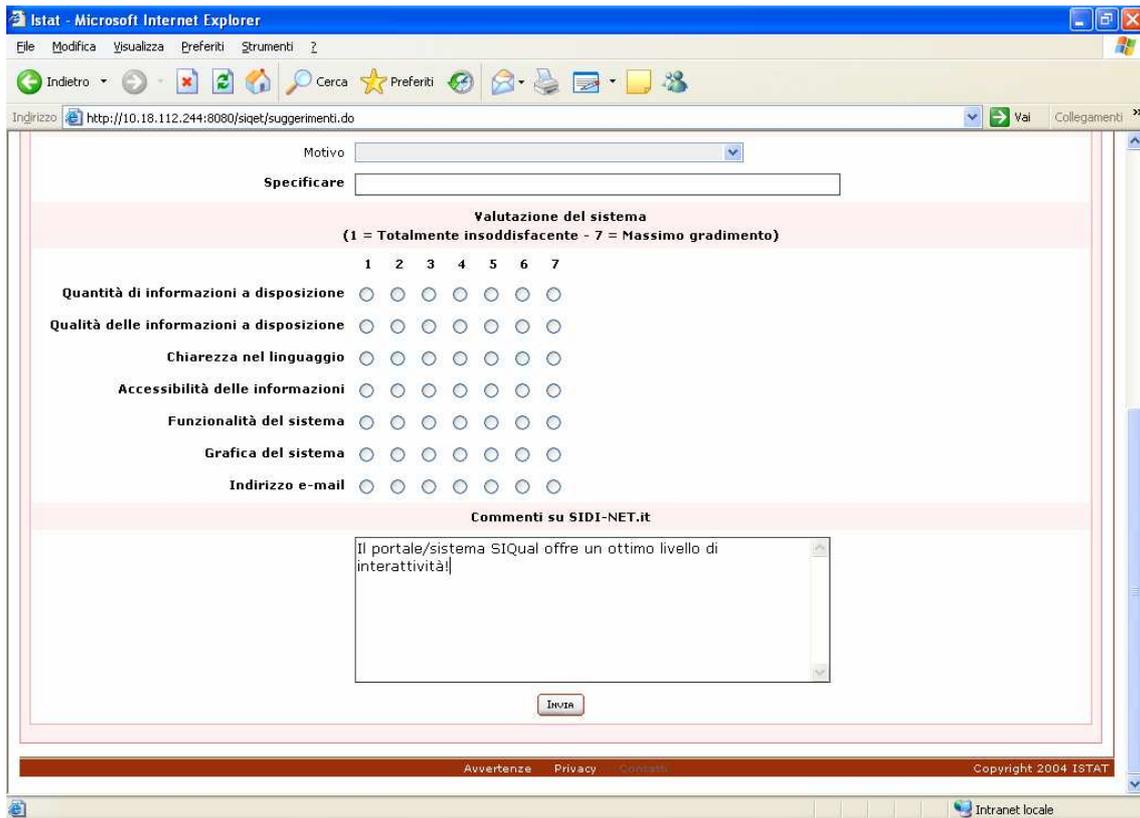


Figura 3.3 Pagina dei suggerimenti

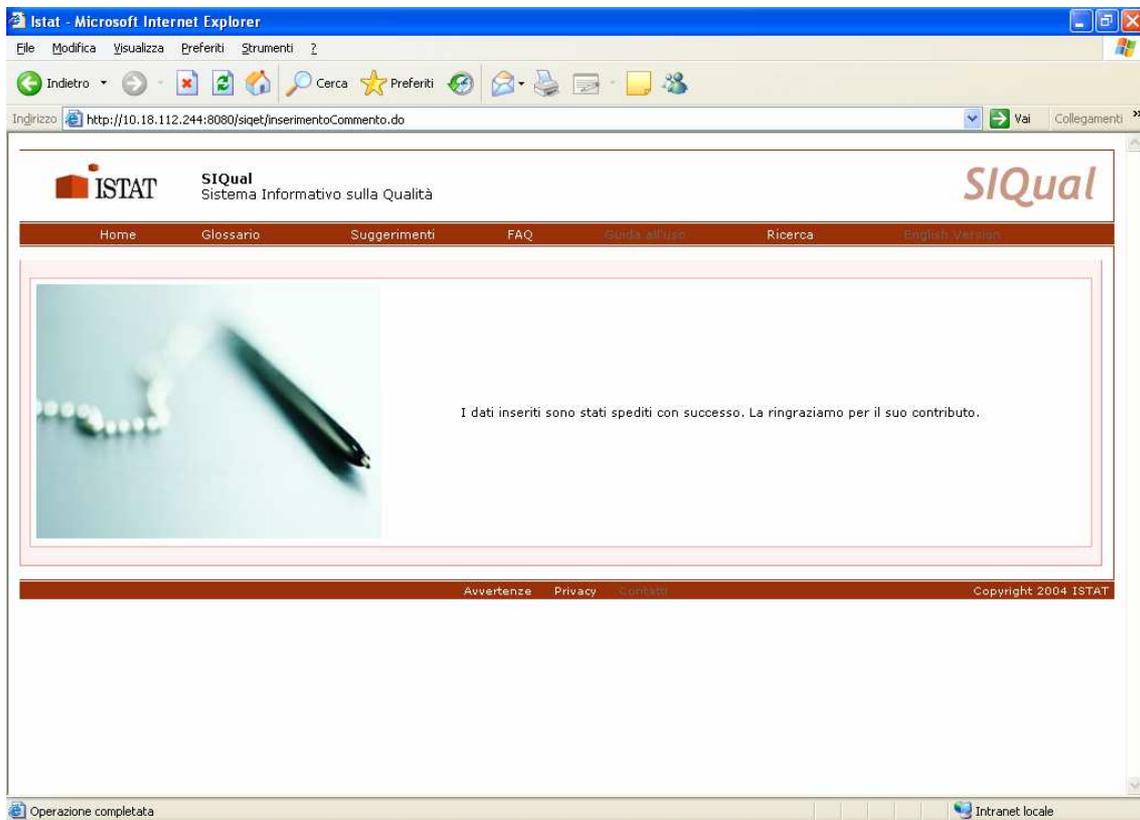


Figura 3.4 Pagina di esito

## Capitolo 4

### Il sistema di acquisizione delle tavole ASI: AcqAsi

Come già accennato, la funzione di acquisizione delle tavole dell'ASI (Annuario Statistico Italiano) è stato un obiettivo iniziale al quale si è rinunciato in relazione a sviluppi paralleli di altri progetti.

Ciò nonostante l'esperienza prototipale del sistema AcqAsi rappresenta una possibile risposta ad alcune esigenze legate al trattamento di macrodati quali tavole e grafici. Parte dell'esperienza AcqAsi è confluita nel progetto Sisco.T (Servizio Informativo Statistico Comunale.Tavole) avente l'obiettivo di fornire ai piccoli e medi Comuni facenti parte del Sistan uno strumento che agevolasse la produzione di opuscoli statistici, Annuari, e sezioni web a partire da dati già diffusi dall'Istituto.

Il primo prototipo, realizzato secondo gli obiettivi iniziali e denominato SIQeT (Sistema Informativo Qualità e Tavole), prevedeva l'interrogazione per metadati di rilevazione e di elaborazione e per le tavole dell'Annuario. Per poter acquisire le tavole dell'ASI in modo standard è stato sviluppato il sistema AcqAsi, brevemente descritto successivamente. Una volta acquisite le tavole, ne era previsto il caricamento nel database del sistema SIQeT.

#### 4.1 Il sistema di navigazione SIQeT

In primo luogo può essere definito lo schema di navigazione del sistema inizialmente ipotizzato (Figura 4.1).

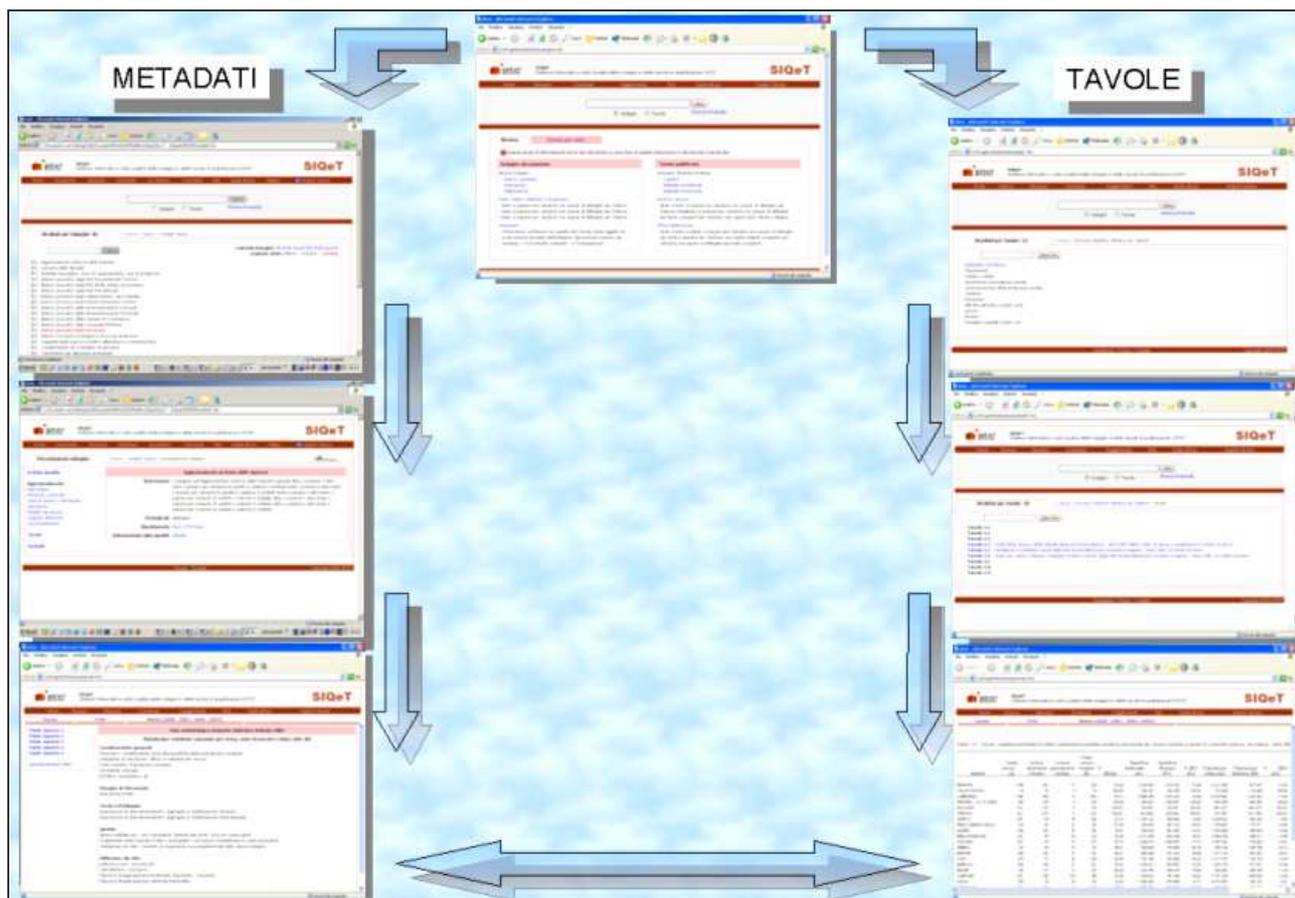


Figura 4.1 Schema di navigazione

L'utente può procedere secondo due linee di interesse: la documentazione del processo (metadati) o le tavole statistiche dell'ASI.

Procedendo verso la documentazione di processo, che comprende sia la documentazione sintetica rappresentata dalle note ASI, sia i suoi approfondimenti, rappresentati dai metadati contenuti nel sistema

Sidi, (relativi alle operazioni ed alle azioni di controllo adottate nel processo di produzione statistico), l'utente può selezionare il processo statistico di interesse mediante alcuni criteri:

1. denominazione
2. fenomeni osservati
3. unità di analisi
4. tipo di campionamento
5. periodicità

A partire dalla documentazione di processo visualizzata, l'utente può continuare il suo percorso e visualizzare, se presente, la corrispondente tavola ASI. A questo punto può risalire attraverso i metadati di tavola (anno di riferimento, livello territoriale, parola chiave) e tornare all'home page.

Procedendo secondo la seconda linea di interesse, lo schema di navigazione permette all'utente di cercare la tavola attraverso 5 modalità di ricerca tra le tavole dell'ASI:

1. Per capitolo
2. Per titolo della tavola
3. Per parola contenuta nella tavola
4. Riferimento temporale
5. Riferimento territoriale

Una volta individuata la tavola di interesse è possibile visualizzarla o scaricarla in formato MsExcel. Inoltre è possibile visualizzare sia la documentazione sintetica sia la documentazione completa estratta da Sidi (Figura 4.1).

## **4.2 Sistema di acquisizione AcqAsi**

Mentre i metadati di processo sono forniti da Sidi, nel sistema prospettato vi era l'esigenza del caricamento degli elementi di tavola. Le tavole dell'ASI sono disponibili infatti su file MsExcel e costituiscono un oggetto dotato delle sue caratteristiche che possono essere rilevanti per un utente esterno

Lo schema di acquisizione si basa sugli elementi essenziali che costituiscono la metainformazione della tavola:

1. Volume in cui la tavola è pubblicata
2. Numero capitolo
3. Numero tavola
4. Titolo Tavola
5. Testo contenuto
6. Area principale
  - a. Articolazione territoriale
  - b. Articolazione temporale
7. Area secondaria
  - a. Articolazione territoriale
  - b. Articolazione temporale
8. Note (comprendente la fonte e le eventuali note)
9. Flag di suddivisione su più pagine

Inoltre interessava costruire un DB per la ricerca testuale nei vari campi della tavola, che permettesse una veloce individuazione delle tavole di interesse.

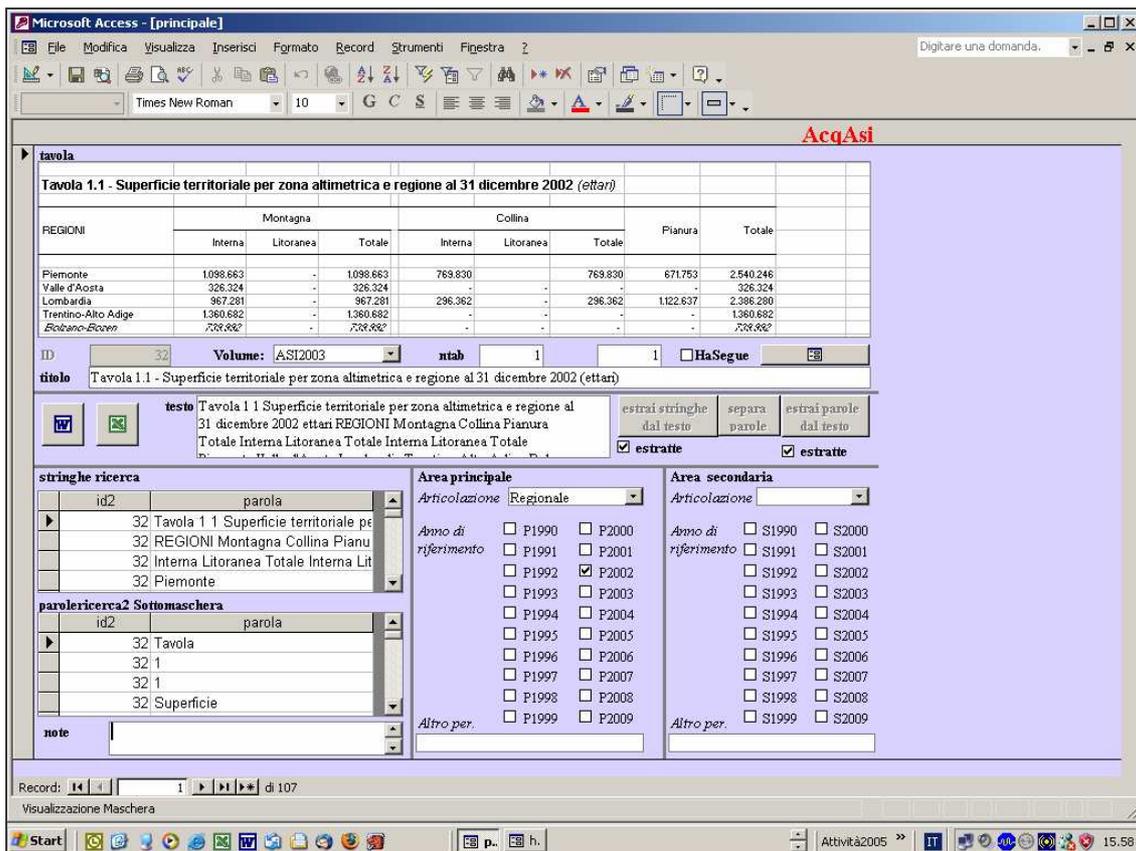


Figura 4.2 Software AcqAsi

Il sistema sfrutta lo standard di fatto MsOffice nel senso che essendo gli oggetti da trattare fogli elettronici formato MsExcel l'applicazione AcqAsi realizzata in MsAccess offre una serie di funzionalità integrate mediante l'utilizzo delle proprietà OLE, tra cui l'estrazione automatica delle parole contenute per alimentare il DB di ricerca testuale. In questo modo l'acquisizione delle tavole risulta essere un processo che presenta un buon grado di automazione.

Per esempio l'estrazione degli elementi testuali dalle tavole ASI in formato MsExcel, viene effettuata automaticamente tramite macro VBA lanciate attraverso MsAccess.

Il prodotto è stato progettato e realizzato per ambienti Windows 2000/XP su database Ms Access. Comprende un database MsAccess (AcqAsi.mdb) per l'archiviazione dei dati che inoltre contiene il software. Le maschere e i report sono state costruite utilizzando esclusivamente i controlli messi a disposizione dall'ambiente di sviluppo di Access e la parte "codice" è stata scritta in VBA. I riferimenti alle altre librerie sono: Visual basic for applications, Microsoft Access 11.0 Object Library, OLE Automation, Microsoft ActiveX Data Object 2.1 Library, Microsoft Office 11.0 Object Library, Microsoft Excel 11.0 Object Library. Le principali tabelle contenute nel database dei dati sono: *Tavola*, in cui vengono archiviate le tavole statistiche (MsExcel) dell'ASI e gli elementi informativi delle tavole; *Parola e parole*: in cui vengono archiviate le singole parole e le frasi.

Operativamente, le fasi dell'acquisizione risultano essere:

1. Selezione della tavola e caricamento automatico di tutti elementi testuali necessari per la ricerca testuale
2. Caricamento manuale del titolo della tabella, del numero tabella, del capitolo, dell'aggregazione principale e secondaria, degli anni di riferimento principali e secondari, della disposizione su più pagine e delle eventuali note
3. Validazione semiautomatica degli elementi caricati

Infine il collegamento tra gli elementi informativi della tavola e i metadati contenuti in Sidi e nel sottosistema Asimet viene garantito dalla chiave formata da numero capitolo, numero tavola ed anno dell'ASI, informazione presente tra i metadati nella parte di predisposizione delle note metodologiche.

## Capitolo 5

---

### Sviluppi futuri

Il sistema è stato sviluppato in moduli, per poter essere rilasciato in modo graduale e incrementale. Una parte che sicuramente merita una riflessione e uno studio approfondito e sulla quale si vuole completare il sistema riguarda la funzionalità di ricerca dei processi in modo flessibile incrociando criteri diversi. Tale ricerca, presente in modo molto dettagliato in SIDI-TOP, ad una prima analisi risulta essere poco idonea agli utenti esterni. Le analisi fino ad ora effettuate indirizzano verso lo sviluppo di una sorta di data warehouse di interrogazione sui metadati. Approfondimenti sono però necessari per intraprendere questa strada.

E' stato evidenziato come la funzionalità del sistema di scelta guidata per area e settore tematico potrebbe essere utilizzata nel portale come criterio generale di accesso ai dati. A tale riguardo, bisogna notare che le aree e settori utilizzati nel SIQual sono quelle del PSN, e non corrispondono completamente a quelle di diffusione già definite per il portale. Una lavoro di raccordo tra le due diverse classificazioni si renderà necessario qualora si decida che SIQual assuma questo ruolo di accesso ai dati nel portale.

Il sistema documenta le pubblicazioni prodotte nell'ambito dei processi produttivi dell'Istituto, specificando anche il tipo di pubblicazione (on line, cartacea ...). Quando on line, queste pubblicazioni sono direttamente disponibili nel sito dell'Istat interrogando l'area del Catalogo on-line. Purtroppo non è ancora stato possibile effettuare l'integrazione tra i due sistemi e quindi tale funzionalità non è integrata nel SIQual.

Al momento, quando dei sistemi informativi di dati sono disponibili sul sito dell'Istat, da SIQual è possibile accedere alla homepage di tali sistemi. Dovrà essere ulteriormente studiata la possibilità di accedere ad aree di questi sistemi informativi di dati più specificamente inerenti alla rilevazione o alla elaborazione e predisporre il link da ciascun sistema di dati verso SIQual.

Per un sistema di questo tipo, appare molto rilevante l'attenzione verso gli utenti di altri paesi, e quindi lo sviluppo della versione in inglese.

Da un punto di vista informatico, SIQual è stato predisposto per poter funzionare una volta che la base di dati sia stata tradotta.

Infine, la parte più qualificante di SIQual riguarda la documentazione sulla qualità. A tale scopo, le note dell'ASI, compilate ad oggi solo per le rilevazioni ed elaborazioni che partecipano all'annuario, dovranno essere estese a tutti i processi. La parte di software applicativo che consente la gestione delle schede di qualità è stata predisposta, in modo da poter includere anche un sottoinsieme di indicatori (quelli definiti da Eurostat) o tutti o parte degli indicatori standard di SIDI. E' necessario sottolineare però che è un scelta politica dell'Istituto quella di definire gli standard di documentazione della qualità per gli utenti esterni. Mentre infatti SIDI-TOP il sistema di navigazione delle indagini sulla Intranet include indicatori di tipo quantitativo, tale scelta potrebbe non essere idonea per gli utenti esterni sul portale dell'Istituto.

## Riferimenti Bibliografici

---

Apache. <http://www.apache.org/>

Apache Tomcat. <http://tomcat.apache.org/>

Apache Struts <http://struts.apache.org/>

Brancato G. (2004). *Il Sistema ASIMET per la gestione delle Note Metodologiche dell'Annuario Statistico Italiano*. Istat, Documenti, n.3, Anno 2004.

[http://www.istat.it/dati/pubbsci/documenti/Documenti/doc\\_2004/2004\\_3.doc](http://www.istat.it/dati/pubbsci/documenti/Documenti/doc_2004/2004_3.doc)

Brancato G., Simeoni G. (2004) *Tesauri del Sistema Informativo di Documentazione delle Indagini (SIDI)*. Istat, Documenti n.1, Anno 2004.

[http://www.istat.it/dati/pubbsci/documenti/Documenti/doc\\_2004/2004\\_1.doc](http://www.istat.it/dati/pubbsci/documenti/Documenti/doc_2004/2004_1.doc)

Brancato G, Pellegrini C, Simeoni G. e Signore M. (2004) *Standardising, Evaluating and Documenting quality: the implementation of Istat information system for survey documentation - SIDI*. European Conference on Quality and Methodology in Official Statistics. Mainz, 24-26 May 2004. CD ROM

Brancato G., Pellegrini C, D'Amore G. (2004) *IT solutions and information systems in support to quality*. Conferinta Nationala: Statistica Oficiala în Slujba Societatii, Bucuresti, 12-14 Iulie 2004.

[http://www.insse.ro/evenimente/conferinta/Rezumate/ISTAT/En/ENAbstract\\_Brancato\\_Damore.pdf](http://www.insse.ro/evenimente/conferinta/Rezumate/ISTAT/En/ENAbstract_Brancato_Damore.pdf)

De Francisci S., Renzetti M., Sindoni G, Tininini L. (2005) *La modellizzazione dei processi nel sistema informativo generalizzato di diffusione dell'Istat*. Istat, Documenti n.15, Anno 2005.

[http://www.istat.it/dati/pubbsci/documenti/Documenti/doc\\_2005/2005\\_15.pdf](http://www.istat.it/dati/pubbsci/documenti/Documenti/doc_2005/2005_15.pdf)

Eclipse. <http://www.eclipse.org/>

IMF. (2001) *Fourth Review of the Fund's Data Standards' Initiatives Supplement on the Data Quality Assessment Framework* <http://www.imf.org/external/np/sta/dsbb/2001/supp.htm>

Oracle. <http://www.oracle.com/index.html>