



DOCUMENTI ISTAT

SISTEMA INFORMATIVO TERRITORIALE SULLA GIUSTIZIA: *strumenti di gestione e manutenzione*

Massimiliano Renzetti* e Annamaria Urbano**

* ISTAT – DPTS / DCMT / Servizio metodologie, tecnologie e software per la produzione dell'informazione statistica

** ISTAT – DPTS / DCIS / Servizio Giustizia

SOMMARIO
PREMESSA
1. IL SISTEMA INFORMATIVO TERRITORIALE SULLA GIUSTIZIA
1.1. IL SOTTO-SISTEMA DI PRODUZIONE
1.2. IL SOTTO-SISTEMA DI DIFFUSIONE
2. IL PROCESSO DI PRODUZIONE E DIFFUSIONE
2.1. IL PROCESSO SENZA UNO STRUMENTO DI GESTIONE
2.2. IL PROCESSO CON UNO STRUMENTO DI GESTIONE
3. IL SISTEMA DI GESTIONE E MANUTENZIONE DEL SITG
4. DESCRIZIONE FUNZIONALE DEL SISTEMA DI GESTIONE E MANUTENZIONE
4.1. FUNZIONALITÀ PER LA GESTIONE DELLE UTENZE
4.2. FUNZIONALITÀ PER LA GESTIONE DELLE FONTI DATI
4.3. FUNZIONALITÀ PER LA GESTIONE DELLE CATEGORIE
4.4. FUNZIONALITÀ PER LA GESTIONE DEGLI INDICATORI
4.5. FUNZIONALITÀ PER LA GESTIONE DELLE TAVOLE STATISTICHE
4.6. FUNZIONALITÀ PER LA GESTIONE DELLA DOCUMENTAZIONE
5. DESCRIZIONE TECNICA DEL SISTEMA DI GESTIONE E MANUTENZIONE
5.1. ARCHITETTURA DEL SISTEMA
5.2. ARCHITETTURA DEI DATI
5.2.1. MICRODATI
5.2.2. MACRODATI
5.2.3. KNOWLEDGE BASE
5.2.4. TAVOLE E DOCUMENTAZIONE
5.3. ARCHITETTURA DI DISPIEGAMENTO
6. IL CICLO DI VITA DELLE FONTI DATI, DEGLI INDICATORI E DELLE TAVOLE
6.1. CICLO DI VITA DELLE FONTI DATI
6.2. CICLO DI VITA DEGLI INDICATORI
6.3. CICLO DI VITA DELLE TAVOLE
7. SCENARI DI UTILIZZO DEL SISTEMA DI GESTIONE E MANUTENZIONE
7.1. INSERIMENTO DI UNA NUOVA INDAGINE
7.1.1. ATTIVITÀ PRELIMINARI
7.1.2. CARICAMENTO DEI MICRODATI
7.1.3. CREAZIONE DELLA FONTE DATI
7.1.4. DEFINIZIONE E CALCOLO DEGLI INDICATORI
7.1.5. DEFINIZIONE, GENERAZIONE E PUBBLICAZIONE DELLE TAVOLE
7.1.6. PUBBLICAZIONE DELLA DOCUMENTAZIONE
7.2. MANUTENZIONE DI UN'INDAGINE ESISTENTE
7.2.1. INSERIMENTO DI NUOVI MICRODATI
7.2.2. MODIFICHE SUGLI INDICATORI
7.2.3. MODIFICHE SULLE TAVOLE
7.2.4. MODIFICHE SULLA STRUTTURA DELL'INDAGINE
7.2.5. MODIFICHE SULLA DOCUMENTAZIONE
8. CONSIDERAZIONI CONCLUSIVE
BIBLIOGRAFIA

Sommario

Il presente documento descrive architettura, tecnologia e funzionalità di una nuova applicazione informatica realizzata per la gestione e l'aggiornamento della banca dati "Sistema Informativo Territoriale sulla Giustizia (*SITG*)", dedicata alla diffusione su web¹ delle statistiche giudiziarie a diversi livelli territoriali.

Lo strumento presentato migliora notevolmente il processo di produzione e diffusione delle tavole statistiche, supportando gli statistici nelle attività di definizione e organizzazione dei contenuti statistici, nonché di gestione del relativo ciclo di vita. Esso consente, infatti, di automatizzare tutte le fasi di elaborazione dei dati a partire dalla definizione delle specifiche fino alla produzione e visualizzazione di indicatori e tavole statistiche.

Il sistema, inoltre, può essere facilmente esportabile anche in contesti di indagine diversi da quelle sulla giustizia per alimentare e gestire altre banche dati.

Abstract

The paper aims to describe architecture, technology and functionalities of new software applications realized for managing and updating the data collection "Judicial statistical information system (*SITG*)", born to disseminate ever more judicial statistics on line¹ by different Italian territorial units.

The management tools improve widely the process of statistical tables production and publication on the Internet, allowing researchers to set up and organize statistical information and manage it throughout its life cycle. Actually several software functions enable to easily define and modify the structure, content, and organisation of published tables and documents. At the same time, a statistical table definition process allows researchers to specify information to be included for each table, rules to be applied for calculation and layouts to be used for presentation.

Besides the system can be used for different surveys in order to create and manage other statistical data banks.

¹ <http://giustiziaincifre.istat.it>

Premessa

L'obiettivo del presente documento è quello di fornire una descrizione dell'architettura, della tecnologia e delle funzionalità proprie del sotto-sistema per la gestione e manutenzione del "Sistema Informativo Territoriale sulla Giustizia"², nonché delle ragioni che ne hanno motivato la realizzazione.

Il resto del documento è strutturato nel modo seguente. Il paragrafo §.1 illustra gli scopi ed i contenuti del sistema informativo territoriale sulla giustizia, mentre il paragrafo §.2 descrive nel dettaglio il processo di produzione e diffusione dell'informazione statistica ad esso relativa. Il paragrafo §.3 fornisce, sia da un punto di vista statistico, sia sotto il profilo informatico ed organizzativo, le motivazioni che hanno portato alla realizzazione del sotto-sistema di gestione e manutenzione. I paragrafi §.4 e §.5 riportano le caratteristiche funzionali e tecniche del sotto-sistema di gestione e manutenzione; il paragrafo §.6 descrive il ciclo di vita dei diversi elementi, mentre il successivo paragrafo §.7 presenta i principali scenari di utilizzo. Il conclusivo paragrafo §.8, infine, illustra i benefici che il sotto-sistema di gestione e manutenzione consente di ottenere nel processo di produzione e diffusione delle tavole statistiche relative alle indagini (attuali e future) del Servizio Giustizia e la possibile applicabilità del sistema stesso ad altri ambiti di indagine.

² Il sistema informativo è stato progettato e realizzato da un team Istat (Deliberazioni n. 28/02/DISS e 53/02/DISS), supportato da risorse informatiche del Raggruppamento Temporaneo di Imprese (RTI), coordinate da Lucia Veltri.

1. Il Sistema Informativo Territoriale sulla Giustizia

Il “Sistema Informativo Territoriale sulla Giustizia” (nel resto del presente documento indicato con l’acronimo SITG) diffonde su web una serie di indicatori statistici relativi alla quantità e qualità della domanda e dell’offerta di giustizia in Italia.

Gli utenti Internet hanno a disposizione un vasto insieme di tavole statistiche predefinite che possono consultare in modo molto flessibile, scegliendo area e materia di interesse, anno di riferimento, contesto e livello di dettaglio territoriale opportuni. La fiancata è costituita generalmente dal territorio, mentre la testata cambia secondo la materia e l’argomento.

In base alla progettazione concettuale statistica, il sistema informativo diffonde informazioni relative alle seguenti aree tematiche³:

- ✓ **Giustizia Civile**
 - ✓ *Separazioni*
 - ✓ *Divorzi*
 - ✓ *Minori: tutela, affidamenti e adozioni*
 - ✓ *Lavoro, previdenza e assistenza*
 - ✓ *Protesti e fallimenti*
 - ✓ *Cognizione e provvedimenti vari*
- ✓ **Giustizia Amministrativa**
 - ✓ *Primo grado – TAR*
 - ✓ *Grado di appello*
- ✓ **Attività Notarile**
- ✓ **Giustizia Penale**
 - ✓ *Suicidi e tentativi di suicidio*
 - ✓ *Delitti denunciati dalle Forze dell’Ordine*
 - ✓ *Criminalità*
 - ✓ *Criminalità minorile*
 - ✓ *Condannati per delitto*
- ✓ **Materia Penitenziaria**

Per ciascuna area sono disponibili un insieme di tavole statistiche ed una serie di informazioni di supporto (schede di indagine, pubblicazioni, normativa, glossario, etc.)

Allo stato attuale, il SITG rende disponibili su web 2.292 indicatori inclusi in 132 famiglie di tavole che, tenendo conto dei possibili contesti e livelli di dettaglio territoriali e degli anni di riferimento disponibili, danno luogo a 2.890 tavole statistiche distinte.

È anche disponibile una sezione *Confronti Internazionali* contenente informazioni relative al funzionamento dei sistemi giudiziari di alcuni Paesi europei ed una serie di tavole statistiche di confronto.

Dal punto di vista informatico, l’architettura del SITG è costituita da due sotto-sistemi principali, che sono descritti nel seguito del presente paragrafo.

1.1. Il sotto-sistema di produzione

Il sotto-sistema di produzione è dedicato alla gestione e manutenzione dell’informazione statistica nell’arco del suo ciclo di vita che, per ciascuna indagine trattata, ha inizio con la disponibilità dei microdati validati e si conclude con la generazione e validazione delle tavole statistiche da diffondere.

³ Il SITG è ovviamente costruito in maniera tale da poter agevolmente essere esteso, in momenti successivi, alla diffusione di informazioni relative ad aree o materie diverse da quelle inizialmente previste.

Tale sotto-sistema è incentrato su una serie di strumenti software di ausilio per il ricercatore nell'elaborazione dei microdati e nella definizione, generazione, validazione e pubblicazione delle tavole previste dal piano statistico di diffusione e della relativa documentazione a corredo:

1. *SITG-Load*⁴: applicativo client/server in ambiente Microsoft Visual Basic per il caricamento dei microdati validati;
2. *SITG-Admin*⁵: applicativo client/server in ambiente Oracle Forms per la gestione e manutenzione delle fonti dati trattate, degli indicatori e delle tavole statistiche;
3. *SITG-Doc*: applicativo web-based basato su tecnologia JSP per la gestione e manutenzione della documentazione a corredo delle tavole statistiche.

Il resto del presente documento è principalmente dedicato all'esame di *SITG-Admin*, applicativo che costituisce il nucleo centrale del sotto-sistema di produzione e lo strumento principale attraverso il quale effettuare tutte le attività di gestione e manutenzione dei contenuti informativi statistici del SITG. Saranno inoltre fornite alcune informazioni in merito a *SITG-Doc* (applicativo attualmente ancora in fase di realizzazione), mentre tutti i dettagli relativi a *SITG-Load* saranno presentati nell'ambito di un successivo lavoro ad esso specificamente dedicato.

1.2. Il sotto-sistema di diffusione

Il sotto-sistema di diffusione (nel resto del presente documento denominato più brevemente *SITG-Web*) permette di rendere disponibile in ambiente intranet ed Internet l'insieme dei contenuti statistici relativi alle indagini sulla giustizia. Esso fornisce contenuti statistici secondo la struttura e le funzionalità di seguito riportate:

1. **Consultazione Dati**. Contiene un insieme prestabilito di tavole statistiche.
 - 1.1. *Dati prima del 2000*. Contiene una serie di tavole che possono essere visualizzate e/o salvate in formato MS Excel.
 - 1.2. *Dati dal 2000*. L'utente può navigare fra le diverse tavole disponibili scegliendo la materia, l'anno di riferimento dei dati e la gerarchia (giudiziaria o amministrativa) ed il livello di dettaglio territoriale. La fiancata di ogni tavola è generalmente il livello territoriale, attraverso il quale è possibile navigare gerarchicamente. Dove previsto, infatti, l'utente può aumentare (ottenendo informazioni più analitiche) o diminuire (ottenendo informazioni di carattere più generale) il livello di dettaglio territoriale. Può inoltre cambiare anno, tavola e materia rimanendo nello stesso contesto territoriale. L'utente ha anche la possibilità di scaricare in formato MS Excel la singola tavola scelta.
2. **Download**. In questa sezione è possibile scaricare in formato MS Excel tutte le tavole statistiche appartenenti ad una specifica materia di interesse e ad un anno.
3. **Confronti Internazionali**⁶. Offre una panoramica sul funzionamento di alcuni sistemi giudiziari europei e sulle attività degli organismi internazionali in materia giudiziaria. Fornisce, inoltre, informazioni per orientare l'utente nell'approfondimento della materia. La sezione si compone di tre aree principali:
 - 3.1. *Paesi*. Descrive l'organizzazione e il funzionamento del sistema giudiziario in cinque paesi europei⁷ mentre la voce "Altri paesi" offre soltanto dei link ai relativi siti web contenenti statistiche e informazioni sulla giustizia

⁴ L'applicativo *SITG-Load* è stato progettato e sviluppato da Marco Broccoli.

⁵ L'applicativo *SITG-Admin* è stato progettato e sviluppato con il supporto di risorse informatiche del Raggruppamento Temporaneo di Imprese (RTI), coordinate da Lucia Veltri.

⁶ La sezione "Confronti Internazionali" è stata realizzata da Maura Steri nell'ambito del progetto "Integrazione e armonizzazione dei dati sulla giustizia in ambito europeo e internazionale" (contratto N. 374/T del 10 dicembre 2002 e contratto N. 352/T del 10/12/2004)

⁷ Attualmente Francia, Germania, Spagna, Inghilterra e Galles

- 3.2. *Confronti statistici*. Quest'area è dedicata alla presentazione di tabelle statistiche di confronto per materia tra i diversi Paesi, corredate da note esplicative sui diversi criteri di rilevazione dei dati.
- 3.3. *Organismi Internazionali*. Illustra le attività statistiche e i progetti condotti dal Consiglio d'Europa, dall'Unione europea e dalle Nazioni unite in materia di giustizia.
4. **Documentazione**. Si tratta della sezione in cui sono contenute tutte le informazioni di supporto per la consultazione delle tavole statistiche presenti nella sezione *Consultazione Dati*. Tali informazioni appartengono alle categorie di seguito elencate:
- 4.1. *Schede Informative di Indagine*.
 - 4.2. *Classificazioni*.
 - 4.3. *Avvertenze*.
 - 4.4. *Normativa*.
 - 4.5. *Pubblicazioni ISTAT*.
 - 4.6. *Glossario*.
 - 4.7. *Istruzioni agli uffici giudiziari*.
 - 4.8. *Altri dati ISTAT sulla giustizia*.

Relativamente alla consultazione delle sezioni "Dati dal 2000", "Download" e "Documentazione", *SITG-Web* costituisce, un'evoluzione del sistema *DaWinci-PD*⁸, esteso con una serie di funzionalità⁹.

⁸ Web warehouse originariamente sviluppato nell'ambito del censimento della popolazione e delle abitazioni del 2001 per la pubblicazione su *web* dei dati provvisori e di quelli relativi alla popolazione legale.

⁹ Ad esempio navigazione per gerarchie concettuali dei dati e dei metadati, navigazione temporale dei dati, navigazione per gerarchie territoriali multiple, documentazione contestualizzata rispetto alla tavola statistica selezionata, download di tutte le tavole disponibili per ciascuna materia.

2. Il processo di produzione e diffusione

2.1. Il processo senza uno strumento di gestione

Prima di analizzare l'architettura e le funzionalità del sotto-sistema di gestione e manutenzione del SITG, risulta utile illustrare nel dettaglio il processo di produzione e diffusione delle indagini sulla giustizia.

Nell'ambito del progetto di realizzazione del SITG, conclusosi nel novembre 2004, è stato possibile analizzare dettagliatamente e nell'ambito di molte indagini relative al Servizio Giustizia, il processo di gestione della tavole statistiche che, a partire dalla disponibilità dei microdati validati per una specifica indagine, consentiva di definire, produrre, validare e pubblicare un insieme di tavole statistiche in ambiente web. Tale processo, preceduto da una indispensabile fase di progettazione concettuale statistica, era stato inizialmente modellato nel modo seguente:

0. *Progettazione concettuale statistica.* Si tratta del cuore dell'attività di analisi statistica, nell'ambito della quale vengono ideati, definiti, calcolati, verificati e validati gli indicatori e le tavole statistiche che ad essi fanno riferimento. Tale attività viene svolta anche tramite il supporto di alcuni strumenti software come SAS, Access o SQL*Plus, che consentono allo statistico di lavorare direttamente il file dei microdati validati. Una volta definiti tutti gli indicatori relativi ad una specifica indagine, questi vengono opportunamente composti all'interno di un insieme di tavole statistiche, dando così luogo ad un collezione di specifiche (inizialmente prodotte nella forma di un documento .xls) da utilizzare nei successivi passi 3-7.
1. *Caricamento dei microdati.* I microdati validati relativi ad una specifica indagine (tipicamente disponibili su file testuale piatto, in formato .xls o MS Access) vengono caricati nella base di dati del sistema informativo tramite l'applicativo *SITG-Load*, che provvede a popolare la corrispondente porzione della base di dati. Questo passo del processo, ovviamente, può avvenire soltanto in seguito al disegno ed alla creazione di un'opportuna componente dello schema della base di dati, a partire dalla disponibilità del tracciato record e del modello di indagine.
2. *Produzione dei macrodati.* I microdati caricati nella base di dati vengono elaborati tramite l'esecuzione di un insieme di stored procedures PL/SQL, tramite le quali vengono prodotti e memorizzati nella base di dati i corrispondenti macrodati. Tali macrodati saranno poi utilizzati come base per il calcolo degli indicatori che faranno parte delle tavole statistiche. Anche questo passo, come il precedente, richiede la preventiva definizione e creazione di un'apposita componente dello schema della base di dati.
3. *Definizione degli indicatori.* Questo passo ed i due successivi (4 e 5) fanno riferimento al processo di definizione delle specifiche informatiche relative a tutti gli indicatori e le tavole concepite nella fase di progettazione concettuale statistica. Si tratta, quindi, di una mera attività di formalizzazione dei risultati ottenuti nel corso del precedente passo 0. Nei passi 3 e 4, ciascun indicatore relativo all'indagine considerata viene definito¹⁰, calcolato, verificato e validato, sempre con l'ausilio degli strumenti software precedentemente menzionati.
4. *Calcolo degli indicatori.* In questa fase un insieme di stored procedures PL/SQL determina i valori associati ai vari indicatori applicando le specifiche di calcolo di cui al passo precedente.

¹⁰ In questo caso, la fase di definizione consiste nella mappatura delle specifiche fornite rispetto al file dei microdati nella corrispondente struttura all'interno della base di dati.

5. *Definizione delle tavole.* In questo passo, in modo del tutto analogo rispetto a quanto avviene per gli indicatori, le specifiche fornite per le tavole statistiche vengono mappate in un insieme di specifiche informatiche rispetto alla struttura della base di dati.
6. *Codifica delle tavole.* Sulla base delle specifiche formalizzate nei passi precedenti, viene attivata una fase di sviluppo software che, nella sostanza, produce un insieme di stored procedures PL/SQL che implementano le specifiche di cui sopra.
7. *Generazione delle tavole.* Questo passo consiste nell'esecuzione del codice sviluppato al passo 6 e nella memorizzazione all'interno della base di dati di tutti i valori calcolati per le tavole definite nelle specifiche. Ciascuna tavola prodotta può essere a questo punto visualizzata in ambiente intranet, secondo un'interfaccia utente identica a quella del sistema di consultazione disponibile su internet. In tal modo, è possibile verificare che i risultati ottenuti per ciascuna tavola siano, sotto tutti i punti di vista, pari a quelli attesi.
8. *Pubblicazione delle tavole.* Questo passo consente di trasferire i dati di ciascuna tavola relativa all'indagine considerata dal database server interno (rispetto al quale è stata effettuata la verifica di cui al passo 7) verso il database server esposto su internet, così da rendere di fatto disponibili agli utenti le tavole in questione.

All'interno del processo appena descritto sono state rilevate, nell'ambito delle indagini del Servizio Giustizia, alcune criticità che vengono di seguito descritte:

- ✓ *Produzione dei macrodati.* I problemi legati a questa fase del processo erano sostanzialmente connessi ad eventuali errori presenti nelle procedure software di produzione dei macrodati. La questione principale, in questo caso, era che eventuali errori nelle procedure software non venivano facilmente rilevati, se non al momento della visualizzazione delle tavole statistiche, cioè nella fase finale del processo sopra descritto. Per tale ragione, diventava di primaria importanza il tentativo di ridurre al minimo il lasso di tempo intercorso fra la disponibilità dei macrodati di una specifica indagine (al termine del passo 2) e la possibilità di visualizzare le relative tavole statistiche (passo 7). Come si vedrà, uno dei benefici principali del sistema di gestione e manutenzione è stato proprio quello di consentire la riduzione di tale intervallo di tempo.
- ✓ *Codifica delle tavole.* Il passo di codifica era fondamentalmente incentrato su due aspetti, da un lato quello del calcolo degli indicatori che fanno parte della tavola a partire dai macrodati dell'indagine, dall'altro quello della produzione di tutte le informazioni a livello di sistema che consentono alla componente di visualizzazione web di presentare ciascuna tavola per tutti i livelli di dettaglio territoriale previsti. Entrambi gli aspetti erano fortemente legati ad una componente manuale dell'attività, la prima legata al vero e proprio processo di codifica a livello software delle specifiche statistiche, la seconda direttamente dovuta ad un oneroso processo di data entry. Ciò significa che la probabilità che, nella fase di trasformazione delle specifiche fornite nelle vere e proprie tavole statistiche, si potessero generare errori risultava, potenzialmente, molto elevata. Tali attività di sviluppo software e data entry, inoltre, richiedevano l'intervento di risorse informatiche che, in ogni caso, potevano produrre un risultato da sottoporre al vaglio delle risorse statistiche in tempi non certo brevi (e che potevano superare la soglia del mese solare ed oltre).

Il motivo principale per il quale le criticità espresse rappresentavano una notevole fonte di inefficienza per il processo di gestione delle tavole statistiche sopra descritto, era dato dal fatto che eventuali errori potevano essere rilevati soltanto al momento della visualizzazione delle tavole per la verifica, indipendentemente dal momento in cui questi si fossero prodotti. Ciò si doveva, oltre che a quanto sopra esposto, anche

al fatto che le specifiche delle tavole venivano tipicamente prodotte a valle di un'attività statistica con lavorazione dei microdati incentrata sui files piatti e non sul corrispondente contenuto della base di dati, il che, ovviamente, rappresentava un problema nel caso in cui si verificassero errori nella fase di caricamento dei microdati o di produzione dei macrodati. Il fatto di utilizzare direttamente i files piatti per il calcolo degli indicatori e la definizione delle specifiche, per altro, poteva anche essere fonte di errori soltanto "apparenti". Questo accadeva quando, al momento della verifica di una tavola, si rilevava un disallineamento fra i valori calcolati a partire dai microdati disponibili su file piatto e quelli memorizzati nella base di dati, scoprendo, soltanto a valle di un'attenta analisi, che l'errore era stato commesso durante la fase di definizione delle specifiche.

In tutti i casi, comunque, gli eventuali errori potevano essere rilevati anche a distanza di settimane e la loro rimozione poteva comportare uno o più ricicli che, nel caso peggiore, potevano ripartire addirittura dal primo passo del processo, quello di caricamento dei microdati.

2.2. Il processo con uno strumento di gestione

È proprio per risolvere le questioni appena esposte che si è pensato di realizzare un'applicazione ad hoc per la gestione e manutenzione del sistema informativo in grado di migliorare l'efficienza del processo sopra descritto e di supportare lo statistico dal momento in cui sono disponibili i microdati validati, fino alla produzione e diffusione delle tavole stesse. Tale sistema è nato con l'obiettivo di rendere per quanto possibile automatiche (e quindi svincolate da interventi a carico di risorse informatiche) le fasi di elaborazione dei dati che vanno dal momento della definizione delle specifiche fino alla produzione e visualizzazione della tavole statistiche, cercando di mantenere inalterato il grado di indipendenza dello statistico rispetto agli strumenti informatici da utilizzare per le proprie attività di ideazione, definizione, calcolo, verifica e validazione delle tavole stesse.

In tal modo, il processo di produzione e diffusione risulta modificato come di seguito descritto (cfr. fig.1):

0. *Progettazione concettuale statistica.* Questo passo rimane inalterato rispetto alla precedente descrizione del processo.
1. *Caricamento dei microdati.* Questo passo rimane inalterato rispetto alla precedente descrizione del processo.
2. *Produzione dei macrodati.* Questo passo rimane inalterato rispetto alla precedente descrizione del processo, anche se le relative procedure vengono attivate tramite il sistema di gestione e manutenzione.
3. *Definizione delle variabili.* A questo punto, un utente informatico che sia a conoscenza della struttura della base di dati sottostante, provvede a definire, per l'indagine considerata, un insieme di variabili statistiche, ciascuna delle quali direttamente associata ad una o più tabelle presenti nella componente macrodati della base di dati. In questo modo, gli utenti statistici potranno agire indirettamente sui macrodati contenuti nella base di dati tramite un'interazione diretta con le variabili statistiche dell'indagine in questione.
4. *Definizione degli indicatori.* In questa fase l'utente statistico procede alla definizione dell'insieme degli indicatori relativi all'indagine in esame, individuati in fase di progettazione concettuale. Per ciascuno di essi, deve essere fornita tutta una serie di informazioni (come il formato numerico, la regola di calcolo, le gerarchie ed i livelli territoriali e gli intervalli temporali di validità) che permettono al sistema di determinarne tutti i valori associati.

- Al termine del passo di definizione, l'utente può decidere se limitarsi a salvare la definizione stessa, ovvero calcolare e visualizzare direttamente l'indicatore in esame, per verificarne la validità.
5. *Definizione delle tavole.* L'utente statistico procede alla definizione delle tavole statistiche relative all'indagine considerata, componendo gli indicatori disponibili secondo criteri di omogeneità rispetto alle materie ed agli argomenti trattati. A tale scopo, per ciascuna tavola devono essere forniti tutti gli elementi costitutivi, ovvero, nella sostanza, il titolo, la lista ordinata degli indicatori (con i relativi titoli di testata corrispondenti) e l'insieme delle note esplicative.
 6. *Generazione delle tavole.* Questo passo consiste nell'esecuzione di un insieme di procedure software che, a partire dai macrodati dell'indagine in esame, calcola il valore di tutti gli indicatori¹¹ che fanno parte delle tavole coinvolte e provvede a popolare le opportune tabelle della base di dati con i valori così ottenuti. A questo punto, l'utente statistico può accedere alla versione disponibile sulla intranet dell'Istituto del sotto-sistema di diffusione, per poter visualizzare tutte le tavole statistiche prodotte e verificarne la correttezza.
 7. *Pubblicazione delle tavole.* In questo modo l'utente statistico ha modo di rendere visibili su Internet le tavole già definite, prodotte, visualizzate e verificate in ambiente intranet. In sostanza, viene eseguita una procedura software che provvede a replicare i dati relativi alle tavole da pubblicare dal database server interno a quello esposto su Internet.

Confrontando la struttura dei due processi illustrati (prima e dopo l'introduzione del sistema di gestione e manutenzione) risulta abbastanza facile rilevare i vantaggi connessi all'introduzione dell'applicazione in esame. In primo luogo, si riduce drasticamente il tempo che intercorre tra la definizione delle specifiche e la rilevazione di eventuali errori. Tale tempo, infatti, non deve più tener conto di alcuna attività manuale di sviluppo software o data entry (perché non esiste più il passo di *codifica delle tavole*), ma soltanto dei tempi di esecuzione di alcune stored procedures PL/SQL generate automaticamente dal sistema. Ciò consente allo statistico di poter visualizzare le proprie tavole, direttamente nel formato che sarà reso disponibile agli utenti del sistema informativo, non *settimane*, ma soltanto *ore* (o *minuti*) dopo aver terminato la fase di definizione delle specifiche e senza alcun intervento da parte di risorse informatiche.

L'unico intervento aggiuntivo richiesto da parte di risorse informatiche è relativo al nuovo passo 3 (di *definizione delle variabili*), ma si tratta comunque di un intervento di modesta entità e che può essere considerato come una banale appendice delle attività informatiche di disegno e creazione di basi di dati di cui ai passi 1 e 2.

Come già anticipato, l'impatto del sistema rispetto alle modalità organizzative e di lavoro degli statistici può essere mantenuto molto limitato, perché l'unica attività che deve essere obbligatoriamente svolta in modo diverso rispetto al precedente processo di gestione è quello della definizione degli indicatori e delle tavole che consisteva nella produzione di un documento .xls e che, in presenza del sistema di gestione e manutenzione, viene effettuata inserendo le stesse informazioni tramite alcune semplici forms in ambiente *Oracle Developer*¹². Per il resto, tutte le altre attività statistiche di ideazione, calcolo, verifica e validazione degli indicatori (e delle tavole) possono essere svolte nel modo tradizionale e con gli strumenti ritenuti più opportuni. Si tenga presente, tuttavia, che utilizzare il sistema di gestione e manutenzione come strumento software di supporto per tali attività comporterebbe l'ulteriore vantaggio di eliminare completamente i cosiddetti errori "apparenti" precedentemente citati, dal

¹¹ A meno che tale calcolo non sia già stato effettuato al termine del precedente passo 4.

¹² Vale la pena di notare come, a fronte di un leggero cambiamento nella modalità operativa di produzione delle specifiche, si ottenga un notevole vantaggio in termini di efficienza, dal momento che con l'utilizzo del sistema di gestione le specifiche entrano direttamente a far parte della base di conoscenza del sistema informativo e possono essere direttamente utilizzate per la produzione delle tavole, senza alcun intervento manuale di sviluppo software o data entry.

momento che tutte le attività di calcolo sarebbero in tal caso effettuate su una sola copia dei dati (quella contenuta nella base di dati). Inoltre, l'utilizzo del sistema consentirebbe di ridurre drasticamente il tempo di produzione delle specifiche, dal momento che non sarebbe più necessario, per le verifiche sui valori assunti dagli indicatori in esame, scrivere codice (SAS, SQL, etc.) come accade attualmente. Infine, dal momento che tutte le attività che seguono alla definizione delle specifiche sono eseguite direttamente dal sistema di gestione e manutenzione, la probabilità che si verifichino errori in questa fase è, di fatto, ridotta a zero.

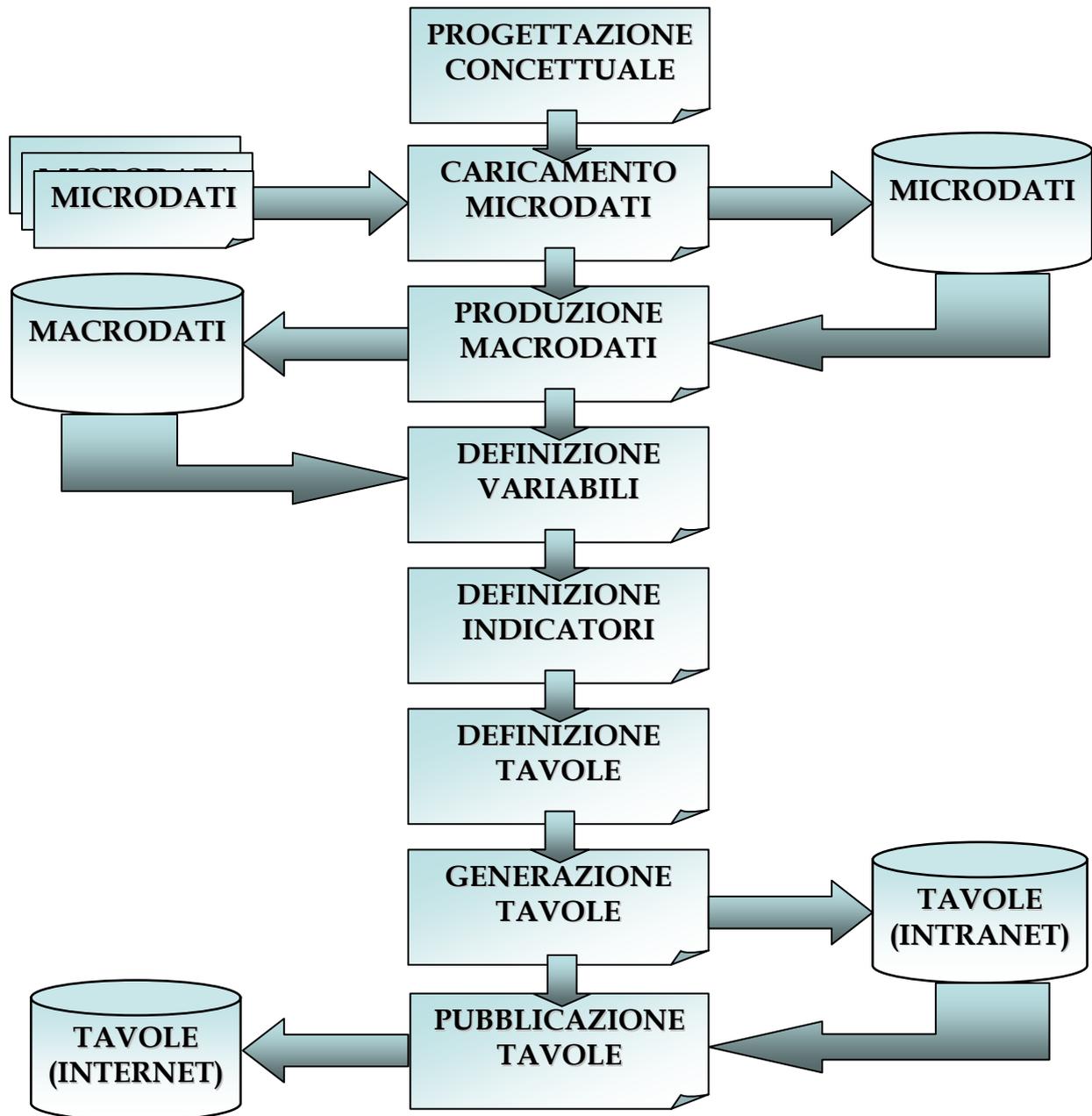


Fig.1 – SITG: il processo di produzione e diffusione.

3. Il sistema di gestione e manutenzione del SITG

Come risulta evidente dall'analisi del processo di produzione e diffusione di cui al precedente paragrafo §.2, la manutenzione e l'aggiornamento del sistema informativo in esame sono questioni assolutamente non marginali rispetto alla sua creazione iniziale; al contrario, ne garantiscono la sopravvivenza con un necessario carattere di attualità ed efficacia informativa.

Il SITG richiede, pertanto, una gestione interattiva, flessibile, efficiente ed efficace, non soltanto delle attività relative alla creazione dei diversi contenuti statistici (indicatori, tavole, documentazione, etc.), ma anche di quelle connesse alla variabilità nel tempo di tali contenuti.

Queste considerazioni hanno condotto alla realizzazione di un vero e proprio sistema informatico autonomo e finalizzato alla gestione e manutenzione delle diverse componenti informative statistiche del sistema, ad esclusione della sezione dedicata ai confronti internazionali. Tale sistema informatico, da un'altra prospettiva, costituisce un ponte tra la progettazione concettuale statistica dei contenuti informativi e le attività informatiche che ne consentono la diffusione su web, affinché il ricercatore sia sufficientemente autonomo in tutte le attività di inserimento e aggiornamento di tutti gli elementi informativi coinvolti.

Il sistema in questione, inoltre, semplifica la già accennata possibilità di estendere il sistema informativo ad altre aree di indagine attualmente non previste, consentendo l'immediata "traduzione" in procedure software della progettazione e definizione dei nuovi contenuti statistici. Per altro, ciò rende il sistema facilmente esportabile anche in contesti del tutto diversi da quello relativo alle indagini sulla giustizia.

Nei successivi paragrafi §.4 e §.5 vengono presentate le funzionalità e le caratteristiche tecniche del sistema di gestione e manutenzione, con particolare riferimento, come già accennato, alle due componenti *SITG-Admin* e *SITG-Doc*.

4. Descrizione funzionale del sistema di gestione e manutenzione

Il presente paragrafo è dedicata alla descrizione delle principali funzionalità del sistema di gestione e manutenzione. Tali funzionalità sono state suddivise nelle seguenti aree principali:

- ✓ *Gestione delle utenze*: contiene tutte le funzionalità relative alla gestione degli utenti e dei relativi profili di accesso all'interno del sistema.
- ✓ *Gestione delle fonti dati*: contiene tutte le funzionalità relative alla gestione delle fonti dati definite all'interno del sistema¹³.
- ✓ *Gestione delle categorie*: contiene tutte le funzionalità relative alla gestione dell'organizzazione dell'informazione statistica (dati e documentazione) all'interno del sotto-sistema di diffusione *SITG-Web*, secondo criteri di suddivisione dell'informazione stessa in aree e materie, con eventuali ulteriori livelli di raffinamento.
- ✓ *Gestione degli indicatori*: contiene tutte le funzionalità relative alla gestione degli indicatori statistici.
- ✓ *Gestione delle tavole statistiche*: contiene tutte le funzionalità relative alla gestione delle tavole statistiche, dalla fase di definizione a quella di diffusione su web.
- ✓ *Gestione della documentazione*: contiene tutte le funzionalità relative alla gestione della documentazione associata alle tavole statistiche (schede di indagine, norme legislative, pubblicazioni, classificazioni, etc.)¹⁴.

Come è lecito attendersi, non tutte le funzionalità possono essere utilizzate da tutti gli utenti. Al contrario, all'interno del sistema di gestione sono previsti tre diversi ruoli che, al momento della definizione di un nuovo utente, consentono di attribuire a tale utente tutti e soli i privilegi che si ritengono necessari. I tre ruoli previsti sono:

- ✓ *gestore*: consente di utilizzare le funzionalità di gestione delle utenze;
- ✓ *amministratore*: consente di utilizzare le funzionalità di gestione delle fonti dati;
- ✓ *statistico*: consente di utilizzare le funzionalità per la gestione dell'informazione statistica (categorie, indicatori, tavole e documentazione).

4.1. Funzionalità per la gestione delle utenze

Tali funzionalità possono essere utilizzate soltanto da un utente cui sia stato attribuito il ruolo di *gestore* e consentono di creare un utente, modificare o eliminare un utente precedente definito e attribuire a ciascun utente un opportuno profilo di accesso. Tale profilo di accesso è determinato, in sostanza, dal ruolo attribuito all'utente (*gestore*, *amministratore*, *statistico*) e dai privilegi di accesso alle categorie (aree, materie e sotto-materie) secondo le quali risultano organizzati i contenuti statistici.

¹³ Tali funzionalità sono incluse in *SITG-Admin*, ad eccezione di quelle relative al caricamento dei microdati validati, che sono rese disponibili da *SITG-Load*.

¹⁴ Tali funzionalità sono incluse in *SITG-Doc*, attualmente in fase di sviluppo.

Nome	Descrizione
CreaUtente	Definisce un nuovo utente del sistema. A tale utente, identificato tramite un identificativo e una password, possono essere attribuiti uno o più dei profili sopra descritti.
ModificaUtente	Permette di modificare una o più caratteristiche (password, profilo, etc.) di un utente precedentemente definito nel sistema.
EliminaUtente	Permette di eliminare un utente precedentemente definito nel sistema. Questa funzionalità può essere attivata solo per utenti che non siano proprietari di alcun oggetto (indicatori, tavole, etc.).

4.2. Funzionalità per la gestione delle fonti dati

Tali funzionalità possono essere utilizzate soltanto da un utente cui sia stato attribuito il ruolo di *amministratore* e consentono di creare una nuova fonte dati, ovvero di modificare, validare o eliminare una fonte dati precedentemente definita. Nella creazione o nella modifica di una fonte dati, l'attività principale consiste nella definizione di un'opportuna corrispondenza fra le tabelle che, nello schema della base di dati, sono associate alla fonte dati stessa e le variabili statistiche ad essa relative. Una volta conclusa e verificata la definizione di una nuova fonte dati, l'utente ha la possibilità di dichiararla *valida* il che, nella sostanza, la rende disponibile agli utenti statistici per le successive attività di definizione di nuovi indicatori e tavole statistiche su tale fonte dati.

Nome	Descrizione
CreaFonteDati	Definisce una nuova fonte dati nel sistema. Nel corso del processo di definizione della fonte dati, l'utente amministratore stabilisce una serie di corrispondenze fra le tabelle dei fatti e delle dimensioni presenti nella base di dati e relative alla fonte dati considerata e l'insieme delle variabili quantitative e qualitative che saranno rese disponibili all'utente statistico.
ModificaFonteDati	Permette di modificare uno o più elementi di una fonte dati precedentemente definita nel sistema.
EliminaFonteDati	Permette di eliminare una fonte dati precedentemente definita nel sistema. Questa funzionalità può essere attivata solo per fonti dati sulle quali non siano definiti indicatori.
ValidaFonteDati	Si tratta dell'ultimo passo di definizione della fonte dati e la rende disponibile agli utenti statistici, i quali potranno definire su di essa indicatori e tavole statistiche.
CaricaFonteDati	Effettua il caricamento della fonte dati a partire dai microdati relativi ad uno specifico anno di riferimento.

4.3. Funzionalità per la gestione delle categorie

Tali funzionalità possono essere utilizzate soltanto da un utente cui sia stato attribuito il ruolo di *statistico* e consentono di creare una nuova categoria, ovvero di modificare, spostare (all'interno della struttura ad albero delle categorie) o eliminare una categoria precedentemente definita.

Nome	Descrizione
CreaCategoria	Definisce una nuova categoria (area, materia o sotto-materia) nel sistema. Tale funzionalità consente di creare e mantenere la struttura ad albero secondo la quale i dati e la relativa documentazione risultano organizzate all'interno del <i>SITG-Web</i> .
ModificaCategoria	Permette di modificare uno o più elementi di una categoria precedentemente definita nel sistema.
EliminaCategoria	Permette di eliminare una categoria precedentemente definita nel sistema. Questa funzionalità può essere attivata solo per categorie all'interno delle quali non siano definiti oggetti (indicatori, tavole, etc.) o altre categorie.
SpostaCategoria	Permette di spostare una categoria precedentemente definita nel sistema da un punto all'altro dell'albero delle categorie.

4.4. Funzionalità per la gestione degli indicatori

Tali funzionalità possono essere utilizzate soltanto da un utente cui sia stato attribuito il ruolo di *statistico* e consentono di creare un nuovo indicatore, ovvero di modificare, calcolare o eliminare un indicatore precedentemente definito.

Utilizzando tali funzionalità, il ricercatore può definire indicatori statistici *assoluti* o *derivati*. Nel primo caso si tratta di indicatori direttamente definiti su una o più variabili di una fonte dati; nel secondo caso, al contrario, l'indicatore viene definito come funzione di uno o più indicatori (assoluti o derivati) precedentemente definiti, eventualmente su fonti dati distinte.

Per ciascun indicatore, inoltre, il sistema consente di indicare il formato numerico da utilizzare (intero o decimale) di specificare le gerarchie ed i livelli di dettaglio territoriali disponibili, indicando per ciascuna gerarchia e per ciascun livello di dettaglio il relativo intervallo temporale di validità.

Per ciascun indicatore definito, è poi possibile calcolare i valori assunti su tutte le gerarchie ed i livelli di dettaglio territoriali disponibili, semplicemente specificando l'anno di riferimento rispetto al quale effettuare il calcolo.

Un'ultima funzionalità molto interessante nell'area della gestione degli indicatori è quella relativa alla gestione delle eccezioni che, nella sostanza, consente di gestire situazioni nelle quali un indicatore, per un dato anno di riferimento e per uno specifico contesto e livello di dettaglio territoriale, debba assumere valori diversi da quelli calcolati mediante la semplice applicazione delle specifiche di calcolo.

Nome	Descrizione
CreaIndicatore	Definisce un nuovo indicatore. Un indicatore può essere di tipo assoluto o derivato. Nel primo caso si tratta di un indicatore direttamente definito su una o più variabili di una fonte dati; nel secondo caso, al contrario, l'indicatore viene definito come funzione di uno o più indicatori (assoluti o derivati) precedentemente definiti. Per ogni indicatore definito, il sistema memorizza la procedura per effettuare il calcolo (sia che si tratti di una semplice query SQL, sia nel caso in cui si tratti di un'espressione matematica contenente, eventualmente, una o più istruzioni SQL).
ModificaIndicatore	Permette di modificare uno o più elementi di un indicatore precedentemente definito nel sistema.
EliminaIndicatore	Permette di eliminare un indicatore precedentemente definito nel sistema. Questa funzionalità può essere attivata solo per indicatori che non facciano parte di alcuna tavola statistica.
CalcolaIndicatore	Calcola e visualizza, per uno specifico anno di riferimento, tutti i valori associati ad un indicatore precedentemente definito. L'utente ha anche la possibilità di salvare i risultati ottenuti in un file.
Eccezioni	Permette di associare manualmente ad un indicatore qualsiasi tipo di eccezione ("-", "..", "...", "n.d.", "n.c." o uno specifico valore numerico), per uno o più anni e unità territoriali (identificate da gerarchia, livello di dettaglio e contesto) di riferimento ¹⁵ .

4.5. Funzionalità per la gestione delle tavole statistiche

Tali funzionalità possono essere utilizzate soltanto da un utente cui sia stato attribuito il ruolo di *statistico* e consentono di creare una nuova tavola statistica, ovvero di modificare, eliminare, generare o pubblicare una tavola precedentemente definita.

Per ogni tavola, l'utente statistico ha la possibilità di definire tutti gli elementi costitutivi: nome, testata (intesa sia come etichette da assegnare alle colonne di testata, sia come indicatori da associare a ciascuna colonna) e note.

Per ciascuna tavola definita, è poi possibile, semplicemente specificando l'anno di riferimento, effettuare la generazione, che consiste nel calcolare tutti i valori assunti su tutte le gerarchie ed i livelli di dettaglio territoriali disponibili per gli indicatori appartenenti alla tavola in questione. Al termine della fase di generazione, la tavola può essere visualizzata tramite il sotto-sistema di diffusione disponibile sulla intranet dell'Istituto (*SITG-Web*) ed essere così verificata e validata dal ricercatore.

Rispetto a ciascuna tavola validata, può essere poi attivata una funzionalità di pubblicazione, il cui risultato consiste semplicemente nell'esecuzione di una procedura software che provvede a replicare i dati relativi alla tavola in questione dal database server interno a quello esposto su Internet.

Vale la pena di sottolineare, infine, come il sistema permetta anche di gestire più versioni per una singola tavola, cosa che concede al ricercatore la possibilità di variare nel tempo gli elementi costitutivi della tavola stessa (titolo, etichette della testata, insieme degli indicatori coinvolti, note, etc.).

¹⁵ Il significato dei simboli, riportato nelle avvertenze, è il seguente:

- *linea (-)*: quando il fenomeno non esiste; quando il fenomeno esiste e viene rilevato, ma i casi non si sono verificati; quando il dato relativo deriva da un rapporto tra due frequenze nulle
- *due punti (.)*: per i numeri, assoluti o relativi, che non raggiungono la metà della cifra significativa dell'ordine minimo considerato
- *quattro punti (...)*: quando il fenomeno esiste, ma i dati non si conoscono per qualsiasi ragione
- *n.d.*: non diffondibile, quando il dato non può essere diffuso
- *n.c.*: non calcolabile, quando il dato relativo deriva da un rapporto in cui il numeratore è maggiore di zero, ma il denominatore è uguale a zero.

Nome	Descrizione
CreaTavola	Definisce una nuova tavola statistica relativa ad una specifica categoria (area, materia o sotto-materia). Per ogni tavola, l'utente statistico ha la possibilità di definire tutti gli elementi costitutivi: titolo, testata, gerarchie e livelli territoriali disponibili (che saranno utilizzati per la generazione automatica delle fiancate), note.
ModificaTavola	Permette di modificare uno o più elementi di una tavola statistica precedentemente definita nel sistema.
EliminaTavola	Permette di eliminare una tavola statistica precedentemente definita nel sistema.
GeneraTavola	Calcola, per uno specifico anno di riferimento, tutti i valori associati a tutti gli indicatori presenti nella tavola in questione (nel caso in cui questi non siano già stati precedentemente calcolati) e rende così la tavola consultabile in ambiente intranet tramite il sistema di navigazione <i>SITG-Web</i> . In tal modo, l'utente statistico ha la possibilità di vedere il risultato del proprio lavoro esattamente nello stesso modo (sia sotto il profilo dei dati, sia dal punto di vista grafico) in cui sarà, eventualmente, pubblicato su Internet.
PubblicaTavola	Rende disponibile all'utenza Internet una tavola statistica per uno o più anni di riferimento.

4.6. Funzionalità per la gestione della documentazione

SITG offre alla comunità Internet non soltanto un set di tavole statistiche, ma anche tutta una serie di informazioni a corredo dei dati pubblicati (si veda, a tale proposito, quanto riportato in §.1.2). Nella sostanza, le informazioni pubblicate possono consistere in semplici ipertesti all'interno di una normale pagina web (è quanto si verifica per le *avvertenze*, le *classificazioni*, le *istruzioni agli uffici giudiziari* e gli *altri dati ISTAT sulla giustizia*), ovvero una serie di elementi informativi (denominati *oggetti*) consultabili singolarmente. È questo il caso delle *schede informative di indagine*, della *normativa*, delle *pubblicazioni ISTAT*, a loro volta suddivise in *pubblicazioni* ed *editoria divulgativa*, e delle voci di *glossario*.

Le funzionalità per la gestione della documentazione consentono di creare, modificare ed eliminare qualsiasi tipo di oggetto. In generale, a ciascun oggetto è associato un *identificativo* univoco, una *classe* (normativa, pubblicazione, voce di glossario, etc.) ed una *categoria* (area civile, area penale, etc.) di appartenenza e una serie di *proprietà* che ne caratterizzano il contenuto informativo¹⁶. Al momento della creazione di un nuovo oggetto (o della modifica di un oggetto già pubblicato), il sistema consente di definirne (o ridefinirne), attraverso un'interfaccia utente del tutto analoga a quella del sistema di diffusione, tutti gli elementi costitutivi e di effettuarne immediatamente la pubblicazione su web.

¹⁶ Nel caso delle voci di glossario, inoltre, è possibile associare il singolo oggetto ad un preciso insieme di tavole, realizzando così un glossario sensibile al contesto che, nella consultazione di una generica tavola, consente di visualizzare tutte e sole le voci di glossario pertinenti alla tavola stessa.

SISTEMA INFORMATIVO TERRITORIALE SULLA GIUSTIZIA:
strumenti di gestione e manutenzione

Nome	Descrizione
CreaOggetto	Definisce un nuovo oggetto, permettendo all'utente di definirne tutte le caratteristiche (classe e categoria di appartenenza, proprietà, etc.). Al termine dell'operazione, l'oggetto risulta visibile tramite il sistema navigazione disponibile in intranet.
ModificaOggetto	Permette di modificare uno o più elementi di un oggetto precedentemente definito nel sistema.
EliminaOggetto	Permette di eliminare un oggetto precedentemente definito nel sistema. Questa funzionalità può essere attivata solo per oggetti non pubblicati.
PubblicaOggetto	Rende disponibile l'oggetto selezionato alla comunità Internet.

5. Descrizione tecnica del sistema di gestione e manutenzione

5.1. Architettura del sistema

Il sistema di gestione e manutenzione, come già anticipato, risulta costituito dai tre sotto-sistemi *SITG-Load*, *SITG-Admin* e *SITG-Doc*. I primi due sotto-sistemi presentano un'architettura software di tipo client-server (cfr. fig.2.b); in particolare, la componente client di *SITG-Load* è stata sviluppata in ambiente *Microsoft Visual Basic*, mentre quella di *SITG-Admin* si basa sulla tecnologia *Oracle Forms*. Entrambi gli applicativi utilizzano come componente server il *RDBMS Oracle*, componente all'interno della quale, nel caso di *SITG-Admin*, risiede buona parte della logica applicativa del software, sotto forma di stored procedures PL/SQL.

Il sotto-sistema *SITG-Doc*, al contrario, sfrutta un'architettura three-tiers del tutto analoga a quella del sotto-sistema di diffusione *SITG-Web* e che prevede l'utilizzo di *Oracle RDBMS* come database server, di *Apache/Tomcat* e della tecnologia *JSP* come application server e di un semplice web browser come interfaccia utente (cfr. fig.2.a).

In ogni caso, tutte le componenti del sistema di gestione e manutenzione e lo stesso sotto-sistema di diffusione condividono una comune architettura dei dati, che viene di seguito illustrata.

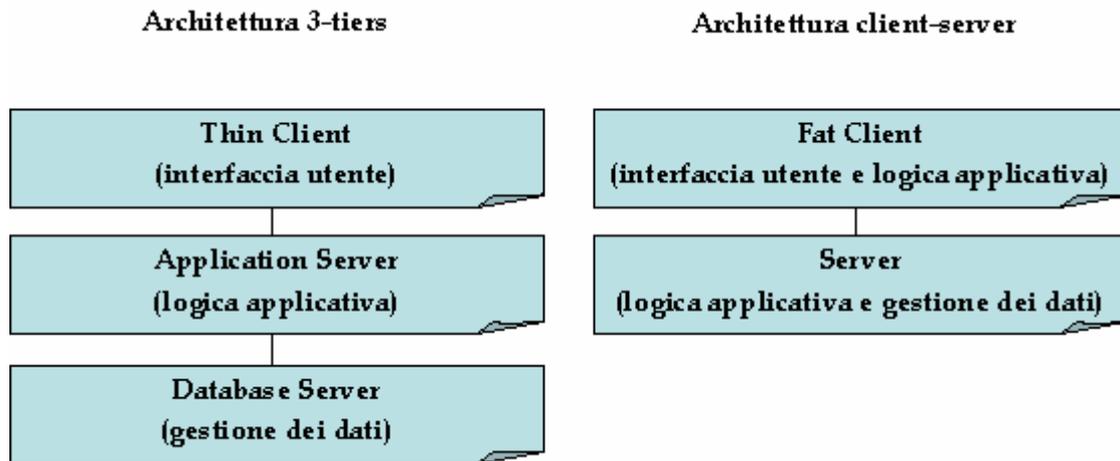


Fig.2.a

Fig.2.b

5.2. Architettura dei dati

La componente dati di SITG è costituita da due basi di dati, la prima della quali include l'intero contenuto informativo del sistema ed è ospitata su un server protetto, mentre la seconda contiene soltanto le informazioni pubblicate su Internet e si trova su un server esposto. Scendendo maggiormente in dettaglio, la componente dati del sistema risulta partizionata nelle quattro aree di seguito riportate:

- ✓ microdati,
- ✓ macrodati,
- ✓ knowledge base,
- ✓ tavole e documentazione

Nel resto del presente paragrafo vengono illustrati i dettagli delle aree ora individuate.

5.2.1. **Microdati**

L'area dei microdati è costituita da un insieme di sotto-schemi, uno per ciascuna fonte dati. Ciascun sotto-schema presenta una struttura variabile in funzione della complessità della fonte dati (indagine) cui si riferisce ma che, in generale, si fonda sulle ormai tradizionali strutture *a stella* e *a fiocco di neve*, tipiche dei sistemi di data warehouse. A titolo di esempio, in fig.3 è riportato lo schema dei microdati relativo alla fonte dati *Contenzioso TAR*¹⁷. In ogni caso, l'area in questione contiene tutti i microdati validati di ogni fonte dati facente parte del sistema e ciò, evidentemente, consente di ottenere in prospettiva qualsiasi tipo di indicatore definito su tali microdati.

Dell'area microdati fanno anche parte i cosiddetti *thesauri*, ossia un insieme di tabelle contenenti le modalità specifiche di ciascuna variabile qualitativa legata alle fonti dati.

Come già accennato, questa porzione della base di dati viene popolata, per ciascuna fonte dati, a partire dal file dei microdati validati tramite l'applicazione *SITG-Load*.

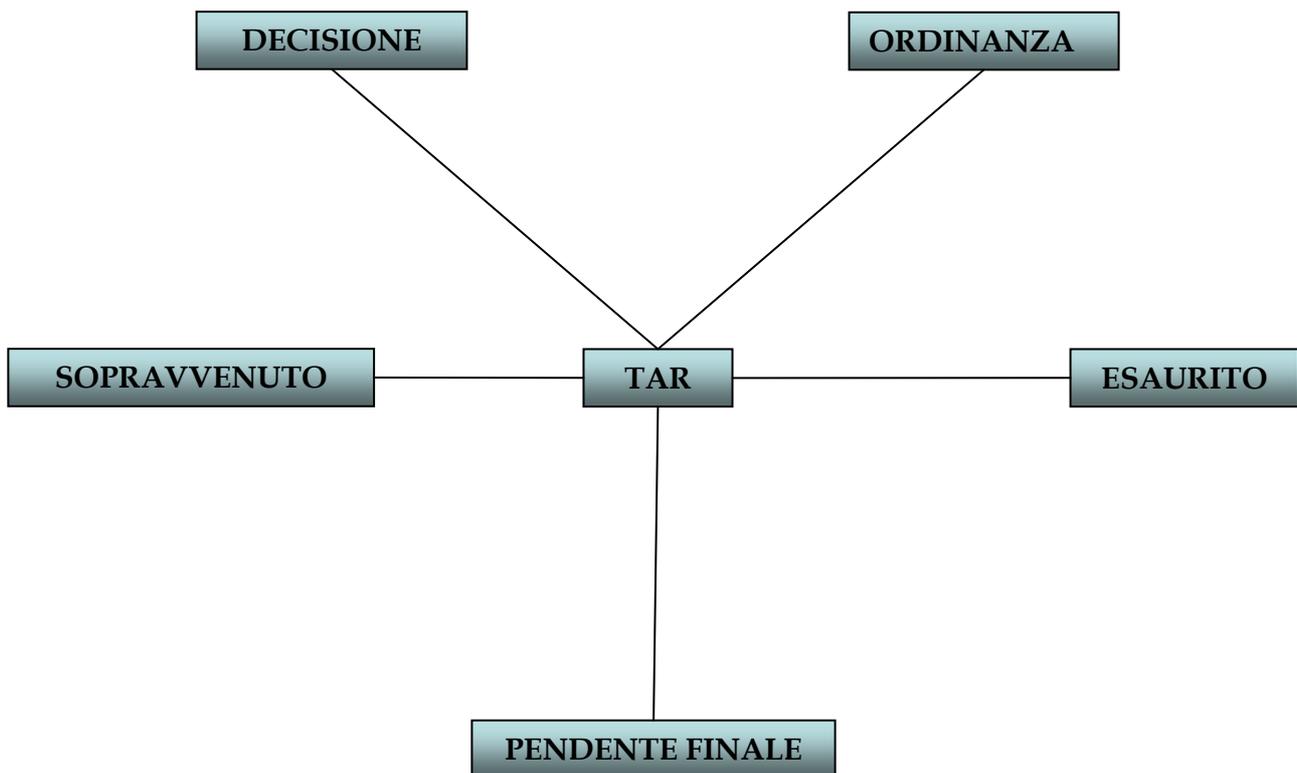


Fig.3 – Contenzioso TAR – schema dei microdati.

5.2.2. **Macrodati**

Anche l'area dei macrodati, come quella dei microdati, è costituita da un insieme di sotto-schemi, uno per ciascuna fonte dati, aventi struttura *a stella* o, nel caso

¹⁷ Per ulteriori informazioni o esempi si consulti la documentazione di progetto del "Sistema Informativo Territoriale sulla Giustizia".

più generale, *a fiocco di neve*. Tali schemi presentano alcune caratteristiche ricorrenti che si riportano di seguito:

- ✓ La tabella dei fatti si riferisce all'oggetto di analisi della fonte dati (nel caso di schemi a fiocco di neve esistono più tabelle dei fatti, ciascuna delle quali associata ad uno specifico oggetto di analisi della fonte dati corrispondente) e contiene i valori delle variabili quantitative rilevate e le chiavi esterne verso le opportune tabelle delle dimensioni.
- ✓ Ciascuna tabella delle dimensioni si riferisce ad una variabile di tipo qualitativo e contiene tutte le modalità che tale variabile può assumere. Particolari tabelle che appartengono a questa categoria sono quelle relative alla dimensione temporale ed a quella territoriale (secondo le due gerarchie *amministrativa e giudiziaria*).

Tale struttura consente di calcolare facilmente i valori associati ad un indicatore, tramite una semplice query di selezione sulla tabella dei fatti (per indicatori basati su variabili quantitative) o di join fra la tabella dei fatti ed una o più tabelle delle dimensioni (per indicatori basati su variabili quantitative e qualitative)¹⁸.

A titolo di esempio, in fig.4 è riportato lo schema dei macrodati relativo alla fonte dati *Contenzioso TAR*¹⁹.

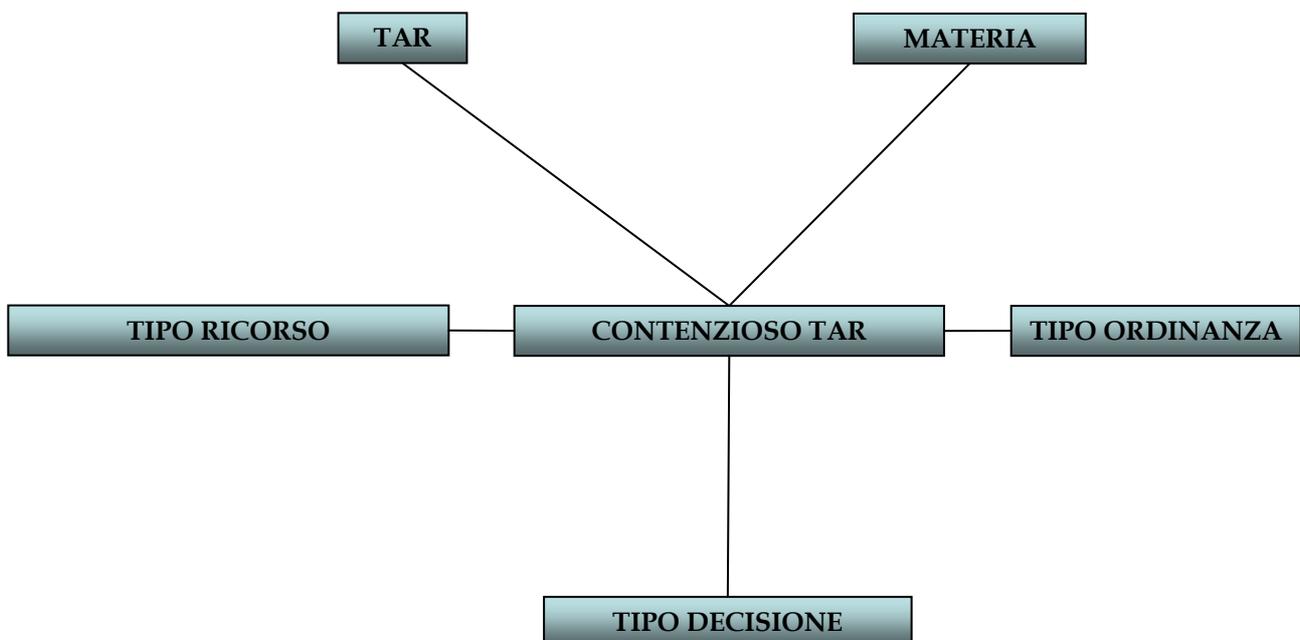


Fig.4 – Contenzioso TAR – schema dei macrodati.

Analogamente a quanto si verifica per i thesauri della componente microdati, è possibile individuare diversi livelli di condivisione delle dimensioni:

- ✓ Dimensioni comuni a tutte le fonti dati del sistema (es. *dimensione territoriale e dimensione temporale*).
- ✓ Dimensioni comuni a tutte le fonti dati di un'area (es. *sintetica del delitto per l'area penale e tipo di affidamento per l'area civile*).

¹⁸ In entrambi i casi, ovviamente, è prevista un'operazione di join con le due tabelle dimensionali relative, rispettivamente, al tempo ed al territorio.

¹⁹ Per ulteriori informazioni o esempi si consulti la documentazione di progetto del "Sistema Informativo Territoriale sulla Giustizia".

- ✓ Dimensioni proprie di una specifica fonte dati (es. *testi legislativi* per la fonte dati *criminalità*, *frequenza delle visite* per la fonte dati *separazioni* e *tipo di ordinanza* per la fonte dati *TAR*).

Tutte le dimensioni presentano le eventuali dimensioni gerarchiche in forma denormalizzata come nel caso delle dimensioni territoriali e delle classificazioni a più livelli (es. *condizione professionale* nelle fonti dati *separazioni* e *divorzi*).

Le dimensioni le cui modalità corrispondono a intervalli di valori (ad esempio *classi di età*) hanno una struttura a livelli in cui la classificazione relativa agli intervalli più stretti è quella referenziata dalle tabelle dei fatti, in modo tale da consentire l'aggiunta di ulteriori intervalli anche in un secondo momento.

Questa porzione della base di dati viene popolata, per ciascuna fonte dati, a partire dai corrispondenti microdati e tramite l'esecuzione di una serie di stored procedures PL/SQL attivabili da *SITG-Admin*.

5.2.3. Knowledge base

Si tratta della porzione della base di dati su cui, in sostanza, si fonda il sistema di gestione e manutenzione. La knowledge base contiene tutte le informazioni persistenti relative agli oggetti definiti all'interno del sistema: utenze, categorie, fonti dati, indicatori e tavole statistiche. Per ciascuno di tali oggetti, nella knowledge base vengono memorizzate tutte le informazioni che ne consentono la gestione durante tutto l'arco del suo ciclo di vita; tali informazioni, ovviamente, si differenziano in funzione del tipo di oggetto considerato:

- ✓ *utenza*: nome, password, ruoli attribuiti, privilegi di accesso alle categorie disponibili;
- ✓ *categoria*: nome, descrizione, referente statistico, posizione all'interno dell'albero delle categorie;
- ✓ *fonte dati*: nome, descrizione, associazioni fra tabelle della base di dati e variabili statistiche, lista delle procedure di produzione dei macrodati a partire dai corrispondenti microdati validati;
- ✓ *indicatore*: nome, descrizione, tipo (assoluto o derivato), categoria di appartenenza, gerarchie e livelli di dettaglio territoriali disponibili (con relativi periodi temporali di validità), regole per il calcolo dell'indicatore, formato numerico;
- ✓ *tavola statistica*: titolo, versione, lista delle etichette di testata, lista degli indicatori coinvolti.

Questa porzione della base di dati viene popolata interattivamente e automaticamente dagli utenti statistici e tramite l'utilizzo di *SITG-Admin*.

5.2.4. Tavole e documentazione

Questa componente deriva direttamente dalla base di dati del sistema *DaWinci-PD* e contiene tutte le informazioni ed i dati relativi alle tavole statistiche presenti nella sezione "*Dati dal 2000*" (generate e/o pubblicate) ed alla relativa documentazione associata. Per quanto attiene alle tavole statistiche, sono gestite due categorie di informazioni:

- ✓ informazioni relative alla definizione della tavola: si tratta, sostanzialmente, di informazioni come il titolo della tavola, i nomi delle colonne in testata e delle righe in fiancata e le note;
- ✓ informazioni relative ai dati contenuti nella tavola: si tratta dei veri e propri valori che fanno parte della tavola in esame.

Per quanto riguarda la documentazione a corredo delle tavole statistiche presenti nel sistema, in questa componente sono contenute tutte le informazioni che definiscono ciascun oggetto (scheda informativa, pubblicazione, norma legislativa, etc.) presente nel sistema e, nel caso in cui esistano files esterni alla base di dati associati a specifici documenti, un link ai files stessi.

Questa porzione della base di dati viene popolata indirettamente dagli utenti statistici, tramite l'utilizzo di *SITG-Admin* (per le tavole generate e/o pubblicate) e *SITG-Doc* (per la documentazione).

Come accennato all'inizio del presente paragrafo, la componente dati di SITG è costituita da due basi di dati. Con riferimento alle quattro aree appena descritte, è possibile precisare che la base di dati interna le contiene tutte, mentre quella esposta è costituita esclusivamente da una replica parziale della sola partizione *tavole e documentazione*, ovvero dalle sole informazioni relative alle tavole statistiche ed ai documenti pubblicati.

5.3. Architettura di dispiegamento

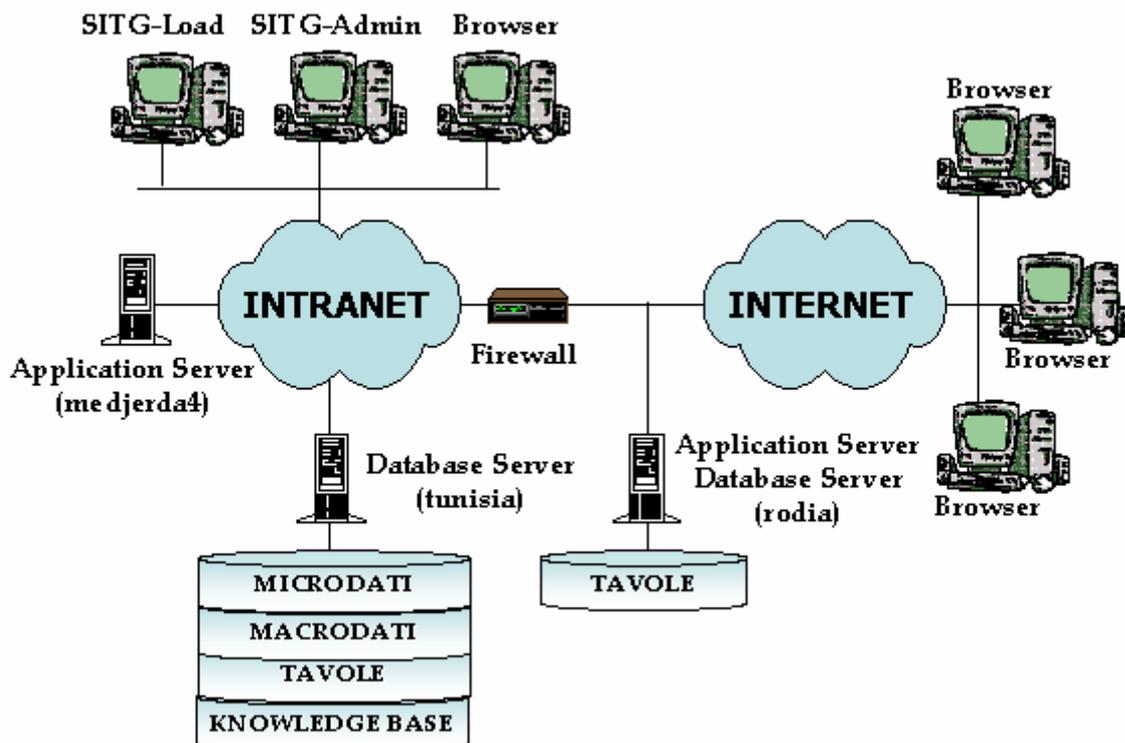


Fig.5 – SITG: architettura di dispiegamento.

In fig.5 è riportata l'architettura di dispiegamento di SITG. Come si può notare, esistono fisicamente due distinti sotto-sistemi di diffusione *SITG-Web*: uno accessibile esclusivamente nell'ambito della intranet di Istituto (utilizzato tramite browser dagli statistici per la verifica e validazione di tutte le tavole prodotte) ed uno disponibile su

SISTEMA INFORMATIVO TERRITORIALE SULLA GIUSTIZIA:
strumenti di gestione e manutenzione

Internet (contenente le sole tavole pubblicate, con la relativa documentazione a supporto).

Il sotto-sistema di gestione e manutenzione, al contrario, risulta fisicamente costituito dal software e dai dati presenti sul database server e sull'application server interni e dai fat clients *SITG-Admin* e *SITG-Load* e risulta, proprio per questo, accessibile esclusivamente in ambito intranet, da parte dei soli utenti autorizzati.

Come già evidenziato, tutta la componente dati del sistema risiede sul database server interno, mentre su quello esterno è replicata la sola porzione relativa alle tavole statistiche pubblicate ed alla relativa documentazione.

6. Il ciclo di vita delle fonti dati, degli indicatori e delle tavole

Nell'ambito del processo di produzione e diffusione del SITG descritto nel precedente paragrafo §.2, tutti gli oggetti definiti e gestiti all'interno del sistema attraversano varie fasi, diverse in funzione del particolare tipo di oggetto considerato, che ne caratterizzano lo specifico ciclo di vita all'interno del sistema. Allo scopo di illustrare alcuni tipici scenari reali di utilizzo del sistema di gestione e manutenzione nel successivo paragrafo §.7, si descrive di seguito il ciclo di vita relativo alle tre categorie di oggetti fondamentali in ambito SITG: le fonti dati, gli indicatori e le tavole statistiche.

6.1. Ciclo di vita delle fonti dati

In fig.6 è riportato il diagramma degli stati relativi al ciclo di vita di una fonte dati:

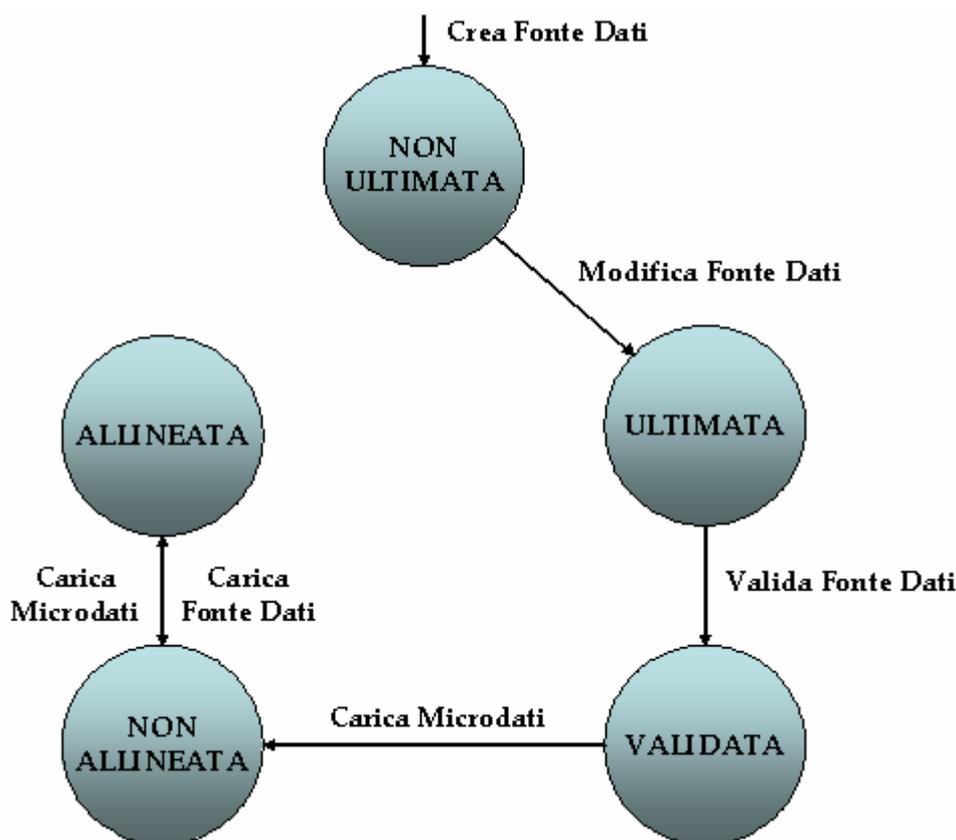


Fig.6 – Ciclo di vita delle fonti dati.

Nel momento in cui un utente con profilo di amministratore crea una nuova fonte dati, questa entra nello stato iniziale di **NON ULTIMATA**, che sta ad indicare che la fonte dati esiste ma che non tutte le variabili disponibili sono state definite. A questo punto, inizia una fase di successive modifiche della fonte dati attraverso le quali viene specificata da un lato la sequenza di procedure software che il *SITG-Admin* dovrà attivare per la produzione dei macrodati a partire dai microdati validati e dall'altro l'associazione fra i campi delle tabelle della base di dati coinvolte e le variabili che saranno rese disponibili agli utenti statistici.

SITG-Admin rileva automaticamente la situazione in cui tutte le variabili sono state definite e porta la fonte dati nello stato **ULTIMATA**. A questo punto, l'utente

amministratore deve verificare che la fonte dati in esame sia stata correttamente definita in ogni dettaglio e, conseguentemente, portarla nello stato VALIDATA.

Nel momento in cui i microdati relativi alla fonte dati in esame vengono caricati nel sistema (tramite l'applicativo SITG-Load) per uno o più anni di riferimento, la fonte dati stessa risulterà non allineata per gli anni in questione. A questo punto, è possibile attivare la produzione dei macrodati per tutti gli anni per i quali risultano disponibili i relativi microdati validati. Al termine di tale processo, la fonte dati risulterà, rispetto agli anni per i quali esso è stato compiuto, nello stato ALLINEATA, stato finale del suo ciclo di vita²⁰.

6.2. Ciclo di vita degli indicatori

In fig.7 è riportato il diagramma degli stati relativi al ciclo di vita di un indicatore:

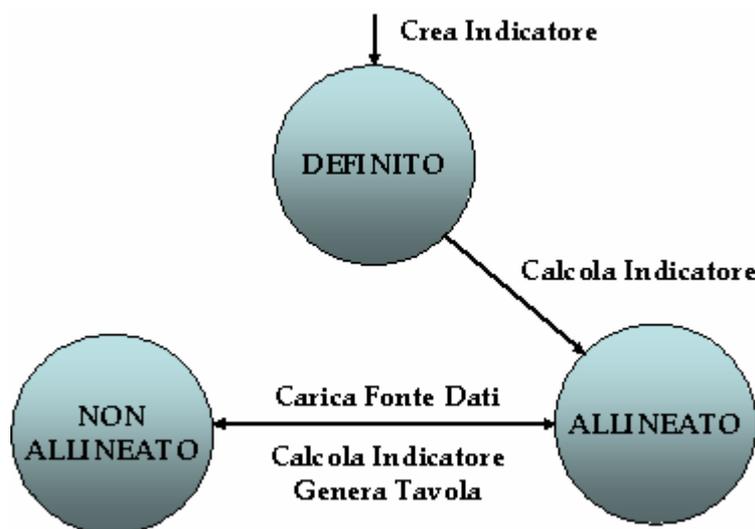


Fig.7 – Ciclo di vita degli indicatori.

Nel momento in cui un utente con profilo di statistico crea un nuovo indicatore, questo entra nello stato iniziale di DEFINITO; ciò significa che sono presenti nel sistema le specifiche di calcolo dell'indicatore, ma non i dati ad esso relativi.

A questo punto, l'utente può richiedere che sia effettuato il calcolo dell'indicatore appena definito, semplicemente specificando l'anno di riferimento da usare per il calcolo; l'indicatore entra così nello stato ALLINEATO.

Infine, nel caso in cui per almeno una delle fonti dati rispetto alle quali l'indicatore è stato definito sia di nuovo attivata la procedura di produzione dei macrodati, l'indicatore passerà nello stato di NON ALLINEATO, dal quale potrà tornare allo stato ALLINEATO sia in modo esplicito (nel caso in cui l'utente effettui di nuovo il calcolo dell'indicatore), sia implicitamente (come effetto della generazione di una tavola che lo contiene).

6.3. Ciclo di vita delle tavole

In fig.8 è riportato il diagramma degli stati relativi al ciclo di vita di una tavola:

²⁰ A meno che, per qualsiasi motivo, non venga di nuovo effettuato il caricamento dei microdati, che riporterebbe la fonte dati nello stato NON ALLINEATA per l'anno di riferimento interessato.

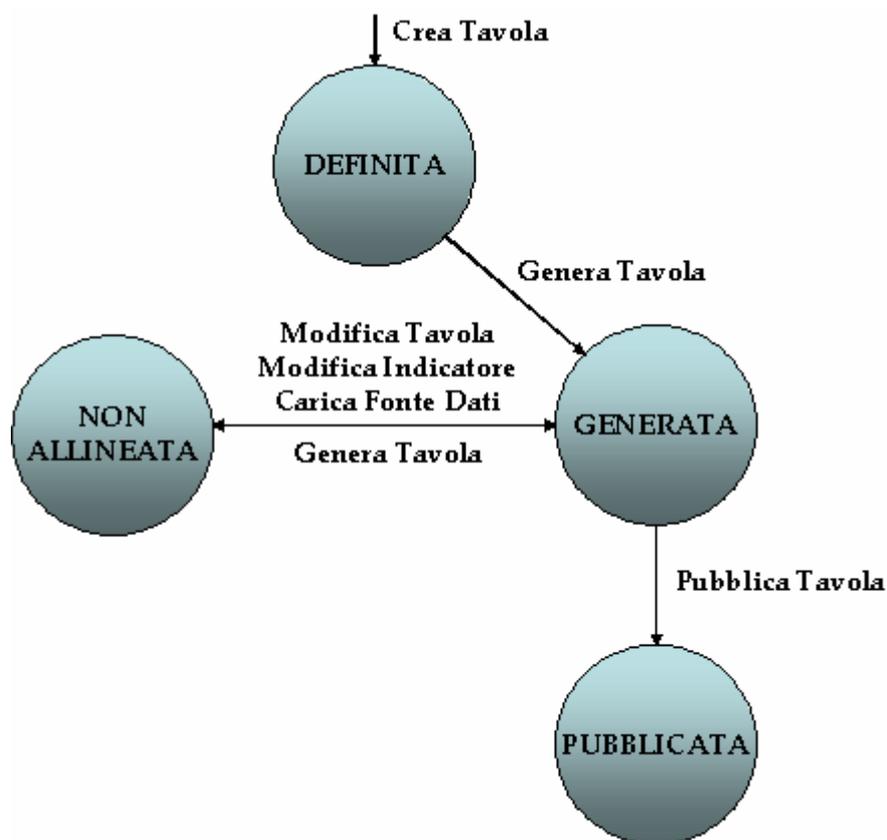


Fig.8 – Ciclo di vita delle tavole.

Nel momento in cui un utente con profilo di statistico crea una nuova tavola, questa entra nello stato iniziale di DEFINITA; ciò significa che sono presenti nel sistema tutti gli elementi che definiscono la tavola (titolo, etichette di testata, indicatori coinvolti, note, etc.), ma non i dati ad essa relativi.

A questo punto l'utente può richiedere che sia effettuata la generazione della tavola appena definita, semplicemente specificando l'anno di riferimento da usare per il calcolo; la tavola entra così nello stato di GENERATA per l'anno di riferimento selezionato.

A questo punto, possono intervenire due ulteriori transizioni di stato. La prima deriva da modifiche nella definizione della tavola o degli indicatori in essa contenuti; la seconda si verifica nel caso in cui per almeno una delle fonti dati rispetto alle quali la tavola è stata definita, sia di nuovo attivata la procedura di produzione dei macrodati. In entrambi i casi, la tavola passerà nello stato di NON ALLINEATA e potrà tornare allo stato di GENERATA semplicemente richiedendo la generazione della tavola in esame.

Infine, dopo averne verificata la correttezza, l'utente statistico può rendere disponibile la tavola in esame alla comunità Internet, portando così la tavola stessa nello stato di PUBBLICATA per uno specifico anno di riferimento.

È importante notare, infine, come ad ogni tavola siano in generale associate più versioni, ciascuna delle quali contraddistinta da una propria definizione. Ciò consente di gestire in modo semplice situazioni nelle quali una tavola subisca variazioni nel corso del tempo²¹ ed implica che tutte le transizioni di stato precedentemente illustrate si riferiscono, in effetti, ad una specifica versione della tavola in questione.

²¹ Si pensi a variazioni nel titolo, nel numero o tipo di indicatori contenuti nella tavola, nel livello di dettaglio territoriale disponibile, etc.

7. Scenari di utilizzo del sistema di gestione e manutenzione

Nel presente paragrafo saranno descritti i due scenari principali di utilizzo del sistema di gestione e manutenzione, ovvero il processo di inserimento di una nuova indagine (includere le tavole e la documentazione ad esse associata) ed il processo di manutenzione di un'indagine già presente nel sistema.

7.1. Inserimento di una nuova indagine

7.1.1. Attività preliminari

Nell'ipotesi in cui sia voglia inserire una nuova indagine nel sistema, sarà preventivamente necessario espletare una serie di attività di carattere informatico al fine di predisporre le opportune strutture dati e procedure software per poter gestire la nuova fonte dati. Scendendo maggiormente nel dettaglio, le attività da effettuare sono le seguenti:

1. creazione delle strutture dati relative ai microdati della nuova indagine;
2. creazione delle strutture dati relative ai macrodati della nuova indagine;
3. sviluppo delle procedure software per il caricamento dei microdati relativi alla nuova indagine;
4. sviluppo delle procedure software per la produzione dei macrodati relativi alla nuova indagine.

A questo punto, è possibile procedere con l'inserimento della nuova indagine nel sistema.

7.1.2. Caricamento dei microdati

Il primo passo da compiere è quello dell'attivazione, tramite l'applicativo *SITG-Load*, delle procedure software per il caricamento dei microdati. Al termine dell'esecuzione di tali procedure, i microdati relativi alla nuova indagine saranno memorizzati all'interno delle relative strutture dati.

7.1.3. Creazione della fonte dati

Successivamente è necessario creare, tramite l'applicativo *SITG-Admin*, una nuova fonte dati, che sarà associata alla nuova indagine (cfr. fig.9). A tale fonte dati saranno legate tutte le informazioni necessarie per la sua corretta gestione, ovvero:

1. l'insieme ordinato delle procedure software che dovranno essere attivate per la produzione dei macrodati;
2. il mapping fra gli elementi delle strutture dati coinvolte e le variabili (quantitative e qualitative) che saranno rese disponibili agli utenti statistici.

Al completamento della definizione, la fonte dati sarà *validata* e resa così disponibile per la lavorazione. A questo punto, l'utente statistico può procedere alla definizione degli indicatori sulle variabili disponibili per la nuova fonte dati.

Nuova Fonte Dati

Area: Data Mart Dimensioni

Nome Fonte Dati:

Desc. Fonte Dati:

Nome Pacchetto: Nome Procedura:

Nome Flusso:

Procedure	Nome Flusso	Ordine
API_DWG_LOAD_DIM.DWG_LOAD_D3_B25		2
API_DWG_LOAD_DIM.DWG_LOAD_D4_B2520		3
API_DWG_LOAD_FACT.DWG_F_B_SEPARAZIONI	M_GIUBM2520_FLUSSI	4
API_DWG_LOAD_FACT.DWG_F_B_SEP_FIGLIMINORI	M_GIUBM2520_FLUSSI	5

Nome File:

Nome Tabella	Tipo	Gerarchia	Livello	In Descrizione
DWG_F1_B_SEPARAZIONI	Fatti Principali			
DWG_MAC_F2_SEP_FIGLIMINORI	Fatti Secondari			
DWG_D1_PROVINCIA	Dimensioni	Amministrativa	Province	PROVINCIA
DWG_D3_B_CIRCONDARIO	Dimensioni	Giudiziaria	Circondari	CIRCONDARIO
DWG_D1_CITTADINANZA_MOGLIE	Dimensioni			CITTADINANZA

Record: 1/? <DSC> <DBG>

Fig.9 – SITG-Admin – creazione di una nuova fonte dati.

7.1.4. Definizione e calcolo degli indicatori

In fig.10 è riportata la form principale di accesso da parte dell'utente statistico al *SITG-Admin*. Come si vede, gli oggetti gestibili (categorie, indicatori e tavole statistiche) sono organizzati secondo una struttura ad albero definita dallo stesso utente. Più precisamente, l'albero delle categorie (aree, materie e sotto-materie) riflette esattamente la struttura secondo la quale sono pubblicati i dati e la relativa documentazione nel sotto-sistema di diffusione *SITG-Web*; quello delle tavole, per ovvi motivi, presenta un'identica organizzazione, mentre quella degli indicatori ha una propria struttura che ha il solo scopo di permettere all'utente statistico di organizzare secondo i criteri ritenuti più opportuni gli indicatori disponibili. Allo scopo di definire l'insieme degli indicatori associati alla nuova fonte dati, l'utente statistico non deve far altro che selezionare la cartella degli indicatori nella quale intende inserire i nuovi elementi (o, in alternativa, creare una nuova cartella) e selezionare la funzionalità di creazione di un nuovo indicatore. Viene così attivata la form di creazione di un nuovo indicatore (cfr. figg.11.a e 11.b) che consente di inserire tutte le necessarie informazioni in modo molto semplice ed intuitivo e più precisamente:



Fig.10 – SITG-Admin – form di accesso al sistema.

1. *nome* (sia in italiano, sia in inglese);
2. *codice identificativo*;
3. *formato decimale*;
4. *fonte dati* di provenienza (richiesta esclusivamente per gli indicatori di tipo assoluto);
5. *tipo* (assoluto o derivato);
6. *tipo di formula* (indica se le specifiche di calcolo saranno definite indirettamente, combinando opportunamente le variabili o gli indicatori disponibili e le eventuali condizioni di selezione, oppure indicando esplicitamente il codice SQL da eseguire);
7. *flag di diffusione* (flag booleano che permette di inserire o meno l'elemento in questione nell'elenco degli indicatori disponibile nell'ambito del sotto-sistema di diffusione);
8. *flag di indicatore costante* (flag booleano che permette di specificare se l'elemento in questione assume valore costante su tutte le unità territoriali o meno);
9. *specifiche di calcolo* (fornite in modo congruente rispetto a quanto indicato al precedente punto 6);
10. *territorio* (permette di indicare, per ciascun anno di riferimento, tutte le gerarchie ed i livelli di dettaglio territoriali disponibili).

SISTEMA INFORMATIVO TERRITORIALE SULLA GIUSTIZIA:
strumenti di gestione e manutenzione

The screenshot shows the 'Indicatori' window in the SITG-Admin application. The main configuration area includes:

- Nome:** Numero minori denunciati MF
- Nome Eng:** (empty)
- Sigla:** MI0001
- Formato decimale:** 12345
- Fonte Dati:** MINORENNI DENUNCIATI
- Tipo Indicatore:** Indicatore assoluto, Indicatore Derivato
- Tipo Formula:** Formula, Sql
- Indicatori Diffusione:** No, Si
- Costante:** Costante, Non Costante
- Specifiche di calcolo:** Numero di MINORI DENUNCIATI dove TIPO ARCHIVIAZIONE != Perché autore ignoto
- Specifica / Gerarchia:**
 - Specifica:** ANNO (highlighted), DATA DELITTO, DATA DENUNCIA, DATA ISCRIZIONE, DATA RICHIESTA, DATA RICHIESTA PRECEDENTE
 - Gerarchia:** ABITAZIONE, CAPO DELITTO, CITTADINANZA, CONTINENTE, DELITTO, DELITTO - SINTETICA
- Condizione di SELECT:** TIPO ARCHIVIAZIONE != Perché autore ignoto

At the bottom, the status bar shows 'Record: 1/1' and '<OSC> <DBG>'.

Fig. 11.a – SITG-Admin – creazione di un nuovo indicatore.

This screenshot shows the same 'Indicatori' window, but with the 'Gerarchia' tab selected. The configuration is as follows:

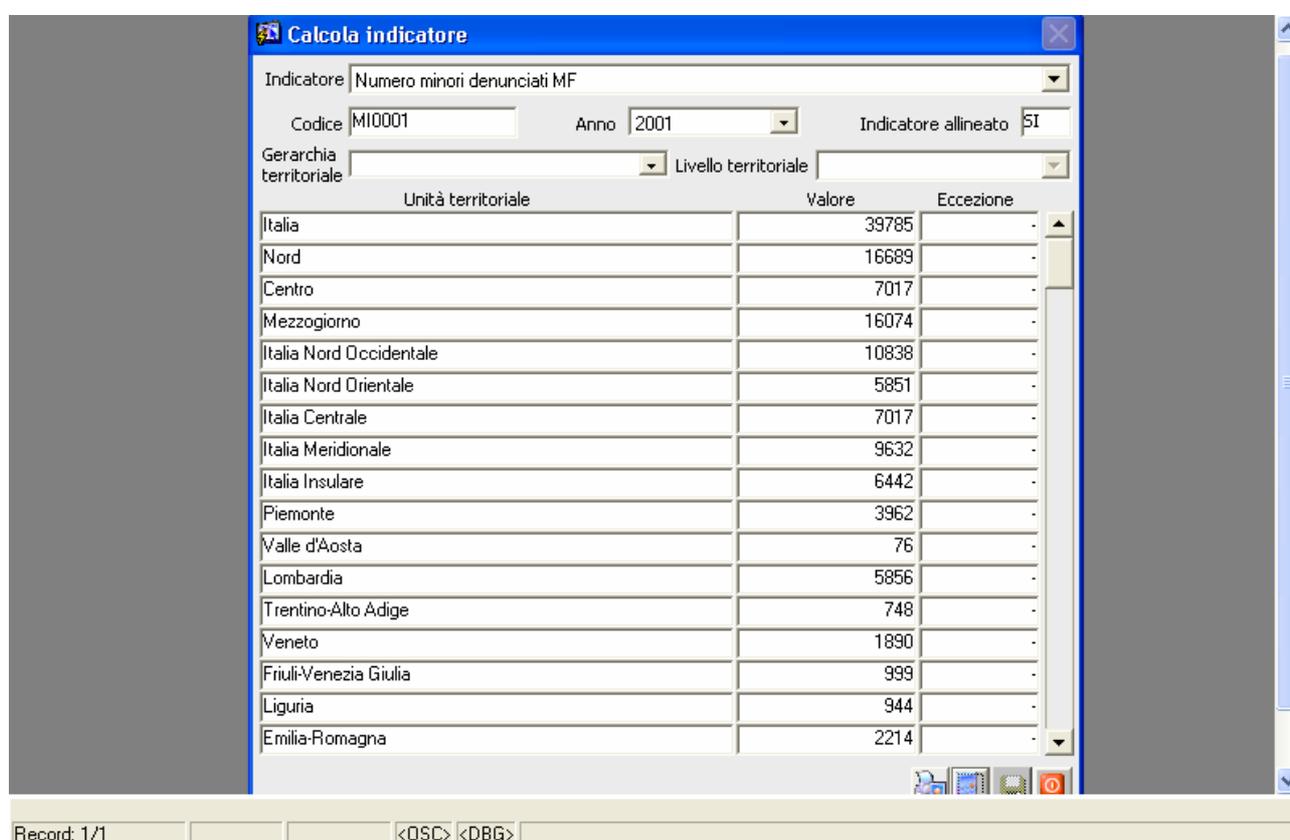
- Nome:** Numero minori denunciati MF
- Nome Eng:** (empty)
- Sigla:** MI0001
- Formato decimale:** 12345
- Fonte Dati:** MINORENNI DENUNCIATI
- Tipo Indicatore:** Indicatore assoluto, Indicatore Derivato
- Tipo Formula:** Formula, Sql
- Indicatori Diffusione:** No, Si
- Costante:** Costante, Non Costante
- Specifiche di calcolo:** Numero di MINORI DENUNCIATI dove TIPO ARCHIVIAZIONE != Perché autore ignoto
- Specifica / Gerarchia:**
 - Filtri:** Gerarchia: Gerarchia Territoriale Amministrativa, Livello territoriale: Italia
 - Gerarchia/Livello Indicatore:**

Gerarchia	Livello territoriale	Anno Inizio	Anno Fine
<input type="checkbox"/> Gerarchia Territoriale Amministrativa	Italia	2000	
<input type="checkbox"/> Gerarchia Territoriale Amministrativa	Dati per Ripartizione a 3 livelli	2000	
<input type="checkbox"/> Gerarchia Territoriale Amministrativa	Dati per Ripartizione a 5 livelli	2000	
<input type="checkbox"/> Gerarchia Territoriale Amministrativa	Dati per Regione	2000	
<input type="checkbox"/> Gerarchia Territoriale Amministrativa	Dati per Provincia	2000	
<input type="checkbox"/> Gerarchia Territoriale Giudiziaria	Tutti i Distretti	2000	
<input type="checkbox"/> Gerarchia Territoriale Giudiziaria	Dati per Distretto di Corte di Appello	2000	

The status bar at the bottom shows 'Record: 1/1' and '<OSC> <DBG>'.

Fig. 11.b – SITG-Admin – creazione di un nuovo indicatore.

Una volta salvata la definizione del nuovo indicatore, l'utente ha la possibilità, se lo desidera, di effettuare il calcolo dell'indicatore stesso, per tutte le gerarchie ed i livelli di dettaglio territoriali disponibili per ciascun anno di riferimento (cfr. fig.12). In tal modo, lo statistico può verificare, a livello di singolo indicatore, se i valori calcolati risultano corretti e, in caso contrario, verificare già in questa fase se sono stati commessi errori nelle precedenti attività di caricamento dei microdati, di produzione dei macrodati e di definizione della regola di calcolo dell'indicatore. Per ragioni di efficienza, tutti i calcoli effettuati possono essere salvati, in modo tale da non dover essere ripetuti successivamente (ad esempio durante la fase di generazione della tavola statistica che conterrà l'indicatore in questione). A questo punto, l'utente può passare a definire (ed eventualmente calcolare) tutti gli altri indicatori di interesse per la nuova fonte dati. A tale proposito, il sistema offre anche una funzione di tipo "create like" che consente di definire un nuovo indicatore a partire da un indicatore precedentemente inserito nel sistema, in modo tale da ridurre al minimo l'attività di data entry da parte dell'utente.



Unità territoriale	Valore	Eccezione
Italia	39785	-
Nord	16689	-
Centro	7017	-
Mezzogiorno	16074	-
Italia Nord Occidentale	10838	-
Italia Nord Orientale	5851	-
Italia Centrale	7017	-
Italia Meridionale	9632	-
Italia Insulare	6442	-
Piemonte	3962	-
Valle d'Aosta	76	-
Lombardia	5856	-
Trentino-Alto Adige	748	-
Veneto	1890	-
Friuli-Venezia Giulia	999	-
Liguria	944	-
Emilia-Romagna	2214	-

Fig.12 – SITG-Admin – calcolo di un indicatore.

7.1.5. Definizione, generazione e pubblicazione delle tavole

Dopo aver inserito nel sistema tutti gli indicatori necessari, è possibile definire la tavola che li contenga. Selezionando la categoria di appartenenza²² ed attivando la form di creazione di una nuova tavola (cfr. figg.13.a e 13.b) è possibile inserire tutte le informazioni necessarie, ovvero:

²² Il sistema consente, eventualmente, di definire una nuova categoria (area, materia o sotto-materia) che contenga le tavole relative alla nuova indagine (cfr. §.4.3).

SISTEMA INFORMATIVO TERRITORIALE SULLA GIUSTIZIA:
strumenti di gestione e manutenzione

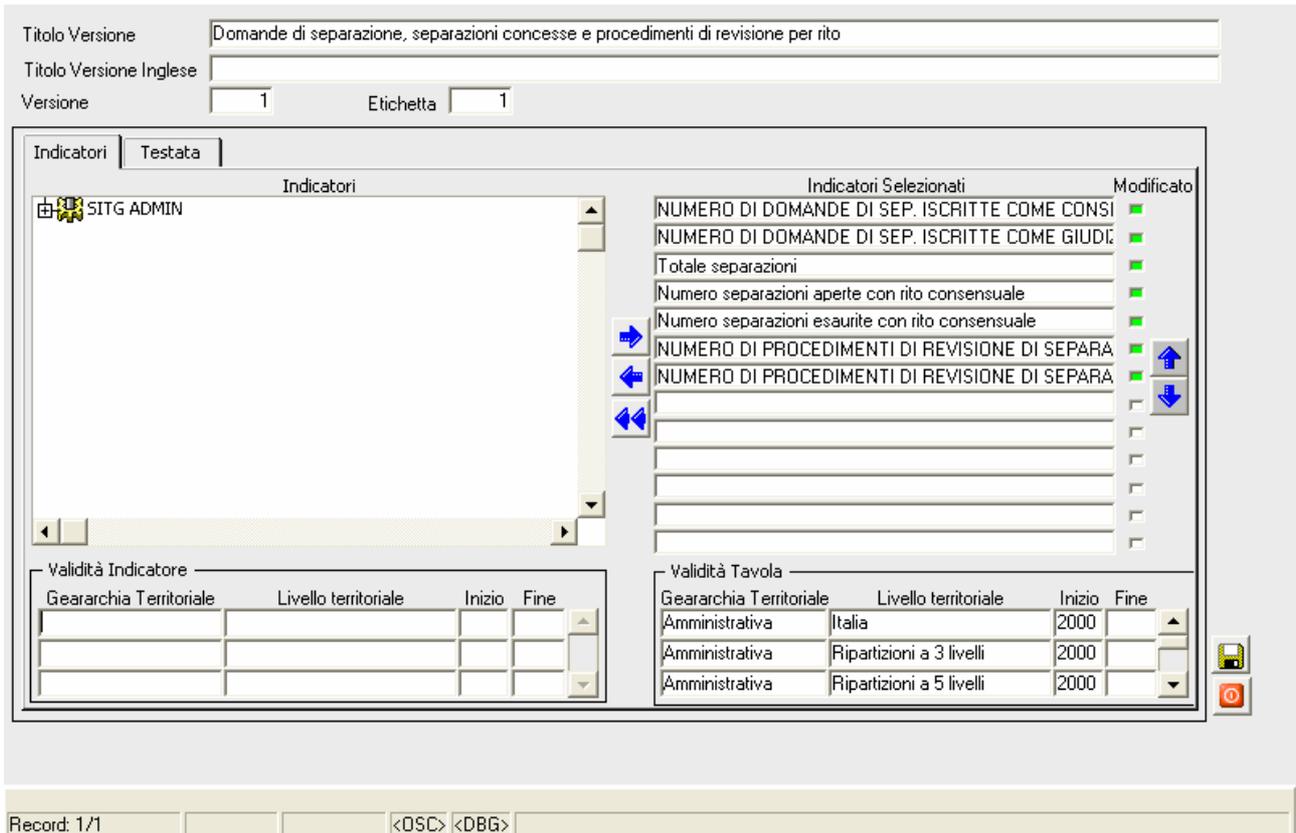


Fig. 13.a – SITG-Admin – creazione di una nuova tavola / scelta indicatori.

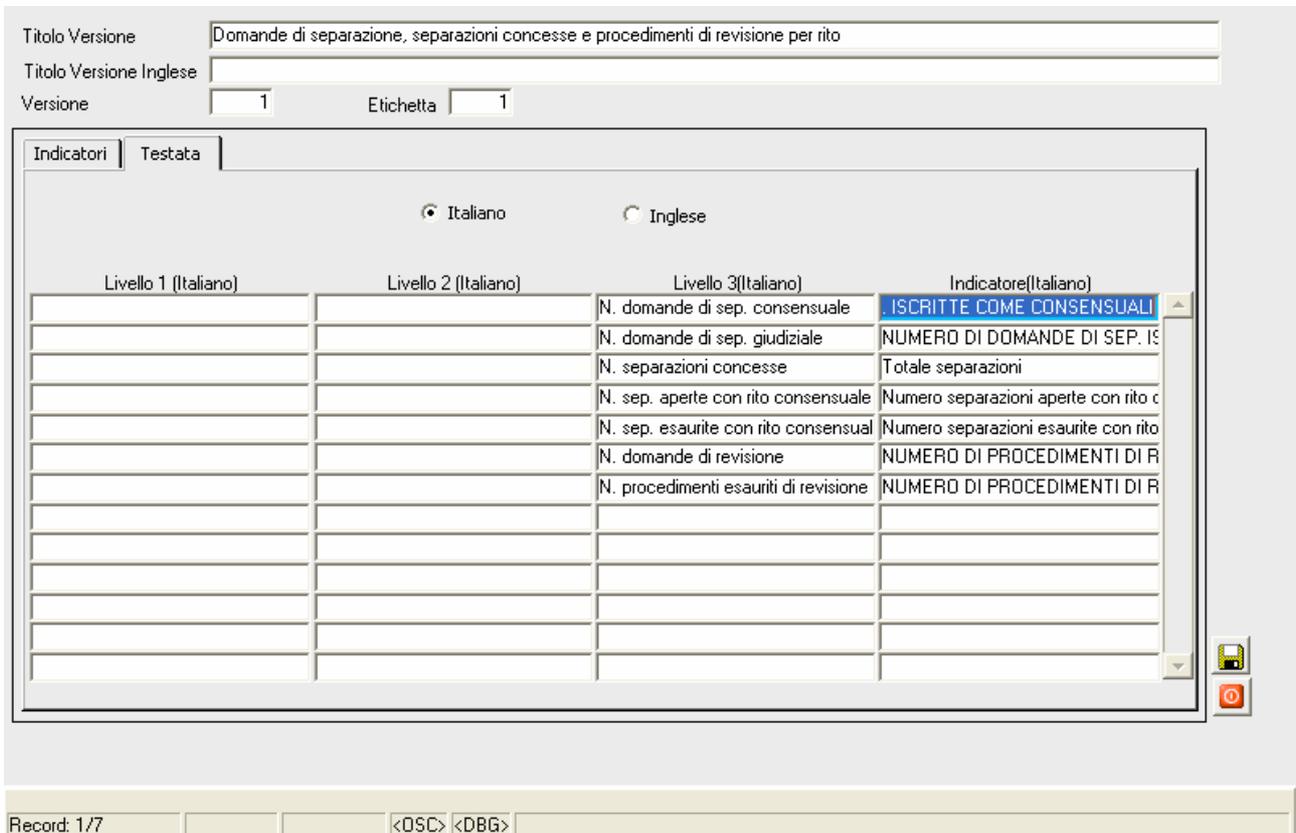


Fig. 13.b – SITG-Admin – creazione di una nuova tavola / definizione etichette di testata.

1. *titolo* (sia in italiano, sia in inglese);
2. *etichetta* (ovvero numero d'ordine nell'elenco delle tavole della categoria di appartenenza);
3. *elenco degli indicatori* (ovvero delle colonne che faranno parte della tavola);
4. *titoli di testata* (sia in italiano, sia in inglese, per ciascuna colonna di cui al precedente punto 3)²³
5. *note* (per ciascuna nota vanno indicati l'anno di riferimento e l'elemento della tavola cui la nota stessa si riferisce).

Si ritiene utile ricordare, a questo punto, come non risulti necessario inserire informazioni relative alla fiancata, che sarà generata automaticamente dal sistema in base alle informazioni relative alle gerarchie ed i livelli di dettaglio territoriali di disponibilità degli indicatori contenuti nella tavola.

Una volta salvata la definizione della nuova tavola, l'utente ha la possibilità di richiederne la generazione per uno specifico anno di riferimento (cfr. fig.14); in tal modo il sistema effettua il calcolo degli indicatori contenuti nella tavola stessa²⁴ e la inserisce nel sotto-sistema di diffusione *SITG-Web* in ambiente intranet, rendendola così immediatamente disponibile per la verifica e validazione da parte del ricercatore.

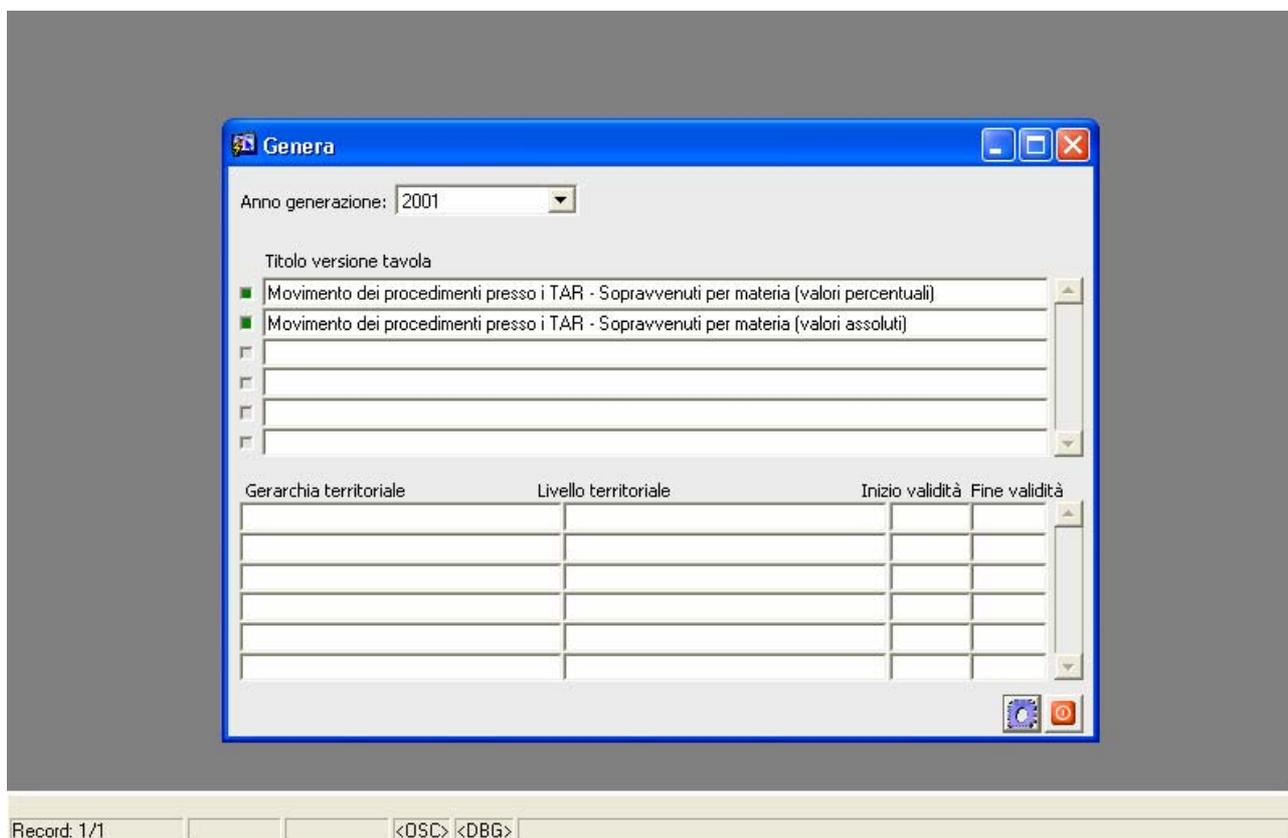


Fig.14 – SITG-Admin – generazione di una tavola.

Terminata la fase di verifica e validazione, l'utente può effettuare immediatamente la pubblicazione della tavola su Internet per l'anno di riferimento di interesse (cfr. fig.15): il sistema replica tutte le informazioni relative alla tavola in questione dalla

²³ In questa fase è possibile specificare eventuali testate multi-livello, per gestire raggruppamenti di colonne.

²⁴ Ad eccezione degli indicatori precedentemente calcolati, per i quali vengono riutilizzati i valori già memorizzati nella base di dati.

base di dati interna a quella esposta, rendendo così di fatto disponibile la nuova tavola alla comunità Internet.

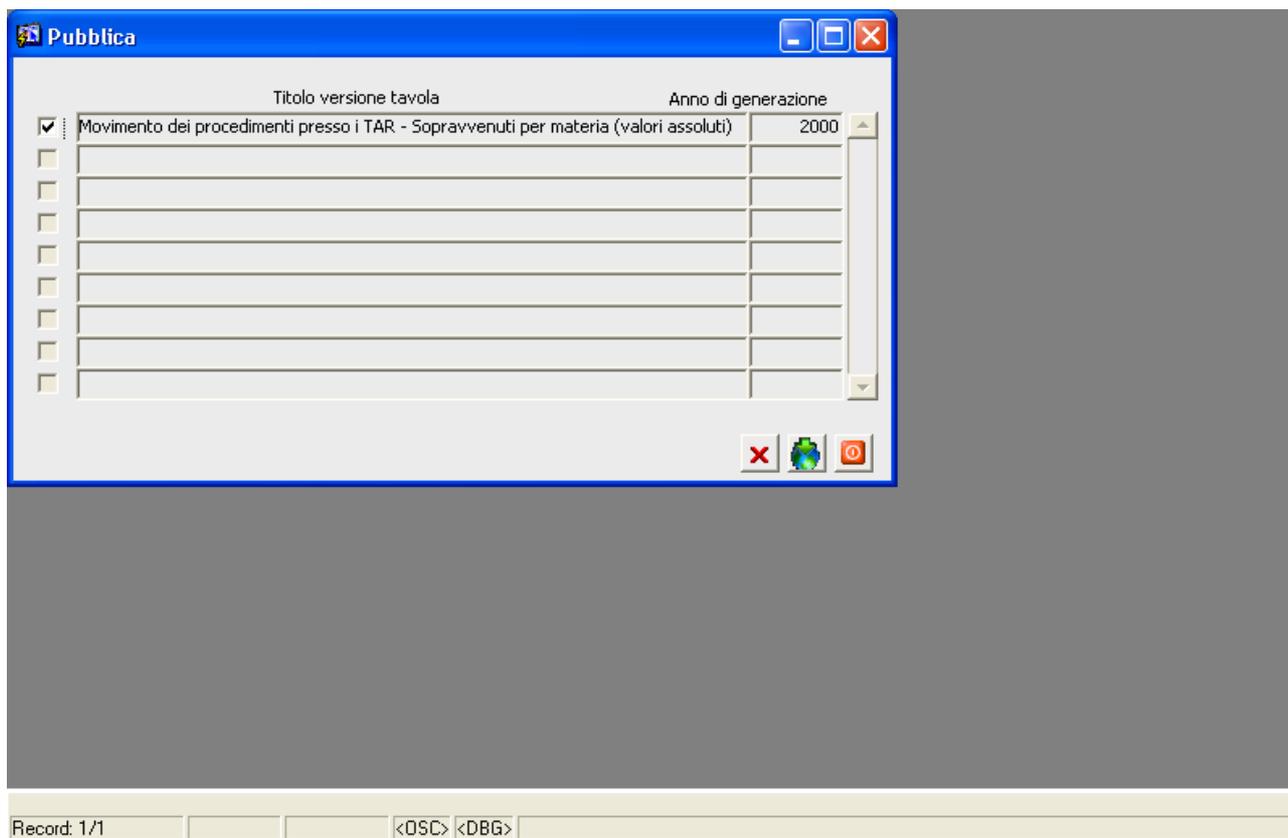


Fig.15 – SITG-Admin – pubblicazione di una tavola.

7.1.6. Pubblicazione della documentazione

Nel trattamento di una nuova indagine sarà certamente necessario inserire nel sistema non soltanto tavole statistiche, ma anche tutta una serie di informazioni a corredo. Sarà indispensabile, infatti, inserire almeno una nuova scheda informativa di indagine e, con ogni probabilità, tutta una serie di nuovi elementi come normative di riferimento dei fenomeni rilevati, pubblicazioni e voci di glossario.

Tali attività possono essere gestite in modo molto semplice tramite *SITG-Doc*, un'applicazione web-based del tutto analoga, dal punto di vista dell'interfaccia utente, al sotto-sistema di diffusione *SITG-Web* e che consente di inserire e pubblicare tutta la documentazione associata ad una nuova indagine. Scendendo maggiormente in dettaglio, le attività da compiere per la pubblicazione della documentazione di indagine sono di seguito riportate:

1. *Inserimento dei files di documentazione.* Alcune tipologie di documentazione (schede informative, normativa e pubblicazioni) prevedono, fra le varie proprietà associate al singolo oggetto, anche la specifica di links verso files che possono essere scaricati o aperti tramite plug-in del browser dall'utente. Ovviamente, affinché ciò possa essere fatto, i files in questione devono trovarsi in un'apposita directory definita sull'application server del sistema di diffusione. Per questo motivo, prima di inserire nel sistema un nuovo elemento informativo, è necessario trasferire tutti gli eventuali files cui tale elemento si

riferisce sull'application server. Ciò può essere fatto, molto semplicemente, tramite qualsiasi applicazione che faccia uso del protocollo ftp.

2. *Inserimento dei nuovi elementi informativi.* A questo punto è possibile inserire i nuovi oggetti (schede informative, pubblicazioni, voci di glossario, etc.) nel sistema in modo molto semplice ed intuitivo, selezionando, per ciascun oggetto, la categoria di appartenenza ed attivando la funzione di inserimento (cfr. fig.16). Il sistema richiederà, secondo la tipologia di oggetto, tutte le informazioni necessarie e, al termine della procedura di inserimento, renderà immediatamente disponibile l'oggetto appena definito in ambiente intranet. A questo punto, lo statistico può verificare tramite SITG-Web che l'oggetto sia stato correttamente definito e, successivamente, effettuare la pubblicazione su Internet: *SITG-Doc* replicherà tutte le informazioni relative all'oggetto in questione dalla base di dati interna a quella esposta, rendendolo così di fatto visibile alla comunità Internet.

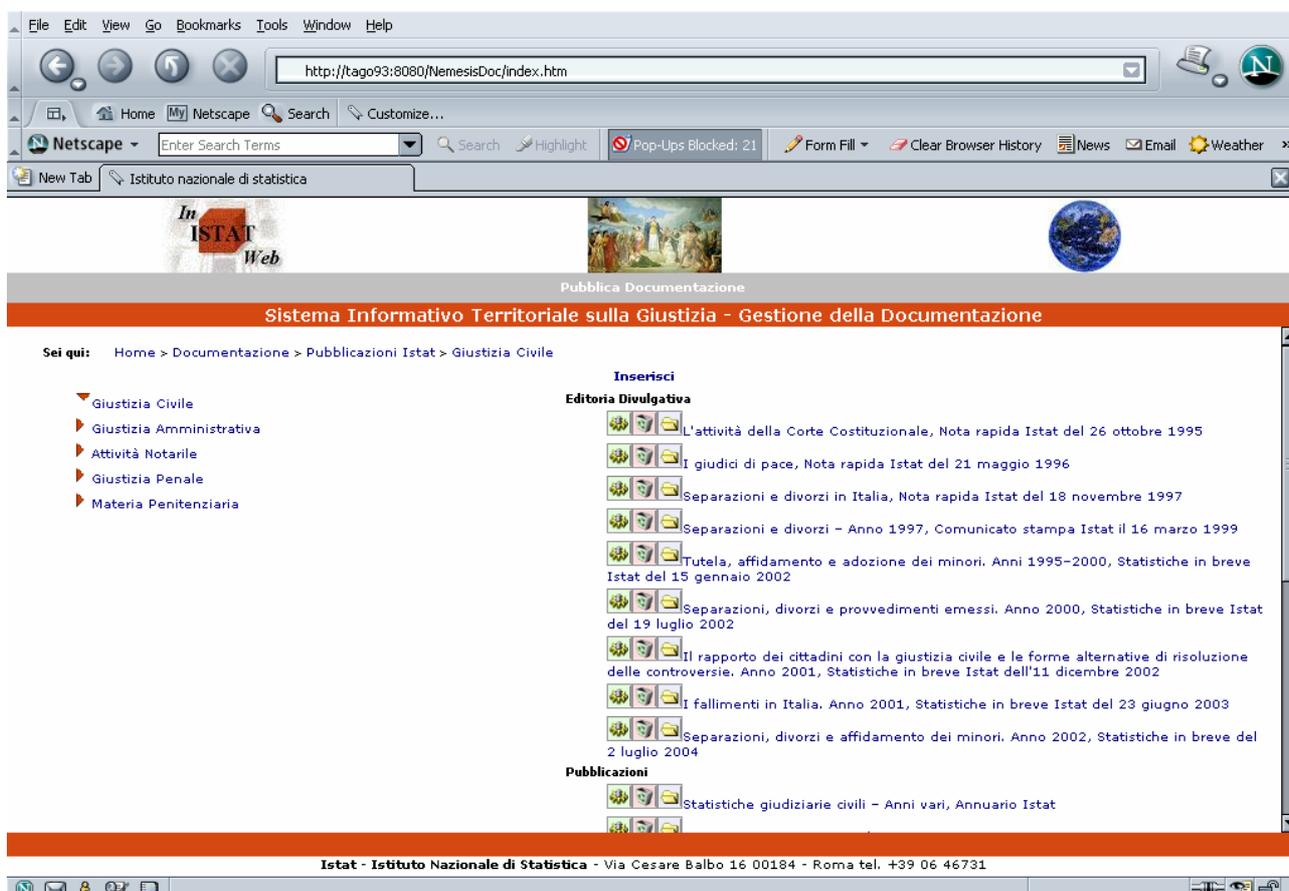


Fig.16 – SITG-Admin – inserimento di un nuovo elemento informativo.

7.2. Manutenzione di un'indagine esistente

Come risulterà evidente fra poco, nella manutenzione di un'indagine, qualsiasi situazione reale potrà, nella sostanza, essere ricondotta ad uno degli scenari od alla combinazione di più scenari fra quelli di seguito descritti.

7.2.1. Inserimento di nuovi microdati

Per tutte le indagini trattate nell'ambito del sistema informativo territoriale sulla giustizia, ogni anno risulterà disponibile un nuovo file di microdati validati da

inserire nel sistema. Nell'ipotesi in cui la struttura dell'indagine non sia variata rispetto all'ultimo caricamento di microdati effettuato ed il ricercatore non intenda modificare in alcun modo il set di tavole da pubblicare, il caso in esame può essere gestito in modo estremamente semplice tramite la sequenza di attività di seguito riportate:

1. caricamento del file dei microdati per il nuovo anno tramite *SITG-Load*;
2. produzione dei macrodati relativi al nuovo anno tramite la funzionalità di caricamento fonte dati dell'applicativo *SITG-Admin*;
3. generazione, tramite *SITG-Admin*, di tutte le tavole associate all'indagine in esame per il nuovo anno;
4. verifica e validazione delle tavole generate al punto 3 tramite il *SITG-Web* in ambiente intranet;
5. pubblicazione in ambiente Internet delle tavole per il nuovo anno tramite l'apposita funzionalità dell'applicativo *SITG-Admin*.

Come è facile osservare, il caso in esame può essere gestito senza alcun intervento di carattere informatico sul sistema, ma semplicemente attraverso alcune attività da parte dell'utente statistico. In definitiva, le tavole per il nuovo anno possono essere rese disponibili alla comunità Internet nel giro di poche ore (necessarie per il caricamento dei microdati, per la produzione dei macrodati e per la generazione delle tavole) o addirittura in pochi minuti (nel caso di indagini caratterizzate da un numero non troppo elevato di record per anno).

7.2.2. Modifiche sugli indicatori

SITG-Admin consente di effettuare tutta una serie di modifiche sui vari elementi costitutivi di un indicatore. Tipicamente, tali modifiche avverranno con maggior frequenza nella fase di definizione di un nuovo indicatore, ovvero prima della sua pubblicazione su Internet, poiché potranno verificarsi cambiamenti nella denominazione dell'indicatore, nella definizione della regola di calcolo da applicare, nel formato da utilizzare, etc.

Tuttavia, le modifiche più interessanti riguardano il momento successivo alla pubblicazione dell'indicatore, ossia i cambiamenti che nel corso del tempo possono intervenire a variare la definizione di un indicatore da un periodo di riferimento all'altro.

Una prima questione da tenere certamente presente a tale riguardo è la variazione, in funzione del tempo, del livello di dettaglio territoriale di validità di un indicatore. Per ciascun anno di riferimento, infatti, ogni indicatore può risultare statisticamente valido per un diverso livello di dettaglio²⁵. *SITG-Admin* consente di gestire tale variabilità in modo molto flessibile, permettendo all'utente statistico, nella definizione di un indicatore, di associare a ciascun livello di dettaglio di ogni gerarchia territoriale disponibile un anno di inizio e fine validità. Ad esempio:

²⁵ Ad esempio, l'indicatore I_x potrebbe essere valido, rispetto alla gerarchia amministrativa, fino al livello regionale per l'anno 2001 e fino al livello provinciale per l'anno 2002.

Indicatore I_x

Gerarchia Territoriale	Livello Territoriale	Inizio Validità	Fine Validità
Amministrativa	Ripartizione (3 modalità)	2000	-
Amministrativa	Ripartizione (5 modalità)	2000	-
Amministrativa	Regione	2001	-
Amministrativa	Provincia	2002	2003
Giudiziaria	Distretto di Corte di Appello	2000	-
Giudiziaria	Circondario	2001	-

In tal modo, nel calcolo dell'indicatore I_x , il sistema genererà, ad esempio, i valori relativi al livello di dettaglio regionale soltanto a partire dal 2001 e per gli anni successivi, mentre quelli a livello provinciale saranno calcolati soltanto per gli anni 2002 e 2003.

Il secondo importante tipo di modifica che può intervenire nel corso del tempo riguarda variazioni nella regola di calcolo dell'indicatore²⁶. In questo caso, il cambiamento può essere gestito secondo due diversi approcci:

- a) *Modificare l'indicatore in esame, variando opportunamente la regola di calcolo.* Il vantaggio di tale soluzione è che l'indicatore così modificato potrà essere immediatamente utilizzato per il calcolo dei valori ad esso associati per i nuovi anni di riferimento, senza alcuna modifica da effettuare sulle tavole coinvolte. Il problema è che lo statistico dovrà avere cura di modificare di nuovo l'indicatore, ripristinando la vecchia regola di calcolo, nel caso in cui volesse ricalcolarlo per gli anni di riferimento precedenti alla modifica.
- b) *Definire un nuovo indicatore identico a quello in esame, tranne che per la regola di calcolo che sarà opportunamente modificata per riflettere il cambiamento intervenuto.* Il "vantaggio", in tal caso, deriva dalla possibilità di associare ad ogni periodo temporale di riferimento un indicatore definito secondo la corretta regola di calcolo. Lo "svantaggio", al contrario, consiste nella necessità, per lo statistico, di modificare le tavole coinvolte per gli anni di riferimento successivi, creando una nuova versione in cui il vecchio indicatore viene sostituito con l'indicatore appena definito.

7.2.3. Modifiche sulle tavole

SITG-Admin consente di gestire, attraverso il meccanismo delle versioni, qualsiasi tipo di modifica sulle tavole in modo molto flessibile. Scendendo maggiormente in dettaglio, ad ogni tavola sono associate, in generale, una o più versioni, ciascuna delle quali può essere separatamente generata e pubblicata per uno specifico anno di riferimento.

Nel momento in cui viene definita una nuova tavola, il sistema le associa automaticamente un codice identificativo ed un numero di versione (pari ad 1). Successivamente, ogni volta che viene effettuata una modifica sulla tavola, al momento del salvataggio, il sistema richiede che l'utente scelga se sovrascrivere l'ultima versione attualmente presente nel sistema, ovvero di salvare la definizione corrente come nuova versione della tavola in esame (con valore pari a quello dell'ultima versione memorizzata incrementato di 1).

Al momento della generazione o della pubblicazione, l'utente può scegliere non soltanto l'anno di riferimento, ma anche la particolare versione della tavola rispetto alla quale sarà effettuata l'operazione. In tal modo, si ha la possibilità di

²⁶ Che potrebbe essere indotta, soltanto per fare un esempio, da mutamenti normativi o altri fattori esterni.

gestire praticamente qualsiasi tipo di modifica possa intervenire nel corso del tempo su una tavola, sia a livello di layout (titolo, titoli di testata, note, etc.), sia per quanto riguarda la lista degli indicatori in essa contenuti.

Si faccia l'ipotesi, per fare un esempio, che una tavola T_x sia stata definita per la prima volta e che sia stata memorizzata nel sistema con codice 001 e versione 1. Si supponga, altresì, che sia stata generata e pubblicata per l'anno 2000 e che siano successivamente intervenuti due cambiamenti (salvati rispettivamente come versione 2 e 3 della tavola 001): il primo nel 2001 (con l'aggiunta di un nuovo indicatore), il secondo nel 2003 (con un cambiamento nel titolo della tavola). Nel caso in cui la tavola in esame sia già stata pubblicata fino al 2001 e soltanto generata per il 2002, la situazione potrà essere riassunta nel modo seguente:

Tavola	Versione	Stato	Anno
001	1	Pubblicata	2000
001	2	Pubblicata	2001
001	2	Generata	2002
001	3	Definita	-

7.2.4. Modifiche sulla struttura dell'indagine

Si tratta, in questo caso, dello scenario di modifica più complesso e di maggior impatto fra tutti quelli possibili perché il cambiamento, questa volta, coinvolge la vera e propria struttura dei dati rilevati. Tale evento può verificarsi sia per motivi interni (ad esempio perché i responsabili dell'indagine hanno stabilito di modificare il modello di rilevazione), sia per cause esterne (mutamenti organizzativi o normativi, solo per fare un paio di esempi). Tali cambiamenti, in ogni caso, si concretizzano, fondamentalmente, nell'aggiunta di una nuova variabile rilevata o nella modifica²⁷ o eliminazione di una variabile precedentemente rilevata.

La gestione di cambiamenti di questa natura comporta necessariamente anche un intervento di carattere informatico sul sistema, i cui dettagli sono riportati di seguito:

1. modifica delle strutture dati relative ai microdati dell'indagine coinvolta;
2. modifica delle strutture dati relative ai macrodati dell'indagine coinvolta;
3. modifica delle procedure software per il caricamento dei microdati relativi all'indagine coinvolta;
4. modifica delle procedure software per la produzione dei macrodati relativi all'indagine coinvolta.

A questo punto, risulta possibile procedere con il trattamento dell'indagine secondo la nuova struttura. In particolare, tutti gli indicatori precedentemente definiti risulteranno ancora utilizzabili²⁸; inoltre, sarà possibile definire nuovi indicatori rispetto alle eventuali nuove variabili rilevate.

²⁷ Una possibile modifica è quella della variazione nel numero e tipo di modalità associate ad una variabile qualitativa.

²⁸ Ovviamente, gli indicatori definiti rispetto a variabili non più rilevate potranno essere calcolati soltanto fino all'anno di ultima rilevazione di tali variabili; mentre per quelli associati a variabili qualitative modificate, il sistema utilizzerà per ogni anno di riferimento le specifiche modalità utilizzate nella corrispondente rilevazione, quindi tali modifiche non avranno alcun effetto sui valori calcolati per gli anni precedenti alle variazioni introdotte.

7.2.5. *Modifiche sulla documentazione*

Tutte le modifiche sulla documentazione possono essere facilmente gestite in modo analogo a quanto precedentemente illustrato in §.7.1.6. In sostanza, la medesima interfaccia web di *SITG-Doc* consente all'utente statistico di apportare modifiche a tutti gli oggetti precedentemente definiti, variandone sia la definizione, sia la categoria di appartenenza.

8. Considerazioni conclusive

Il sistema di gestione e manutenzione rende il processo di produzione e diffusione delle tavole statistiche del Servizio Giustizia più efficiente ed efficace, senza impattare pesantemente sulle precedenti modalità lavorative degli statistici. Questi possono continuare ad usare i propri strumenti tradizionali (*SAS*, *MS Access*, etc.) per costruire e testare (nella fase di progettazione concettuale) indicatori e tavole statistiche, utilizzando il sistema di gestione *SITG-Admin* soltanto per inserire le specifiche di calcolo degli indicatori prescelti e, successivamente, definire, generare e pubblicare le tavole statistiche.

Potrebbero, però, anche svolgere l'intero processo tramite l'ausilio del solo sistema di gestione, calcolando gli indicatori con il *SITG-Admin* direttamente in fase di test.

Inoltre, il sistema fornisce, attraverso le funzionalità per la gestione delle categorie, la possibilità di variare in qualsiasi momento ed in modo estremamente flessibile l'organizzazione delle informazioni pubblicate su web, sia per far fronte alla questione dell'inserimento nel sistema informativo di nuove aree o materie di indagine o, più semplicemente, per l'eventuale ristrutturazione dei contenuti correnti.

Un valore aggiunto del sistema di gestione e manutenzione consiste nella sua possibilità di estensione al trattamento anche di altre indagini, eventualmente esterne al Servizio Giustizia. In effetti, l'unico vincolo che il sistema impone sotto il profilo tecnico è che i macrodati della generica indagine siano rappresentati, all'interno della base di dati, con uno schema *a stella* o, nel caso più generale, *a fiocco di neve*. Tutto il resto, cioè la possibilità di definire nuove variabili statistiche, nuovi indicatori assoluti o derivati sulle variabili, nuove tavole statistiche sugli indicatori e la successiva generazione e pubblicazione delle tavole stesse, può essere ottenuto direttamente e semplicemente, purché sia soddisfatto il vincolo sulla struttura dello schema dei macrodati. Ciò significa che chiunque sia in grado di strutturare i macrodati delle proprie indagini secondo uno schema a stella o a fiocco di neve può usufruire del presente sistema di gestione e manutenzione senza dover fare altro. Tale obiettivo, evidentemente, può essere ottenuto con percorsi diversi in funzione dell'architettura del sistema informatico di partenza.

Si riportano, di seguito, alcuni possibili scenari:

- ✓ Se il sistema informatico sorgente non ha una sua base di dati (perché si tratta, ad esempio, di file piatti trattati direttamente con *SAS*), il sistema di gestione diventa utilizzabile creando un apposito schema a stella o a fiocco di neve per ogni indagine da lavorare e sviluppando una semplice procedura software di caricamento dei files piatti relativi ad ogni specifica indagine. Per altro, anche per quanto riguarda il software di caricamento, l'applicativo *SITG-Load* può essere facilmente adattato per gestire il caricamento di qualsiasi indagine con microdati disponibili su file piatto. Tale scenario, in effetti, è quello che si verifica quando si ha la necessità di aggiungere una nuova indagine sulla giustizia.
- ✓ Nel caso in cui il sistema informatico sorgente abbia una base di dati gestita con un RDBMS diverso da *Oracle*, il sistema di gestione e manutenzione può essere utilizzato con un approccio del tutto analogo a quello appena descritto, con l'unica differenza che la componente software da realizzare sarà di tipo ETL, con estrazione dalla base di dati del sistema sorgente e caricamento in una base di dati in ambiente *Oracle* con schema a stella o a fiocco di neve.
- ✓ Infine, dietro l'ipotesi che il sistema informatico sorgente presenti una componente dati già disponibile in ambiente *Oracle*, il sistema di gestione e manutenzione risulta applicabile in maniera quasi immediata, dal momento che, in questo caso, risulta sufficiente creare, a partire dallo schema di ciascuna indagine da elaborare, un insieme di viste materializzate tali da ricreare uno schema a stella o a fiocco di neve.

Bibliografia

- [1] Renzetti M., Sindoni G., Tininini L., Urbano A. (2004), *The Italian judicial statistical information system*, Physica-Verlag, Atti del 16° Simposio IASC-Compstat 2004, Praga 23-27 agosto 2004.
- [2] Veltri L., Allegra G., Cazora L., Vergnetta M., Vona S. (2004), *Manuale Tecnico: SS0201.04 SITG – ADMIN*.
- [3] Sindoni G., De Francisci S., Paolucci M., Tininini L. (2002), *Experiences in developing a spatio-temporal information systemI*, Research in Official Statistics 5, 45-57.
- [4] Atzeni P., Ceri S., Paraboschi S., Torlone R. (1999), *Basi di dati*.
- [5] Kimball R. (1996), *The data warehouse toolkit*, John Wiley & Sons.
- [6] *Oracle Forms Developer and Reports Developer Release 6i: Guidelines for Building Applications*.
- [7] *Oracle Forms Developer Release 6i: Form Builder Reference*.