

Raccolta dati via Web: un'esperienza di cinque anni

Piero Leone

Terna Spa – Direzione Pianificazione e Sviluppo Rete - Statistiche
email:piero.leone@terna.it

(versione provvisoria)

Sintesi:

Fra i primi in Italia ad utilizzare internet per la raccolta dei dati in adempimento del PSN, la scelta operata, indubbiamente di successo, ha presentato vantaggi notevoli con eliminazione di enormi quantità di carta e semplificazione della raccolta senza ridondanze; si sono evidenziati peraltro anche alcuni svantaggi soprattutto per i fornitori di grandi quantità di dati. Questo problema è stato affrontato nell'ultimo quinquennio mettendo in opera alcune soluzioni rese possibili dall'informatica grazie anche a linguaggi open-source multiplatforma.

Parole chiave:

Statistiche elettriche, Raccolta dati Internet

1. Introduzione

Il settore elettrico in Italia, a partire dal 1999, è stato sottoposto ad un profondo processo di riorganizzazione e liberalizzazione, con il passaggio da un sistema dominato dal monopolista ENEL ad un sistema tendenzialmente concorrenziale. *Terna Rete Elettrica Nazionale Spa* è la società responsabile in Italia della trasmissione e del dispacciamento dell'energia elettrica sulla rete ad alta e altissima tensione su tutto il territorio nazionale; in pratica Terna governa tutto il movimento di elettricità sulle “autostrade elettriche” garantendo che l'energia prodotta dalle centrali giunga fino alle reti dei distributori che la portano nelle nostre case. Oltre a questa prioritaria funzione di servizio pubblico, Terna, tramite il proprio ufficio di statistica, espleta anche quella di responsabile SISTAN delle statistiche ufficiali sul settore elettrico in Italia. Tale funzione è stata svolta fino al 1999 in Enel, poi nel GRTN (Gestore della Rete di Trasmissione Nazionale) e dal novembre 2005 in Terna, a seguito dell'unificazione del GRTN in Terna.

La presente relazione mira a presentare l'esperienza sviluppata a partire dal 2001 di adozione di un innovativo sistema di raccolta dati attraverso Internet. L'obiettivo della relazione è fornire alcuni spunti utili a chi si occupa di raccolta dati a fini statistici per ottimizzare al meglio la progettazione, la realizzazione e la gestione di un sistema di rilevazione informatizzato su Web.

2. Le Statistiche del settore elettrico

La conoscenza puntuale e dettagliata del sistema elettrico, ed energetico in generale, rappresenta un fattore strategico per la politica di sviluppo economico del paese. Per questo motivo, le statistiche del settore elettrico in Italia hanno una lunga tradizione. L'ENEL, sin dalla sua costituzione nel 1963, si è occupata delle statistiche sul settore elettrico, assumendone l'onere ben prima della costituzione del SISTAN.

Numerosi sono gli aspetti descritti dalle statistiche pubblicate dall'ufficio di statistica di Terna; i principali sono raggruppabili in tre gruppi:

1. Consistenza e caratteristiche degli impianti di produzione di energia elettrica
2. Dimensioni e tipologie di produzione di energia elettrica
3. Utilizzi finali di energia elettrica

Il principale prodotto cartaceo è rappresentato dal volume annuale "Dati statistici sull'energia elettrica in Italia", diffuso gratuitamente, e disponibile anche in formato elettronico sul sito www.terna.it nella sezione Dati Statistici

3. La raccolta dati su modelli cartacei fino al 2000

Fino al 2000, i dati statistici sul settore elettrico venivano raccolti seguendo sostanzialmente due fonti: gli archivi gestionali interni ENEL e la rilevazione dell'attività degli altri operatori - imprese municipalizzate ed autoproduttori - tramite moduli cartacei. La rilevazione cartacea costituiva un impegno in termini di risorse non trascurabile. Per dare un'idea dell'impegno:

- Circa 1.500 imprese elettriche rilevate
- Circa 30.000 modelli cartacei personalizzati stampati ogni anno, di cui la metà spediti alle imprese elettriche e il resto a disposizione del personale ENEL per assistenza o eventuali solleciti

4. La raccolta dati via Web dal 2001: il sistema GSTAT

4.1 Caratteristiche di GSTAT

La ristrutturazione del settore elettrico ha reso necessario una riorganizzazione anche delle statistiche, sia in termini di presentazione dei dati, sia, soprattutto, delle modalità di raccolta. Con l'obiettivo di sfruttare al meglio le opportunità rappresentate dal Web come strumento di veicolazione dell'informazione, l'ufficio di statistica dell'allora GRTN ha avviato un innovativo sistema di rilevazione denominato GSTAT. Praticamente, al posto dei vecchi moduli cartacei, i dati statistici richiesti agli operatori elettrici vengono inseriti direttamente tramite apposite maschere on-line di GSTAT, ad accesso controllato. Le caratteristiche principali dell'applicazione GSTAT sono:

1. Accesso tramite User-id e Password assegnata a ciascun Operatore Elettrico
2. Moduli di gestione delle anagrafiche

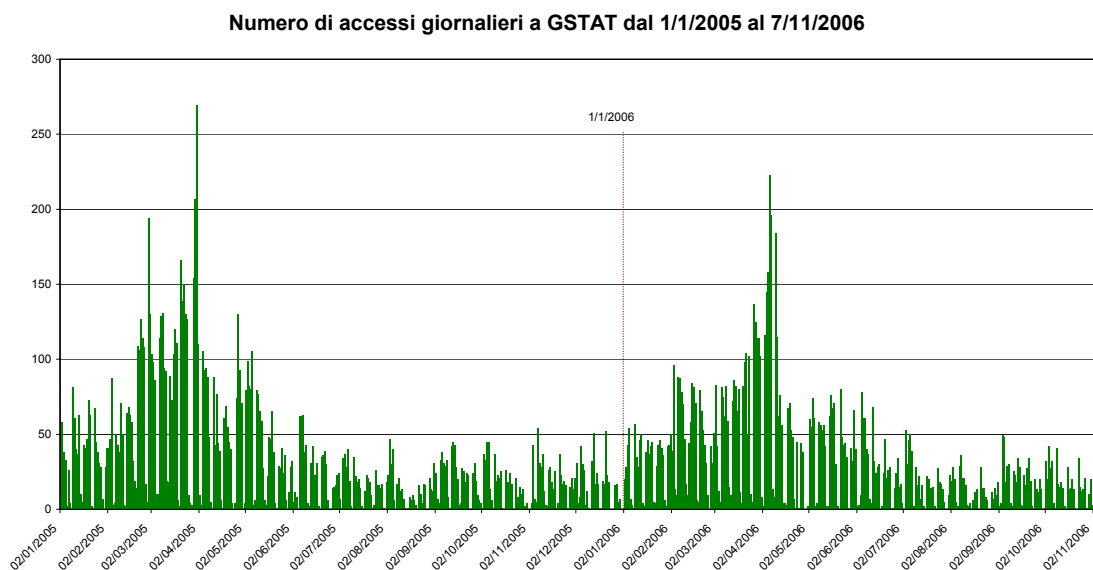
3. Moduli di gestione della consistenza e caratteristiche dei propri impianti di generazione
4. Moduli di inserimento dati di produzione, acquisto, vendita e consegna ai clienti finali dell'energia elettrica
5. Controllo e segnalazione on-line di tutte le incongruenze formali e, dove possibile, sostanziali in fase di inserimento dati

Una caratteristica fondamentale, che rende l'applicazione GSTAT strutturalmente diversa da quelle ampiamente diffuse sul Web per la compilazione di indagini o questionari, è che la personalizzazione dell'accesso a GSTAT comporta la gestione di tutti i dati caratteristici degli impianti di generazione dell'operatore e dei loro "consueti" flussi. In pratica, le videate per il caricamento dati presentano la situazione in essere in termini di parco di generazione, permettendo eventualmente la loro modifica.

4.2 Gli accessi a GSTAT

Per avere un'idea del "carico di lavoro" di GSTAT, in figura 1 è riportato il numero di accessi registrato in ciascun giorno dal 1 gennaio 2005 al 7 novembre 2006.

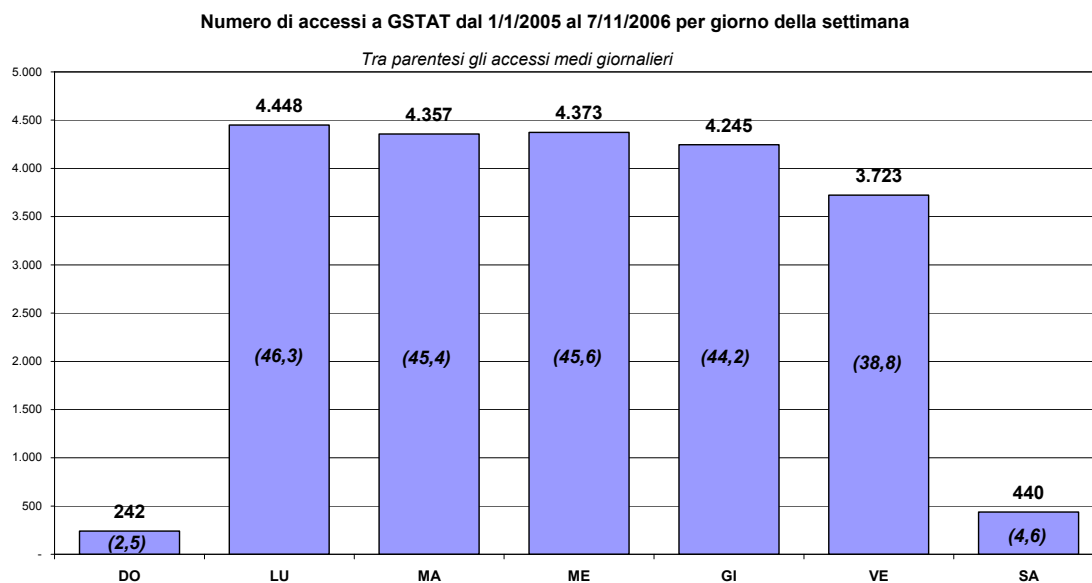
Figura 1



In questo periodo sono stati raccolti i dati per l'elaborazione di due annuari; la concentrazione massima di accessi si ha nel periodo febbraio-aprile, con punte nei giorni a ridosso delle scadenze indicate per il caricamento che superano i 200 accessi. Negli altri periodi dell'anno il caricamento dati è limitato agli operatori più grandi, i quali forniscono le statistiche mensili a carattere provvisorio.

In figura 2 il numero di accessi per lo stesso periodo è suddiviso per giorni della settimana; la cosa più sorprendente, forse, è che ci sono dei soggetti che utilizzano il fine settimana per espletare i loro compiti statistici.

Figura 2



4.3 Lo sviluppo di GSTAT

Dall'esperienza acquisita nello sviluppo di GSTAT si può affermare che è fondamentale l'interazione tra tre diverse "competenze": lo sviluppo del software, la gestione del sistema informatico, la definizione della metodologia statistica. Di seguito, per ognuna di esse sono elencati quelli che si ritengono i fattori che permettono un efficace disegno ed implementazione dell'applicazione.

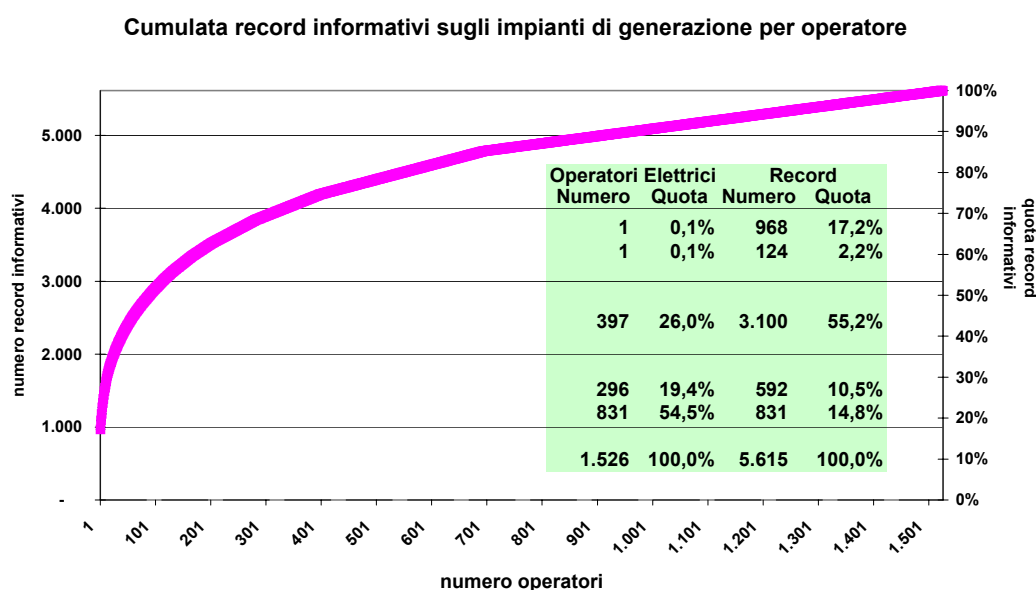
1. **Sviluppo del software:** un'applicazione per la raccolta dati via internet complessa necessita di competenze specifiche in termini di linguaggi di programmazione che normalmente sono esterne all'azienda, proprie di aziende di sviluppo software. E' quindi fondamentale che il ricorso a sviluppatori esterni sia gestito in modo efficiente, soprattutto in termini di specifica per lo sviluppo. L'esperienza di GSTAT suggerisce che è opportuno procedere per gradi, avviando il sistema per step successivi. Molto spesso la realtà da rilevare presenta particolarità ed eccezioni che difficilmente si riesce a prevedere completamente. L'estensione dell'applicazione per gradi e l'affinamento del software "sul campo" rappresentano i fattori chiave dal punto di vista dello sviluppo del software.
2. **La gestione del sistema informatico:** il sistema informatico su cui l'applicazione internet risiede necessita di competenze specifiche, di norma interne all'azienda. Da questo punto di vista, l'esperienza di GSTAT suggerisce di mantenere una costante interazione con i responsabili della gestione del sistema informatico già nella fase di sviluppo dell'applicazione, in quanto l'interazione di un'applicazione di raccolta dati via internet con le politiche della sicurezza e con le politiche di gestione dell'intero sistema informatico (reti, server, sito aziendale, intranet, ecc..) possono condizionare fortemente l'efficiente sviluppo del progetto.

3. **Metodologia statistica:** dal punto di vista metodologico l'esperienza di GSTAT porta a evidenziare alcuni fattori che si ritiene siano in generale importanti per il successo dello sviluppo dell'applicazione:
 - a. Efficiente gestione dei dati di stato
 - b. Assicurare l'unicità di input
 - c. Predisporre utilità a favore degli operatori che inseriscono i dati
 - d. Ricorrere il più possibile a tabelle tipologiche per rendere flessibile la struttura dell'applicativo
 - e. Efficiente gestione del blocco/modificabilità dei dati
 - f. Efficiente gestione della storicizzazione dei dati anagrafici e di stato

4.4 La gestione di GSTAT

Uno degli elementi che ha fortemente condizionato le caratteristiche di GSTAT risiede nella forte concentrazione di numerosità di informazioni statistiche in pochi operatori.

Figura 3



In figura 3 è riportato il numero di record informativi¹ relativi agli impianti di generazione di energia elettrica per ciascun operatore. L'operatore più grande deve gestire le informazioni statistiche relative a 968 record, mentre 831 dei 1.526 operatori elettrici rilevati gestiscono ognuno un solo record, e 296 ne gestiscono 2. Il modo di accesso, caricamento e controllo dei dati necessariamente non può essere lo stesso per operatori con mole di dati così diversi.

Per gestire al meglio questo aspetto si è scelto di mantenere l'unicità di gestione dati in GSTAT, fornendo agli operatori la possibilità di inserire i dati direttamente da file oltre

¹ In pratica un record informativo è un insieme di dati che descrive le caratteristiche tecniche di una sezione di impianto di generazione di energia elettrica oggetto di rilevazione

che manualmente. E' stata quindi sviluppata una funzionalità che permette il caricamento dei dati da un file in formato XML (eXtensible Markup Language). Tale formato, del tutto indipendente dall'ambiente di sistema in uso, permette di recepire in modo semplice e chiaro grandi quantità di dati.

4.5 L'esperienza di GSTAT

Nei cinque anni in cui le statistiche del settore elettrico sono state raccolte tramite GSTAT l'applicativo ha subito successivi ampliamenti e modifiche, sia in termini di tipologia di dati raccolti, sia in termini di affinamento delle modalità di visualizzazione e controllo. In termini schematici le principali considerazioni di carattere generale che l'esperienza di gestione di GSTAT suggerisce sono le seguenti:

- Il passaggio da una raccolta dati su moduli cartacei ad una gestione via Internet ha comportato, oltre ad una riorganizzazione del flusso di gestione dati, anche una riorganizzazione e riqualificazione del personale interno addetto al controllo e validazione. Essi hanno assunto sempre più un ruolo di "consulente" che assiste i soggetti che inseriscono i dati. La principale implicazione è che per svolgere la "consulenza" al meglio si è reso necessario indurre una più approfondita conoscenza dell'intero processo di gestione dei dati, con una sensibilità volta più alla sostanza che alla forma di ciò che i dati rappresentano. In pratica, si è passati da una ripartizione del lavoro per filoni di dati, in cui ciascuno era "esperto" di alcune tipologie, ad una ripartizione per gruppi di operatori elettrici, per cui ciascun validatore segue l'intero processo del gruppo di operatori elettrici di cui è responsabile.
- Un chiaro miglioramento si è registrato, inoltre, anche nel clima di collaborazione tra il personale, in quanto si è passati da una ricerca di riconoscimento individuale improntata alla esclusiva conoscenza di aspetti "di risultato" parziali, ad un riconoscimento individuale più legato alla conoscenza e condivisione di aspetti "di processo".
- Si è percepito un generale innalzamento del livello professionale dei soggetti incaricati di inserire i dati in GSTAT. Tendenzialmente, la compilazione dei modelli cartacei era demandata a personale di segreteria, mentre l'accesso ad un applicativo Internet è stato considerato più "delicato" e bisognoso di personale più esperto.
- La qualità dei dati è risultata notevolmente migliorata, non solo per l'eliminazione di tutte le incongruenze formali direttamente in fase di input, ma anche per la maggiore chiarezza del processo di rilevazione on-line rispetto a diversi modelli cartacei tra loro svincolati.
- La gestione dei solleciti al caricamento si è rivelata molto più agevole, in quanto i riferimenti anagrafici dei soggetti addetti al caricamento sono direttamente aggiornati in GSTAT, e la possibilità di un colloquio costante con gli addetti al caricamento ha permesso "pungolare" efficacemente i ritardatari occasionali o abituali.

In definitiva l'esperienza di GSTAT appare estremamente positiva sotto diversi punti di vista, e può sicuramente costituire un riferimento valido per lo sviluppo di analoghi applicativi per la rilevazione dati via Web.