

**n. 2/2009**

## **Tecniche di cattura dati nei processi di produzione statistica**

*R. Balestrino e A. Gaucci*

Le collane esistenti presso l'ISTAT - *Rivista di Statistica Ufficiale*, *Contributi ISTAT* e *Documenti ISTAT* - costituiscono strumenti per promuovere e valorizzare l'attività di ricerca e per diffondere i risultati degli studi svolti, in materia di statistica ufficiale, all'interno dell'ISTAT, del SISTAN, o da studiosi esterni.

La *Rivista di Statistica Ufficiale* accoglie lavori che hanno come oggetto la misurazione dei fenomeni economici, sociali, demografici e ambientali, la costruzione di sistemi informativi e di indicatori, le questioni di natura metodologica, tecnologica o istituzionale connesse al funzionamento dei sistemi statistici e al perseguimento dei fini della statistica ufficiale.

I lavori pubblicati in *Contributi ISTAT* sono diffusi allo scopo di stimolare il dibattito intorno ai risultati preliminari di ricerca in corso.

I *Documenti ISTAT* forniscono indicazioni su linee, progressi e miglioramenti di prodotto e di processo che caratterizzano l'attività dell'Istituto.

Il Comitato di redazione esamina le proposte di lavori da pubblicare nelle tre collane sopra indicate. Quelli pubblicati nei *Contributi ISTAT* e nei *Documenti ISTAT* sono valutati preventivamente dai dirigenti dell'Istituto, mentre i lavori pubblicati nella *Rivista di Statistica Ufficiale* sono subordinati al giudizio di referee esterni.

---

# DOCUMENTI ISTAT

---

**n. 2/2009**

**Tecniche di cattura dati nei processi  
di produzione statistica**

*R. Balestrino(\*) e A. Gaucci(\*)*

(\*) ISTAT - Servizio Metodologie, tecnologie e software per la produzione dell'informazione statistica

I lavori pubblicati riflettono esclusivamente le opinioni degli autori e non impegnano la responsabilità dell'Istituto

**Contributi e Documenti Istat 2009**

Istituto Nazionale di Statistica  
Servizio Produzione Editoriale

Produzione libraria e centro stampa:  
*Carla Pecorario*  
Via Tuscolana, 1788 - 00173 Roma

**Sommario:** Il presente lavoro intende evidenziare lo stato dell'arte dell'introduzione di tecniche innovative nella fase di raccolta dati delle rilevazioni statistiche condotte dall'Istituto. Viene illustrata la piattaforma web, *indata.istat.it*, dedicata all'acquisizione telematica dei dati primaria e secondaria, attraverso una descrizione dell'architettura, delle funzioni e dei vincoli che si sono individuati a garanzia di un ambiente standard, efficace e sicuro. Vengono forniti dati quantitativi che dimostrano il grado di diffusione delle nuove tecniche tra le diverse classi di rispondenti e per le varie tipologie di rilevazione dei domini statistici oggetto di studio. Sono riportate alcune considerazioni sull'atteggiamento dei rispondenti e sull'utilizzo della piattaforma *indata.istat.it*, e presentate le prospettive di sviluppo della cattura dati, contributo certo all'efficienza e alla qualità dei processi di produzione statistica.

**Parole chiave:** cattura dati, questionario elettronico, Web survey, CAWI

Le collane esistenti presso l'ISTAT - Contributi e Documenti - costituiscono strumenti per promuovere e valorizzare l'attività di ricerca e per diffondere i risultati degli studi svolti, in materia di statistica ufficiale, all'interno dell'ISTAT e del Sistan, o da studiosi esterni.

I lavori pubblicati Contributi Istat vengono fatti circolare allo scopo di suscitare la discussione attorno ai risultati preliminari di ricerca in corso.

I Documenti Istat hanno lo scopo di fornire indicazioni circa le linee, i progressi ed i miglioramenti di prodotto e di processo che caratterizzano l'attività dell'Istituto.

I lavori pubblicati riflettono esclusivamente le opinioni degli autori e non impegnano la responsabilità dell'Istituto.



# INDICE

## **1. TECNICHE INNOVATIVE NELLA FASE DI ACQUISIZIONE DEI DATI**

## **2. WEB SURVEY**

## **3. IL SITO UFFICIALE DI ACQUISIZIONE DATI DELL'ISTITUTO**

3.1 BREVE STORIA DEL SITO

3.2 REQUISITI, VINCOLI E OFFERTA FUNZIONALE

3.3 SPECIFICHE ARCHITETTURALI

3.4 LINGUAGGI, FORMATI E STRUMENTI DI SVILUPPO

3.5 L'APPLICATIVO DATISTAT PER L'ACQUISIZIONE SECONDARIA DEI DATI

## **4. ELEMENTI DI MONITORAGGIO**

4.1 LO SVILUPPO DELLE APPLICAZIONI DI CATTURA DATI

4.2 TASSO DI RISPOSTA ELETTRONICA

4.3 MONITORAGGIO DELLE AZIONI DEI RISPONDENTI

4.4 APPLICAZIONE DEL SOFTWARE WEBALIZER

## **5. PROSPETTIVE**

## **6. ALLEGATO – RILEVAZIONI ATTIVE E TERMINATE**



## 1. TECNICHE INNOVATIVE NELLA FASE DI ACQUISIZIONE DEI DATI

Nell'ambito della statistica ufficiale la funzione di acquisizione dati è di primaria importanza per l'alimentazione dei processi di produzione attraverso cui, a partire dai microdati rilevati, l'informazione statistica viene resa disponibile al Paese. Allo stesso tempo la raccolta dei dati è una fase che incide fortemente sui costi totali della rilevazione: quando poi sono svolte in modalità tradizionale le operazioni di raccolta risultano particolarmente onerose perché comportano spese per i rilevatori, per la riproduzione e consegna di materiale, per la registrazione dei dati, servizi che normalmente vengono acquisiti all'esterno dell'Istituto. Inoltre, le operazioni di acquisizione, rivolgendosi alle unità rispondenti, siano esse imprese, istituzioni, famiglie o individui, implicano una esposizione dell'immagine dell'Istituto: la semplificazione e l'efficacia dei rapporti tra soggetti sociali e pubblica amministrazione riguarda anche la funzione statistica.

Tra i processi statistici che possono trarre vantaggio dall'innovazione della cattura dati vanno incluse non solo le indagini dirette, che generano attività di acquisizione primaria, dovuta alla compilazione di questionari di rilevazione, ma anche tutte quelle attività di acquisizione secondaria che fanno parte dei programmi operativi dell'Istituto e che richiedono l'uso o il riuso di informazione raccolta e/o prodotta da altre fonti.

L'introduzione di strumenti informatici e telematici ha profondamente trasformato negli ultimi decenni le modalità operative attraverso cui si realizza l'acquisizione, producendo benefici in termini di costi, tempestività e qualità dei dati. Le tecniche tradizionali (postali o face-to-face, con modello cartaceo) vengono sempre di più sostituite o affiancate dalle tecniche assistite da computer.

Le tecniche innovative, che stanno trasformando la classica fase di raccolta dati in data capturing, sono andate specializzandosi sulle diverse classi di rispondenti: CATI (Computer Assisted Telephone Interviewing) e CAPI (Computer Assisted Personal Interviewing) si sono indirizzate prevalentemente alle famiglie; Web Survey quasi esclusivamente, per ora, a imprese e istituzioni. E' evidente come la maturazione, a livello nazionale, delle condizioni dei soggetti sociali e delle infrastrutture di rete potrebbe portare a medio termine ad esplorare la possibilità di estendere le tecniche CAWI (Computer Assisted Web Interviewing) anche alle rilevazioni sulle famiglie. Nelle prossime tornate censuarie, per esempio, l'opzione Web sarà sicuramente offerta ai rispondenti. Già hanno avuto luogo d'altra parte esperienze di test su piccole rilevazioni indirizzate a speciali popolazioni di individui (per esempio esperti nazionali sulla materia del mercato del lavoro).

Va evidenziato inoltre che, per necessità o comodità, i diversi strumenti, inclusi quelli tradizionali, spesso vengono associati nello stesso processo di acquisizione di una rilevazione statistica dando luogo alle tecniche note come *mix-mode*, che permettono di introdurre innovazione senza forzare l'atteggiamento dei rispondenti nella fornitura delle risposte. Questa modalità di applicazione delle tecniche di data capturing ha consentito di superare gli iniziali dubbi dei metodologi statistici preoccupati per la bassa copertura del WEB nelle popolazioni di riferimento e la conseguente difficoltà a garantire disegni campionari efficaci.

In altri termini il ruolo della rete Internet è nullo rispetto al passo di estrazione campionaria e influisce unicamente sulla modalità di risposta delle unità statistiche che vengono estratte con metodi tradizionali: questo garantisce la possibilità di utilizzare le procedure inferenziali classiche. Ambito da indagare maggiormente sarebbe l'influenza delle modalità di acquisizione sulla qualità dei dati raccolti.

Per favorire l'innovazione dei processi di raccolta l'Istituto ha istituito un sito Web ufficiale, raggiungibile all'indirizzo <https://indata.istat.it>, dedicato alle operazioni di acquisizione telematica primaria e secondaria, e disponibile per tutti i settori di produzione interessati ad adottare nei propri processi soluzioni innovative basate su Web.

## **2. WEB SURVEY**

Tra le tecniche di data capturing è in crescente diffusione l'uso dei servizi Internet e in particolare il ricorso al WEB.

Offrire al rispondente la possibilità di compilare online il modello di rilevazione presuppone la disponibilità, presso l'amministrazione che conduce l'indagine, di un web server, visibile sulla rete Internet, su cui vengono allestite pagine web contenenti il questionario in formato ipertestuale o altro formato elettronico; i rispondenti possono compilare e inviare il questionario accedendo per mezzo del loro browser all'indirizzo del web server di Istituto. In questo caso la compilazione si svolge online e i dati immessi dal rispondente sono contestualmente resi disponibili nell'ambiente di produzione che viene alimentato quindi con tempestività massima. La dimensione e complessità del questionario e la eventuale attività di editing contestuale all'immissione dei dati condizionano il tempo di connessione e possono influire sulle prestazioni e, pertanto, in generale, può essere conveniente contenere l'attività di controllo in fase di immissione e puntare sulla tempestività nella disponibilità dei dati stessi. Va evidenziato inoltre che le prestazioni di sistema di qualunque applicazione web dinamica dipendono anche dalle caratteristiche dell'ambiente client e dalle condizioni contrattuali con cui gli Internet Providers erogano i servizi ai rispondenti, fattori fuori del controllo dell'Istituto. Anche la variabilità in termini di sistemi operativi e browser utilizzati, provvisti o meno di plug in necessari per aprire gli oggetti da noi predisposti, può costituire un fattore di criticità con cui confrontarsi, ma questo è un problema noto e condiviso tra tutti coloro che offrono servizi su rete Internet per mezzo di applicazioni Web.

In alternativa alla modalità online, può essere talvolta più conveniente che la compilazione avvenga offline, per esempio per ridurre i tempi di connessione del rispondente o per la complessità del questionario stesso, che richiede preventivamente la consultazione di documentazione o archivi presso il rispondente. I servizi WEB sono essenziali anche in questo caso: il web server infatti deve consentire che il rispondente, accedendo al sito predisposto per questa funzione, possa scaricare sulla propria postazione client il questionario elettronico in formato eseguibile, e, in una successiva sessione, possa restituire il questionario compilato, naturalmente in condizioni di sicurezza.

Condizione che facilita il rispondente, ma genera una più complessa attività di sviluppo applicativo, consiste nell'offrire, durante la compilazione online, la possibilità di visualizzare dati pregressi o effettuare salvataggi parziali del questionario, che possano essere ripresi in momenti successivi per completare la compilazione stessa.

E' conveniente sottolineare, ancora di più in tempi di ristrettezza di finanziamenti da dedicare alla funzione statistica, che la cattura dati, come tutte le soluzioni basate sul WEB, oltre alla tempestività nell'acquisizione, garantisce una significativa riduzione dei costi di raccolta: i ruoli umani dell'intervistatore e dell'operatore di data-entry non sono necessari né si devono sostenere costi di riproduzione fisica e movimentazione del modello di rilevazione. Anche sulla qualità dei dati si possono avere effetti positivi, potendosi svolgere un'attività di controllo contestuale alla compilazione ed eliminare le distorsioni dovute all'intervento del rilevatore.

## **3. IL SITO UFFICIALE DI ACQUIZIONE DATI DELL'ISTITUTO**

### ***3.1 Breve storia del sito***

In Istituto la cattura dati via Internet si è andata diffondendo a partire dalla fine degli anni '90. In quegli anni si cercava di mettere a fuoco la materia del data capturing statistico, su cui non c'era esperienza ma solo suggerimenti e segnalazioni provenienti dall'ambito della statistica ufficiale internazionale (Bureau of Census e Statistics Canada soprattutto). Era necessario comprendere e definire la portata del data capturing nello specifico ambito dei processi di produzione statistica e adeguare le conoscenze tecniche di area informatica per essere in grado di sviluppare un'applicazione Web dinamica, che prevedesse

colloquio e scambio di informazione tra centro (Istituto) e periferia (rispondenti). Va ricordato che in quegli anni, se si escludono le nascenti esperienze del mercato elettronico, i siti Web, soprattutto quelli delle pubbliche amministrazioni, erano piuttosto siti *vetrina*, dove l'informazione viaggiava esclusivamente dal centro alla periferia, e sulla base di menu di interazione e ricerca predefiniti.

A cura del Laboratorio Data Capturing, costituito nell'ambito del Dipartimento di informatica, venne allestito nel '98, in via sperimentale, il primo sito web interattivo dedicato all'acquisizione dei dati, e realizzati i primi prototipi di questionari elettronici compilabili in rete Internet. Il sito era ospitato su una macchina IBM con RAM (unità di memoria centrale) di 256MB, che è stata attiva fino alla fine del 2005, quando il sito ospitava già circa 10 applicazioni di cattura dati per rilevazioni correnti.

Naturalmente furono necessari alcuni anni perché la materia diventasse matura e le nuove tecniche perdessero il carattere di sperimentazione ed entrassero a tutto titolo nei processi di produzione delle indagini correnti. Esperienze significative sono andate maturando nello stesso periodo presso alcune Direzioni di produzione statistica, che hanno sviluppato in autonomia un certo numero di siti web dedicati alla cattura dati per le aree statistiche di loro competenza.

Per finalità di standardizzazione e sicurezza, per contenere i costi di sviluppo, gestione e manutenzione evolutiva dell'ambiente web, oltre che per offrire ai rispondenti un'immagine stabile ed omogenea dell'Istituto, nel corso del 2006 sono state definite, e successivamente ratificate dal Dipartimento di Produzione Statistica, le caratteristiche tecnologiche ed organizzative di un ambiente WEB unificato, dedicato alle funzioni di acquisizione dati, basato sui principi di integrazione, uniformità di indirizzo e rispetto della sicurezza.

Il sito unificato è stato reso disponibile allo stesso indirizzo <https://indata.istat.it>, già attivato a cura del Laboratorio Data Capturing, per continuità e per riconoscimento del carattere di trasversalità e standardizzazione che già qualificava il vecchio sito; ha ospitato fino ad ora 38 applicazioni di cattura dati, per altrettante rilevazioni statistiche, ed è in costante sviluppo. Le caratteristiche fisiche e le capacità prestazionali del sistema sono state definite sulla base dell'attività potenziale massima di cattura dati che si possa prevedere nel medio termine.

### ***3.2 Requisiti, vincoli e offerta funzionale***

Per il sistema unificato di acquisizione dell'Istituto sono stati definiti, attraverso un confronto con le realtà di produzione dell'Istituto, i principali requisiti e vincoli da rispettare.

#### **Requisiti funzionali**

- I rispondenti trovano nel sito *indata.istat.it* il punto unico in cui sono disponibili tutti gli strumenti e le informazioni relativi alle indagini che realizzano cattura dati via Web;
- Lo stile di presentazione è omogeneo per qualsiasi indagine;
- I rispondenti si registrano una tantum al sito e si autenticano tutte le volte che accedono al questionario elettronico, utilizzando user-id e password gestite conformemente alla norma vigente;
- Il sistema offre ai rispondenti, come funzione primaria, diverse possibilità di trasmissione dei dati (questionari compilati in linea, questionari compilati fuori linea, prelievo/deposito di file, trasferimento di questionari via fax...);
- Il sistema offre funzioni secondarie e servizi utili alle attività di raccolta sia per i rispondenti sia per gli utenti statistici (ad esempio documentazione relativa all'indagine, assistenza, monitoraggio degli arrivi e gestione degli archivi delle unità rispondenti).

## Requisiti architettureali

- Il sistema offre alti livelli di affidabilità, tali da limitare al massimo la presenza di punti unici di criticità e garantire al massimo la continuità del servizio;
- Il sistema supporta questionari elettronici e applicazioni di acquisizione dati realizzati con strumenti di sviluppo che si considerano standard per l'acquisizione dati in Istituto.

## Vincoli

- L'acquisizione dei dati attraverso il sistema rispetta il "Codice in materia di riservatezza dei dati personali" (D.L. 196/03) ed in particolare tutte le misure di sicurezza definite nel relativo Allegato B e nel Documento Programmatico sulla Sicurezza dell'Istituto;
- Le nuove soluzioni di cattura dati via Web sono sviluppate secondo il modello definito e ospitate sul sito unificato;
- Le applicazioni di acquisizione sviluppate prima della definizione del nuovo sistema, nel momento in cui ne entrano a far parte, si allineano alle misure minime di sicurezza stabilite;
- Le indagini che acquisiscono dati per mezzo della posta elettronica migrano sul nuovo sistema di acquisizione;
- Il sistema non impone vincoli stringenti di presenza di software sulle postazioni client dei rispondenti né impone l'acquisizione e/o installazione di software su di esse, riducendo al minimo il "disturbo statistico".

## Funzioni

Il focus dell'acquisizione via web è sui dati elementari delle indagini correnti; l'ambito della rilevazione statistica dà luogo prevalentemente alla funzione di *acquisizione primaria* (un rispondente trasmette la sua intervista), ma possono avere luogo anche operazioni di *acquisizione secondaria*: per esempio un organo intermedio territoriale trasmette una raccolta di dati corrispondente alle interviste dei rispondenti, oppure una società di mercato, incaricata della raccolta dati in modalità diversa dal Web, trasmette i file dei microdati raccolti.

La funzione di acquisizione secondaria è inoltre piuttosto comune, al di fuori del contesto delle indagini in senso stretto, nell'ambito del trattamento di informazione proveniente da altre fonti, soprattutto amministrative.

Indipendentemente dal linguaggio scelto per la realizzazione del questionario elettronico, le applicazioni di cattura dati devono innanzitutto rispettare uno schema macro-funzionale comune per essere ospitate sul sito ufficiale. Questo significa che, per ogni indagine, ai rispondenti e ai gestori della rilevazione il sistema offre una cornice funzionale che comprende:

### Accesso sicuro

- *Definizione di USER\_ID e generazione di PASSWORD;*
- *Registrazione/ Autenticazione del rispondente;*
- *Modifica della PASSWORD da parte del rispondente;*
- *Ripristino di PASSWORD iniziale su richiesta del rispondente;*

### Orientamento, documentazione e assistenza

- *Accoglienza al rispondente sul sito con presentazione delle finalità del sito e dell'indagine;*
- *Consultazione e stampa di istruzioni e altro materiale;*
- *Assistenza*, per segnalazione di difficoltà o richieste specifiche (per esempio smarrimento password);
- *Gestione FAQ (Frequently Asked Questions);*

### Compilazione del questionario elettronico

- *Personalizzazione del questionario con dati del rispondente già disponibili presso Istat*

- *Compilazione del questionario online e/o offline;*
- *Controllo dell'immissione dei dati;*
- *Invio ricevuta di conferma ricezione del modello;*
- *Scrittura/archiviazione dei dati acquisiti: i dati immessi dal rispondente devono essere scritti temporaneamente in area sicura o archiviati direttamente su DB;*

### **Prelievo/deposito oggetti**

- *Download:* l'esempio più rilevante è il download da parte del rispondente del questionario elettronico, che potrà essere scaricato anche in versione personalizzata;
- *Upload:* l'esempio più rilevante è l'upload, da parte del rispondente, del suo questionario elettronico compilato;

### **Gestione**

Funzioni disponibili agli utenti interni, responsabili delle indagini, per preparare e gestire agevolmente l'ambiente di amministrazione dei rispondenti di un'applicazione di cattura dati: tabella delle unità da rilevare, tabella delle unità che si registrano, tabella delle password iniziali, tabella dei log di accesso e invio dei dati, etc.

- *Creazione, cancellazione, visualizzazione, modifica, liste di tabelle e record;*

### **Monitoraggio**

- *Monitoraggio dell'attività complessiva di cattura dati e del sito di acquisizione;*
- *Monitoraggio della raccolta via WEB per indagine;*

Pur nel rispetto del corredo funzionale di base, raffinando lo schema delle funzioni fino ad arrivare alle funzioni di massimo dettaglio, una diversificazione si può verificare per rispondere alle esigenze dei singoli processi di raccolta; per fare un esempio, il monitoraggio di una rilevazione che usa organi *intermedi* può richiedere micro-funzioni diverse rispetto a una indagine gestita completamente dal *centro*. Altre volte semplicemente sono le specifiche classi di rispondenti che determinano condizioni diverse di cui tenere conto nello sviluppo delle funzioni elementari.

Sulla base delle finalità già descritte e delle funzioni offerte ai rispondenti e ai responsabili di indagine, l'insieme complessivo dei dati trattati nel sistema di acquisizione si può articolare nelle classi riportate di seguito:

#### **Dati di identificazione del rispondente**

Tipicamente sotto forma di coppia USER\_ID / PASSWORD, assegnati dall'Istituto alle unità rispondenti e trattati secondo la norma vigente

#### **Dati di personalizzazione del questionario**

Sono dati, relativi al rispondente, già in possesso dell'istituto, e che vengono preimpostati sul questionario per contenere il disturbo statistico

#### **Dati - oggetto dell'indagine - forniti dal rispondente**

Sono i dati elementari oggetto di compilazione

#### **Dati pregressi relativi alla stessa indagine**

Sono dati acquisiti in tornate precedenti della stessa indagine: in qualche caso può essere utile precompilare il questionario corrente con dati forniti in tornate precedenti per facilitare l'attività di compilazione al rispondente anche attraverso confronto con dati già dichiarati in passato

#### **Dati di amministrazione e monitoraggio della rilevazione**

Sono dati utili alla gestione del processo e al monitoraggio dell'andamento della raccolta: tipicamente comprendono dati sul comportamento dei rispondenti – numero di utenti registrati, numero di prelievi/invii di questionario, ultimo accesso di un rispondente, tasso percentuale di risposta elettronica

### **Dati di amministrazione e monitoraggio del sistema di acquisizione**

Sono dati utili alla gestione complessiva del nuovo sistema di acquisizione: tipicamente comprendono dati sulle applicazioni sviluppate e in produzione; modalità di sviluppo adottate, risorse fisiche coinvolte, interfacce di comunicazione presso i settori statistici, classe di rispondenti interessata

Le classi di dati descritte, in ragione dell'informazione che rappresentano, hanno gradi di riservatezza diversi e richiedono diverse misure di protezione; va evidenziato a questo riguardo che l'architettura hardware e software del sistema *indata.istat.it* è stata progettata per garantire la sicurezza dei dati e agevolare la loro veicolazione verso i settori di produzione interna.

### **3.3 Specifiche architetturali**

Nel progettare l'architettura hardware e software del nuovo sistema di acquisizione si sono tenuti in particolare considerazione i requisiti e vincoli descritti precedentemente.

L'architettura è di tipo *three layer*, in cui si distinguono il livello che si occupa della logica di presentazione dell'applicazione e di terminazione di tutte le comunicazioni con i client (Front End), il livello che si occupa della logica applicativa (Application Server) ed infine il livello che si occupa della logica di accesso ai dati (DB Server). La configurazione complessiva si basa su un'architettura LAMP (Linux, Apache, MySQL, PHP (*Hypertext PreProcessor*)) e prevede un unico punto di accesso visibile dall'esterno e localizzato su un cluster di server esposti sulla rete Internet, che possono alternarsi l'uno con l'altro, rendendo eventuali malfunzionamenti del front-end non rilevabili dai rispondenti.

Tutte le richieste provenienti dai rispondenti che accedono al front-end vengono inoltrate sul server applicativo e sul DataBase Oracle, che sono collocati in una zona a protezione più elevata del sistema *firewall*. Quest'ultimo, in rispetto dei vincoli di sicurezza, è configurato in maniera da permettere il transito delle sole comunicazioni strettamente necessarie all'espletamento del servizio di acquisizione.

Tutte le trasmissioni dai rispondenti al sistema di acquisizione sono cifrate e autenticate da un certificato digitale emesso da un'Autorità di certificazione riconosciuta a livello internazionale.

Il sito *indata.istat.it* costituisce il punto di accesso unico per tutti gli utenti, esterni (rispondenti) ed interni (gestori delle indagini) coinvolti nel processo di *data capturing*. L'organizzazione del sito prevede un'area pubblica, porzione accessibile a tutti quelli che ne conoscono l'indirizzo senza fornitura di credenziali, e un'area privata, a cui accedono solo utenti abilitati, in possesso di credenziali valide, dopo avere effettuato un passo di autenticazione.

L'area pubblica è costituita dalla homepage del sito e dalle pagine di accoglienza delle indagini, accessibili attraverso i link ipertestuali ivi esposti.

La homepage, oltre all'elenco delle indagini trattate, ordinate per dominio statistico di appartenenza, contiene alcuni link a pagine che descrivono la finalità del progetto di cattura dati, danno indicazione su alcuni contatti che potrebbero essere utili ai rispondenti, presentano il sito istituzionale dell'Istituto. Le pagine di accoglienza delle indagini, che fanno parte dall'area pubblica, offrono ai rispondenti una descrizione delle finalità della rilevazione, la possibilità di stampare o scaricare guide, istruzioni e materiale vario di riferimento per l'indagine, numeri di telefono e indirizzi e-mail utili presso il Servizio di produzione che amministra la raccolta.

Fanno parte dell'area privata del sito le pagine attraverso cui i rispondenti scaricano, compilano o inviano il questionario elettronico e le pagine riservate agli statistici amministratori della rilevazione, utenti interni, che possono intervenire sugli archivi anagrafici dei rispondenti e monitorare le operazioni di raccolta.

### ***3.4 Linguaggi, formati e strumenti di sviluppo***

Il sito ufficiale di acquisizione dati, *indata.istat.it*, è stato sviluppato in architettura LAMP e, pertanto, PHP costituisce il linguaggio preferenziale per lo sviluppo di nuove soluzioni di cattura dati da integrare nel sito stesso. In alternativa può essere utilizzato l'ambiente Java. In entrambi i casi devono essere adoperate versioni del software uguali o compatibili con quelle già presenti sul sito *indata.istat.it*: PHP 5.1.2 e Java 1.4.1.

Il codice sviluppato deve, inoltre, essere compatibile con le versioni software usate per il Web Server, in particolare Apache 2.0.52 e Tomcat 5.0.28.

Per la specifica dello stile delle pagine sono utilizzati i fogli CSS2.

Il questionario elettronico costituisce l'oggetto base su cui impiantare l'applicazione di cattura dati: a meno delle situazioni in cui è previsto un semplice trasferimento di dataset, va sviluppato un questionario elettronico che sarà generalmente autocompilato dal rispondente in modalità online oppure offline.

Sul sito *indata.istat.it* per i questionari elettronici i formati attualmente utilizzati sono PDF, HTML e EXCEL, che risultano familiari agli utenti e facilmente trattabili dai più diffusi browser, sia di mercato sia open source.

I questionari in formato PDF vengono generati con il modulo Designer del software Teleform.

Il software Teleform è un sistema di mercato per la generazione e il trattamento di form cartacei ed elettronici, acquisito dall'Istituto nel 2000; è composto di tre moduli principali: Designer, per la generazione dei forms; Reader per l'acquisizione automatica da più canali, e Verifier, per la verifica da operatore dei questionari acquisiti.

I questionari PDF elettronici vengono arricchiti di codice Adobe/javascript per attività di controllo contestuale alla compilazione; il formato PDF ottenuto con il modulo Designer del software Teleform, se stampato, compilato manualmente, e restituito, può essere sottoposto a lettura ottica attraverso il modulo Reader, dotato di motori OCR/ICR (*Optical Character Recognition/ Intelligent Character Recognition*).

I questionari elettronici in formato HTML possono essere generati con il modulo Designer del software Teleform oppure possono essere creati con sviluppo ad hoc.

I questionari elettronici in formato EXCEL vengono scaricati, in forma personalizzata, dai rispondenti, dopo accesso a *indata.istat.it* e passo di autenticazione, e vengono restituiti con un'operazione di upload su una specifica area del sito.

### ***3.5 L'applicativo DATISTAT per l'acquisizione secondaria dei dati***

La produzione statistica ufficiale, nell'ottica della cooperazione tra Amministrazioni pubbliche, del riuso dei patrimoni informativi già disponibili, e del contenimento del disturbo statistico verso i rispondenti, si pone sempre più l'obiettivo di utilizzare fonti esterne di dati, anche di carattere amministrativo, e ridurre, quando è possibile, la costosa attività di rilevazione diretta.

E' prevedibile pertanto un incremento delle attività di acquisizione secondaria dei dati, soprattutto per lo scambio di fonti amministrative e dati raccolti a diverso titolo all'esterno dell'Istituto.

Le operazioni di acquisizione secondaria vengono risolte con l'applicativo DATISTAT, sviluppato internamente proprio con questa finalità.

Datistat è un'applicazione che realizza trasferimento sicuro di file; è scritta in linguaggio PHP e disponibile sulla home page del sito unificato *indata.istat.it*. Datistat fu attivata nel 2004 per offrire un canale sicuro alle operazioni di trasferimento che le società incaricate di indagini CATI dovevano periodicamente eseguire verso l'Istituto. I dati, a volte sensibili, erano usualmente trasmessi tramite

posta elettronica in violazione delle norme. Nel tempo la procedura si è arricchita di nuove funzioni, tra le quali la più importante ha consentito agli utenti esterni di prelevare dataset dallo stesso sito *indata.istat.it*; in questo modo il sito, oltre che l'acquisizione, è in grado di trattare lo scambio dei dati.

Con Datistat il trasferimento fisico dei dataset avviene con operazioni di UPLOAD/DOWNLOAD guidate per l'utente da una cornice applicativa che le rende di facile utilizzo. Le operazioni si svolgono in protocollo HTTPS garantendo così la protezione dei dati nel transito dal client dell'utente al server di Istituto. Datistat ha impatto minimo sulle dotazioni informatiche degli utenti esterni e mantiene un buon livello di autonomia degli utenti interni destinatari della trasmissione dei dataset. Si possono trasferire dataset di qualunque formato; i vincoli sull'ampiezza del dataset sono principalmente dipendenti dalle condizioni contrattuali praticate all'utente esterno dall'Internet provider che gli fornisce il collegamento alla rete Internet, e in questo senso poco controllabili. A livello di configurazione APACHE su *indata.istat.it* è stato definito un parametro massimo di 100 MB per l'ampiezza dei dataset da movimentare. Attualmente l'ampiezza massima effettivamente trattata è stata di 50 MB. Quando i dati da trasferire hanno dimensioni maggiori sono disponibili soluzioni alternative quali, ad esempio, quelle basate sul File Transfer Protocol e disponibili su server diversi da *indata*.

Le funzioni di Datistat, come tutte le altre funzioni di cattura dati disponibili su *indata.istat.it*, sono accessibili con garanzie reciproche, a chi trasmette e a chi riceve i dati, dell'identità del sito (che è regolarmente certificato) e degli utenti che operano sul sito (attraverso la gestione di USER-ID e PASSWORD secondo le norme vigenti).

Ad oggi attraverso Datistat sono stati trasmessi all'Istituto oltre 3000 dataset.

L'ampliamento delle funzioni di Datistat, in particolare il rilascio della funzione di scambio dati, ha reso il sito *indata.istat.it* oggetto di interesse anche per la Direzione Centrale della Diffusione: gli utenti esterni, le cui richieste siano state preventivamente analizzate e autorizzate, si connettono direttamente al sito e prelevano le raccolte di dati oggetto delle loro richieste; si realizza in questo modo una riduzione dei tempi di evasione delle richieste, oltre che delle spese di preparazione e spedizione di supporti (CD, DVD).

## **4. ELEMENTI DI MONITORAGGIO**

### ***4.1 Lo sviluppo delle applicazioni di cattura dati***

Dopo qualche anno di sperimentazione, riflessione e assestamento, a partire dal 2005 il parco delle applicazioni di cattura dati si è sviluppato velocemente, sia relativamente al numero delle rilevazioni trattate, sia per quanto concerne le funzioni offerte agli utenti. L'inizio del periodo fertile per lo sviluppo del data capturing ha coinciso non solo con la maturazione naturale delle competenze tecniche ma anche con alcune misure organizzative che hanno assegnato con chiarezza mandato e obiettivi per la cattura dati.

Ad oggi sul sito *indata.istat.it* sono state rilasciate 38 applicazioni di cattura dati, di cui 4 non sono più attive, perché terminate o perché di natura occasionale (l'elenco delle indagini è riportato in chiusura a questo documento); inoltre 9 nuove applicazioni di cattura dati sono in fase di sviluppo e per numerose altre rilevazioni i settori di produzione statistica hanno già manifestato concreto interesse all'innovazione.

Nell'innovazione della fase di raccolta dei dati quasi tutte le Direzioni di produzione statistica sono rappresentate; sono state inoltre forniti servizi anche a Direzioni centrali di carattere trasversale, quali la Direzione Centrale per l'integrazione e il territorio e la Direzione Centrale per la diffusione dell'informazione statistica, come mostra la tabella 1.

Delle 38 applicazioni già trattate su *indata*, 21 sono indirizzate a imprese, 16 a istituzioni, 1 a individui. L'applicazione che interessa la classe di rispondenti *individui* è una delle più recenti; realizza una ricognizione delle esigenze informative per la costruzione di un sistema informativo sul lavoro ed è indirizzata ad esperti nello specifico settore.

I domini statistici già trattati sul sito sono: Industria e Servizi (il più numeroso in termini di rilevazioni), Agricoltura e Zootecnia, Pubblica Amministrazione e Istituzioni private, Formazione e società, Istruzione, Censimenti, Ambiente e territorio, Prezzi.

L'oggetto "questionario elettronico" viene prodotto secondo due modalità prevalenti: con il software generalizzato commerciale Teleform (in formato PDF o HTML) o come foglio elettronico; parallelamente sono state realizzate applicazioni ad hoc che risultano più flessibili rispetto alla possibilità di sviluppare questionari dinamici – di struttura e dimensione variabile – ma più onerose in termini di risorse da dedicare allo sviluppo. Ci sono inoltre casi in cui per l'acquisizione dei dati non è prevista la compilazione di un questionario bensì l'inoltro di un dataset. Maggiori dettagli sono riportati nella Tabella 3.

**Tabella 1:** *N. di applicazioni per Direzione*

Direzione	Attive	Terminate	Totale
DCCG Censimenti Generali	0	3	3
DCET Esigenze informative, integrazione e territorio	1	0	1
DCIS Statistiche e indagini su istituzioni sociali	1	0	1
DCPC Statistiche sui prezzi e il commercio estero	1	0	1
DCSC Statistiche economiche su imprese, servizi e occupazioni	10	0	10
DCSS Statistiche economiche strutturali	19	1	20
DPTS-AMB Statistiche ambientali e sviluppo sostenibile	2	0	2
<b>Totale</b>	34	4	38

**Tabella 2:** *N. di applicazioni per periodicità della rilevazione*

Periodicità	Attive	Terminate	Totale
Mensile	6	0	6
Trimestrale	2	0	2
Semestrale	1	0	1
Annuale	16	1	17
Biennale	3	0	3
Triennale	2	0	2
Quadriennale	2	0	2
Infrannuale	1	0	1
Occasionale	1	3	4
<b>Totale</b>	34	4	38

**Tabella 3:** *N. di applicazioni per tipologia di questionario elettronico*

<b>Tipologia Questionario elettronico</b>	<b>Attive</b>	<b>Terminate</b>	<b>Totale</b>
Teleform	15	3	<b>18</b>
EXCEL	13	1	<b>14</b>
Upload dataset	4	0	<b>4</b>
PHP	1	0	<b>1</b>
Blaise	1	0	<b>1</b>
<b>Totale</b>	<b>34</b>	<b>4</b>	<b>38</b>

**Tabella 4:** *N. di applicazioni per classe di rispondenti*

<b>Classe rispondenti</b>	<b>Attive</b>	<b>Terminate</b>	<b>Totale</b>
Imprese	20	1	<b>21</b>
Istituzioni	13	3	<b>16</b>
Individui	1	0	<b>1</b>
<b>Totale</b>	<b>34</b>	<b>4</b>	<b>38</b>

#### ***4.2 Tasso di risposta elettronica***

Tutte le rilevazioni statistiche dirette, che siano svolte in modalità tradizionale o innovativa, sperimentano il problema del rifiuto a rispondere. Il fenomeno si spiega con una scarsa motivazione dei rispondenti che spesso si sentono estranei al processo di produzione dell'informazione statistica e d'altra parte non riscontrano vantaggi diretti dalla loro partecipazione alla rilevazione, ma anzi devono dedicare tempo e risorse alla compilazione di modelli talvolta complessi, sottraendoli alle loro attività ordinarie.

Inoltre per alcune classi di rispondenti – ci riferiamo principalmente alle imprese – i problemi sono acuiti dalla caratteristica che una stessa unità partecipa a volte ai campioni estratti di numerose indagini, andando così a moltiplicare l'impegno necessario per rispondere a domande che in certi casi si ripetono identiche nei diversi questionari, tanto da indurre le imprese stesse a chiedere di essere esonerate dalla rilevazione.

Ulteriore fattore critico per la classe delle imprese è il grado di correttezza e completezza dell'archivio di base delle unità statistiche che spesso non è aggiornato e non riporta correttamente informazioni che le imprese hanno già comunicato all'Istituto, per esempio in occasione di un precedente contatto determinatosi per una diversa indagine.

In generale dunque il tasso di mancata risposta totale costituisce un fattore di criticità a cui gli amministratori delle indagini fanno fronte sovradimensionando il campione al momento dell'estrazione per essere certi che al termine del processo le stime prodotte possano comunque essere significative per la popolazione di riferimento.

Esiste parallelamente un problema di mancata risposta parziale, ossia i questionari non vengono compilati totalmente perché ad alcune domande non viene data risposta; anche questo è un fattore critico, più o meno grave a seconda dell'importanza della variabile interessata dalla mancata risposta: in

certi casi i questionari, se non presentano valorizzati alcuni specifici campi di input, sono da considerarsi non validi.

L'adozione di tecniche di data capturing dà un contributo alla qualità dei dati acquisiti ma non risolve il problema della mancata risposta totale: in altri termini chi non rispondeva con le modalità tradizionali non è detto cambi atteggiamento con l'offerta della risposta elettronica. Inoltre è stato verificato che chi risponde in modalità elettronica risponde prima, incrementando così la tempestività nella disponibilità dei dati da sottoporre a elaborazione. Maggiore è l'influenza sulla mancata risposta parziale: avendo a disposizione un questionario elettronico è possibile controllare in fase di input la presenza di risposta in alcuni campi considerati irrinunciabili e condizionare l'avanzamento dell'autointervista alla fornitura di queste risposte.

Il tasso di risposta elettronica, definito per una tornata di raccolta, come rapporto tra numero di questionari elettronici ricevuti e numero totale di questionari validi ricevuti secondo qualunque modalità (generalmente la modalità alternativa di risposta è il questionario cartaceo), dà una misura del contributo della tecnica innovativa alle operazioni di raccolta. In qualche caso, per la classe di rispondenti *istituzioni pubbliche*, o per speciali classi di imprese (per esempio le imprese del settore telecomunicazioni) la modalità di risposta elettronica è stata l'unica possibilità offerta e ha dato luogo a un tasso di risposta piuttosto elevato.

L'esperienza maturata in questi anni ci consente alcune considerazioni di carattere generale.

I rispondenti reagiscono con interesse alle proposte di innovazione: a fronte di sporadiche lamentele per la impossibilità di utilizzare il questionario elettronico (condizione generata dall'obsolescenza dell'ambiente client o dalle politiche di protezione degli internet provider) riceviamo numerosi commenti di gradimento delle nuove modalità.

I tassi di risposta elettronica non sono noti per tutte le applicazioni di cattura dati avviate, in quanto non tutti i settori di produzione ci forniscono dati sulle risposte in modalità tradizionale (parametro necessario per calcolare l'indicatore). Tuttavia va evidenziato che, nell'ambito di una stessa rilevazione di carattere periodico, il numero di questionari ricevuti su *indata* per le diverse tornate è sempre crescente nel tempo. Questo indica sia una fidelizzazione dei rispondenti sia una maturazione progressiva dell'ambiente, che può riguardare il lato *client* (i rispondenti si attrezzano tecnologicamente e culturalmente) ma anche il lato *server* (il nostro sito e le nostre applicazioni di cattura dati diventano più gradevoli e più efficaci).

La classe di rispondenti *istituzioni* mediamente risponde di più all'innovazione rispetto alle *imprese*: soprattutto rispondenti della pubblica amministrazione locale generalmente mostrano maggiore disponibilità.

Complessivamente sono stati acquisiti attraverso il sito *indata* più di 250.000 questionari a partire dall'anno 2004.

A titolo di esempio si riportano i tassi di risposta di alcune rilevazioni. La rilevazione annuale della produzione annuale, Prodcum, è stata trattata a partire dall'anno di riferimento 2004: per le tornate completate (2004, 2005, 2006) i tassi di risposta elettronica sono stati rispettivamente 23%, 44,3%, 57,6%. La rilevazione annuale Sistema Conto Imprese, SCI, è stata trattata a partire dall'anno di riferimento 2003; i tassi conseguiti sono: 36%, 60%, 69%, 92%. Per la rilevazione del Fatturato trimestrale di alcuni servizi, il cui trattamento è iniziato con il I trimestre 2006 si è passati da 1589 questionari iniziali a 2761 nel II trimestre 2008.

La rilevazione annuale sulle imprese di telecomunicazione, di carattere totale, attiva dall'anno di riferimento 2004, si è sempre attestata sul 95%, come è giusto aspettarsi dalla classe di rispondenti interessata. Non mancano esempi meno brillanti: la rilevazione sull'attività alberghiera (Natale, Pasqua e Ferragosto), attiva da 4 anni, dal 5% iniziale ha raggiunto l'11,5% a Natale 2007.

### ***4.3 Monitoraggio delle azioni dei rispondenti***

Per assicurare il regolare svolgimento delle rilevazioni ospitate sul sito *indata* non si può prescindere da un costante monitoraggio delle stesse e, in particolar modo, dal controllo, con cadenza almeno quotidiana, di specifiche azioni dei rispondenti che assicurano lo stato di *buona salute* della rilevazione. Tale controllo consente, da una parte, di avere immediatamente conoscenza di situazioni anomale (malfunzionamenti hw o sw) e, dall'altra, di misurare l'atteggiamento dei rispondenti e valutare l'andamento della raccolta.

Le principali azioni sottoposte a monitoraggio sono le registrazioni dei rispondenti e gli invii eseguiti.

La registrazione è l'operazione iniziale, obbligatoria ed una tantum, che ciascun rispondente è tenuto a fare, nello spazio web dell'indagine che gli compete, per potere successivamente compilare il questionario e inviare i dati. In sostanza si tratta di collegarsi all'indirizzo URL (*Uniform Resource Locator*) dell'indagine, autenticarsi con la userid e la password iniziale, comunicate nella lettera di invito, e sostituire la password iniziale con una personale liberamente scelta. Con la registrazione il rispondente si abilita ad accedere al questionario per la compilazione. Qualsiasi accesso successivo richiede l'autenticazione del rispondente con userid e password personale. Procedure apposite consentono al rispondente di modificare in qualunque momento la password personale o, in caso di smarrimento, di reimpostarla al valore iniziale.

Nelle rilevazioni annuali, o con periodicità più ampia dell'annuale, le registrazioni, anche per i rispondenti che hanno partecipato alla tornata precedente, devono essere nuovamente eseguite perché le password si considerano scadute. Nelle indagini con cadenza diversa, ove gli stessi rispondenti partecipano a più tornate, la registrazione rimane valida fin quando l'impresa o l'ente rimane nell'indagine.

Quando una rilevazione adotta un questionario online, al termine della compilazione i dati immessi dal rispondente sono inoltrati mediante il pulsante *Invia*, sempre presente nel questionario, e memorizzati in tabelle Oracle. I questionari in forma di foglio elettronico, o i dataset, vengono inviati mediante una funzione di *Upload*, presente nello spazio web dell'indagine. Per valutare correttamente le cifre riportate nella tabella che segue, è opportuno precisare che in alcune indagini, in accordo con le indicazioni dei Servizi committenti, è consentito l'invio anche di questionari parzialmente compilati. Tale possibilità si associa, in genere, con modelli che richiedono una compilazione particolarmente onerosa o l'intervento di competenze diverse presenti nei vari uffici delle imprese. Sono altresì consentiti, tranne rare eccezioni, successivi invii per la stessa tornata di indagine quando il rispondente ritiene di aver fornito dati errati o incompleti. Va evidenziato che gli invii multipli sono comunque da considerarsi quantitativamente marginali rispetto al numero totale di invii.

Il numero delle registrazioni è controllato utilizzando le funzioni di gestione delle singole indagini. Per quanto riguarda gli invii, il monitoraggio è eseguito accedendo alle tabelle Oracle di output o, nel caso di invio di dataset o di fogli Excel, con le stesse procedure di gestione.

Nella tabella che segue, per ogni indagine, sono riportati i valori complessivi delle registrazioni e degli invii per tutte le edizioni fino ad ora attivate sul sito. La colonna "Inizio monitoraggio" specifica l'anno, coincidente nella maggior parte dei casi con l'anno di prima esecuzione telematica dell'indagine, a partire dal quale sono calcolati i dati di monitoraggio.

**Tabella 5:** *N. di registrazioni e di invii di questionario per indagine*

<b>Rilevazione</b>	<b>Direzione</b>	<b>Periodicità</b>	<b>Tipo quest.</b>	<b>Rispondenti</b>	<b>Inizio monitor.</b>	<b>Numero registr.</b>	<b>Numero invii</b>
Dati ambientali nelle città	DPTS AMB	Annuale	PDF	Capoluoghi di provincia	2008	109	4.406
Rilevazione sui servizi idrici (Grandi gestori)	DPTS AMB	Annuale	Dataset	Gestori servizi idrici	2008	204	168
Rilevazione sugli archivi dei numeri civici	DCCG	Occasionale	PDF	Comuni	2007	867	2.415
Rilevazione sull'uso dei dati censuari	DCCG	Occasionale	PDF	Organi Sistan	2007	982	1.457
Rilevazioni dei costi censuari dei comuni	DCCG	Occasionale	PDF	Comuni	2007	249	688
Indagine preliminare sistema informativo integrato sul mercato del lavoro	DCET	Occasionale	PDF	Esperti mercato del lavoro	2008	28	17
Indagine triennale sui laureati	DCIS	Triennale	Excel	Università	2006	89	82
Prezzi del settore agricolo	DCPC	Mensile	Dataset	Camere di commercio	2008	75	600
Indagine mensile sul fatturato e ordinativi	DCSC	Mensile	Excel	Imprese	2007	1.824	12.127
Indagine mensile sulla produzione industriale	DCSC	Mensile	HTML	Imprese industriali	2007	1.382	47.240
Indagine sulla capacità degli esercizi ricettivi	DCSC	Annuale	Excel	Regioni, Province, Az. di soggiorno	2008	43	194
Rilevazione annuale del trasporto ferroviario	DCSC	Annuale	PDF	Aziende di trasporto	2005	50	55
Rilevazione mensile movimento clienti strutture ricettive	DCSC	Mensile	Dataset	Regioni, Province, Az. di soggiorno	2007	45	3.000
Rilevazione mensile su occupazione, orari di lavoro e retribuzioni	DCSC	Mensile	PDF	Imprese	2004	416	15.219
Rilevazione mensile sulle vendite al dettaglio	DCSC	Mensile	PDF	Imprese di vendita al dettaglio	2004	830	9.295
Rilevazione periodica sull'attività alberghiera	DCSC	Infrannuale	PDF	Alberghi	2004	1.597	1.307
Rilevazione trimestrale del fatturato e ordinativi	DCSC	Trimestrale	PDF	Imprese	2006	4.779	23.420
Rilevazione trimestrale sui posti vacanti e ore lavorate	DCSC	Trimestrale	PDF	Imprese	2004	5.829	30.445

Rilevazione	Direzione	Periodicità	Tipo quest.	Rispondenti	Inizio monitor.	Numero registr.	Numero invii
International sourcing - Dinamiche e modalità di internazionalizzazione e delle imprese italiane	DCSS	Triennale	PDF	Imprese	2007	3.625	2.372
Rilevazione annuale del sistema dei conti delle imprese	DCSS	Annuale	Excel	Imprese	2004	23.213	15.166
Rilevazione annuale della produzione industriale	DCSS	Annuale	Excel	Imprese	2005	57.137	38.810
Rilevazione annuale delle tecnologie di informazione della Pubblica Amministrazione	DCSS	Biennale	PDF	Pubblica amministrazione	2007	5.124	5.447
Rilevazione annuale delle tecnologie di informazione delle società non finanziarie	DCSS	Annuale	PDF	Imprese non finanziarie	2006	17.180	13.549
Rilevazione annuale delle tecnologie di informazione delle società finanziarie	DCSS	Annuale	PDF	Imprese finanziarie	2006	2.133	1.943
Rilevazione annuale stima provvisoria del valore aggiunto	DCSS	Annuale	Excel	Imprese	2005	4.087	3.464
Rilevazione annuale sui risultati economici delle aziende agricole	DCSS	Annuale	Excel	Aziende agricole	2006	dati non disponibili	
Rilevazione annuale sulla pesca nei laghi e bacini artificiali	DCSS	Annuale	PDF	Camere di commercio	2006	58	931
Rilevazione annuale sulla struttura del costo del lavoro	DCSS	Quadriennale	Excel	Imprese	2005	3.620	1.547
Rilevazione annuale sulla struttura e produzione delle aziende agricole	DCSS	Biennale	Excel	Aziende agricole	2005	21	161
Rilevazione annuale sulle telecomunicazioni	DCSS	Annuale	Excel	Imprese di telecomunicaz	2005	780	837
Rilevazione dei bilanci consuntivi degli enti pubblici	DCSS	Annuale	Excel	Istituzioni pubbliche	2008	495	446
Rilevazione semestrale estimativa della consistenza del bestiame	DCSS	Semestrale	PDF	Camere di commercio	2005	101	441
Rilevazione statistica sulle imprese a controllo estero	DCSS	Biennale	PDF	Imprese a controllo estero	2007	3.332	2.270

Rilevazione	Direzione	Periodicità	Tipo quest.	Rispondenti	Inizio monitor.	Numero registr.	Numero invii
Rilevazione su ricerca e sviluppo negli enti e nelle istituzioni pubbliche	DCSS	Annuale	Excel	Istituzioni pubbliche	2006	1.125	820
Rilevazione su ricerca e sviluppo nelle istituzioni private no profit	DCSS	Annuale	Excel	Istituzioni private	2006	1.183	813
Rilevazione sulla caccia	DCSS	Annuale	Excel	Province	2008	41	35
Rilevazione sulle piccole e medie imprese	DCSS	Annuale	Excel	Imprese, artigiani e professionisti	2008	7.970	6.380
Struttura retribuzioni 2006	DCSS	Quadriennale	Excel	Imprese	2007	7.390	4.342

#### ***4.4 Applicazione del software Webalizer***

Per monitorare l'utilizzo del server *indata* e la sua evoluzione nel tempo viene utilizzato un software freeware, distribuito con licenza GNU (General Public License), largamente conosciuto: Webalizer.

Il prodotto elabora in input i file di log del server Web, prodotti giornalmente dal software Apache e memorizzati sul server stesso, e genera, in output, una serie di tabelle e grafici in formato HTML che possono essere visualizzati con un comune browser.

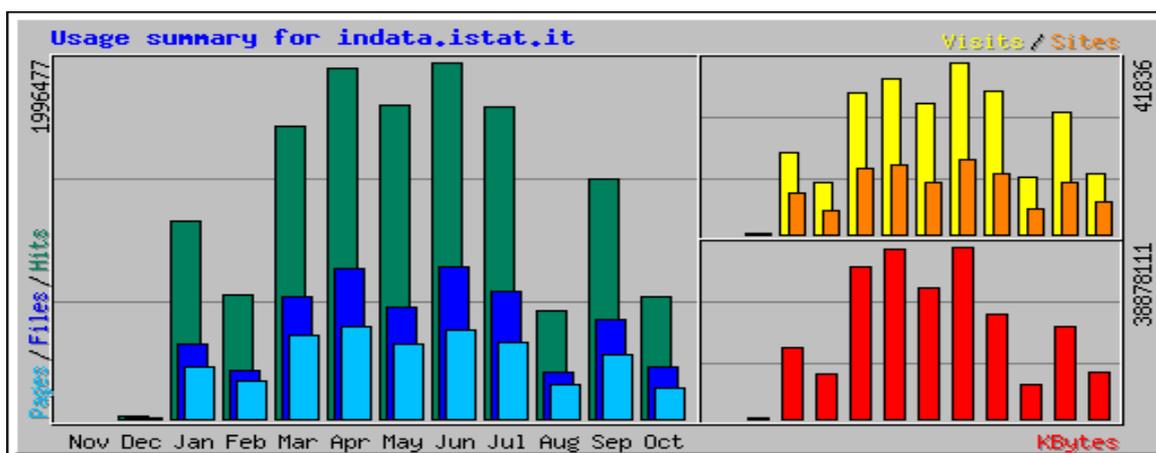
Un file di configurazione, *webalizer.conf*, consente di personalizzare, in parte, gli output prodotti e modificare i valori predefiniti del programma.

Il documento HTML, prodotto da Webalizer, fornisce numerose informazioni, di non sempre facile interpretazione, sul numero degli accessi al sito, la loro distribuzione nel tempo, le pagine visitate, la dimensione in Kbyte delle informazioni trasmesse dal server ai client dei visitatori, il browser utilizzato per accedere al sito ed altre.

La prima pagina del documento presenta un'informazione sintetica relativa agli ultimi dodici mesi dell'intero periodo di osservazione, in forma grafica e tabellare. Nel nostro caso i dati analizzati dal prodotto iniziano con gli ultimi giorni di dicembre 2007 (questo è il motivo del basso valore relativo al mese) sino alla metà di ottobre 2008.

Di seguito è riportato il grafico relativo al sito *indata*:

**Grafico 1:** Utilizzo del sito indata nel 2008



e la corrispondente tabella, con la media giornaliera ed i dati complessivi di ciascun mese:

**Tabella 6:** Utilizzo del sito indata nel 2008

Summary by Month										
Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
<a href="#">Oct 2008</a>	56559	24401	14666	1240	7911	10683086	14883	175996	292817	678713
<a href="#">Sep 2008</a>	44836	18659	11887	984	12619	21019879	29536	356620	559775	1345096
<a href="#">Aug 2008</a>	19515	8461	6234	452	6049	7577656	14013	193278	262300	604979
<a href="#">Jul 2008</a>	56205	23147	13867	1117	14792	23695109	34652	429891	717563	1742361
<a href="#">Jun 2008</a>	66549	28231	16566	1394	18166	38878111	41836	496994	846939	1996477
<a href="#">May 2008</a>	56695	20142	13504	1028	12385	29465728	31869	418635	624417	1757575
<a href="#">Apr 2008</a>	65343	27933	17178	1263	16847	38188263	37903	515346	837995	1960304
<a href="#">Mar 2008</a>	52876	22102	15174	1103	15859	34158576	34195	470424	685167	1639169
<a href="#">Feb 2008</a>	23843	9177	7245	432	5566	10007059	12540	210120	266147	691449
<a href="#">Jan 2008</a>	35542	13626	9326	645	10111	15964597	19996	289119	422431	1101813
<a href="#">Dec 2007</a>	6624	2003	3610	115	118	70366	231	7220	4007	13248
<b>Totals</b>						<b>229708430</b>	<b>271654</b>	<b>3563643</b>	<b>5519558</b>	<b>13531184</b>

Prima di commentare i dati, riportiamo una sintetica descrizione dei termini che appaiono:

**HITS** è il numero totale dei file richiesti al server *indata* per il singolo mese. È da tener presente, per una corretta valutazione del dato, che una singola pagina visualizzata dal browser del visitatore può essere composta da più file. Se la home-page di un'indagine contiene due immagini, il numero degli HITS conteggiati sarà incrementato di 3: il file di testo *index.html* più due file immagini.

**FILES** è il numero totale di file richiesti ed effettivamente inviati al visitatore. Nel caso i file siano già presenti nella memoria cache di sistema del visitatore non vengono inviati dal server; si avrà, quindi, un incremento solo del valore degli HITS, mentre il numero dei FILES rimarrà invariato. La differenza tra HITS e FILES fornisce una indicazione sommaria dei visitatori abituali; maggiore è questa differenza maggiore è il numero di utenti che richiede pagine che sono già nella cache dei loro browser e quindi hanno già visitato almeno una volta la pagina.

**PAGES** è il numero delle pagine visualizzate dai visitatori. Il prodotto considera PAGES i file di formato: Htm, Html, Cgi, Php3, Php4 e altri formati minori. Sono esclusi eventuali files immagine o video contenuti nelle pagine.

**VISITS** è il numero delle visite ricevute dal server. Il valore viene incrementato di uno ogni qualvolta un utente richiede, per la prima volta in un periodo di tempo prefissato (predefinito in 30 minuti), una pagina del sito. Le pagine successive, visualizzate dallo stesso utente all'interno del medesimo periodo, non sono considerate visite ulteriori. Qualora una pagina del sito venga richiesta dallo stesso utente, trascorsi i 30 minuti (o un diverso periodo di tempo impostato nel file di configurazione) dall'ultima richiesta, viene conteggiata una nuova visita. Il normale aggiornamento della pagina non produce, quindi, alcun incremento delle visite. Il parametro *tempo di visita* per il sito *indata* è stato impostato, nel file di configurazione, a 15 minuti, ritenendosi questo un valore più vicino alla reale durata media delle visite dei rispondenti.

**SITES** è il numero degli indirizzi IP diversi che effettuano una o più richieste al server. Questo dato va preso con cautela e richiede qualche precisazione: gli utenti Internet hanno in buona parte indirizzi IP dinamici. Questo significa che un medesimo utente che in tempi diversi, con IP mutato, all'interno del periodo di riferimento (mese, giorno, ecc.), visita il nostro sito può essere conteggiato più volte. Vale anche il contrario: utenti diversi, appartenenti ad una grande azienda o a una medesima istituzione, è il caso ad esempio degli utenti Istat, si presentano in Internet con uno stesso IP uguale per tutti e sarebbero, quindi, conteggiati una sola volta. Parliamo della tecnica NAT - Network Address Translation - che, per superare i limiti degli indirizzi IP disponibili, consente a centinaia o migliaia di computer su una rete privata, di condividere, tramite un firewall o un router, un solo indirizzo IP pubblico.

**KBYTES** rappresenta il numero totale di KBytes trasferiti dal server al computer del visitatore. Ad esempio, se una pagina acceduta è composta da due immagini di 10 e 20 Kb ed il file html della stessa pagina è di 5 Kb, il valore della voce subisce un incremento di 35 Kb.

Chiarito il significato dei termini, vanno evidenziate alcune caratteristiche del sito *indata*, che dipendono dalle proprietà stesse della funzione di raccolta dell'informazione statistica, e che condizionano la lettura dei risultati prodotti da Webalizer.

I valori che si riferiscono al server *indata* dipendono naturalmente dalla tipologia delle rilevazioni Istat e massimamente dal requisito della periodicità: trimestrale, semestrale, annuale, ecc. Ci sono poi anche indagini occasionali (ad esempio quelle realizzate in preparazione del Censimento 2011) che possono iniziare in un mese qualsiasi dell'anno. Anche il numero dei rispondenti è molto variabile: passiamo da un minimo - 66 rispondenti - nel caso dell'Indagine preliminare sul sistema informativo del mercato del lavoro, ad un massimo - 103.040 - per la rilevazione sulle piccole e medie imprese. Variabile è anche il periodo nel quale le indagini rimangono accessibili per l'invio dei questionari. Anche il numero giornaliero dei questionari inviati ha un andamento tipico: sostenuto all'inizio dell'indagine, decrescente poi e, infine, con picchi corrispondenti ai solleciti inviati dal servizio.

L'andamento della raccolta è poi influenzato, ovviamente, dall'orario di lavoro, dai cicli settimanali, e dai periodi di attività dei rispondenti che, ricordiamo, sono, per le nostre indagini, in massima parte imprese private ed istituzioni.

In altri termini, alcune variazioni nell'arco dell'anno, che a prima vista potrebbero essere interpretate come disaffezione da parte degli utenti, sono invece fisiologiche e temporanee e rispecchiano

L'andamento delle attività delle indagini, compresi picchi e riduzioni di attività. Potrebbe avere maggior senso confrontare i valori dell'utilizzo del sito in periodi corrispondenti per anni diversi ma va tenuto conto che il numero delle rilevazioni non è costante, poiché vengono rilasciate continuamente nuove applicazioni, e che lo svolgimento di una stessa rilevazione non riesce a rispettare sempre la relativa pianificazione.

Ciò premesso, già nel grafico n.1 è facilmente osservabile il calo marcato nel mese di agosto, caratterizzato dalle tipiche ferie estive e la bassa media giornaliera di dicembre (per questo mese sono presenti nei log di input solo gli ultimi giorni dell'anno) coincidente con la pausa di fine anno. Sono invece elevati i valori di aprile (inizio della raccolta per il periodo gennaio – marzo delle rilevazioni trimestrali) e giugno (inizio della Rilevazione sulle piccole e medie imprese: indagine con un numero molto elevato di rispondenti).

Nella pagina iniziale del documento HTML di Webalizer, selezionando uno dei mesi visualizzati nella prima colonna a destra della tabella, si accede alle informazioni di dettaglio relative a quel periodo. Tra le varie tabelle e grafici disponibili, a livello mensile, ad esempio settembre 2008, possiamo vedere la tabella seguente:

**Tabella 7:** *Statistica mensile di utilizzo nel mese di settembre 2008*

Monthly Statistics for september 2008		
Total Hits	1345096	
Total Files	559775	
Total Pages	356620	
Total Visits	29536	
Total KBytes	21019879	
Total Unique Sites	12619	
Total Unique URLs	5946	
Total Unique Referrers	1228	
Total Unique User Agents	685	
	Avg	Max
Hits per Hour	1868	18643
Hits per Day	44836	113260
Files per Day	18659	37674
Pages per Day	11887	21302
Visits per Day	984	1860
KBytes per Day	700663	2404400

Rispetto a quelli già descritti, nella tabella appaiono alcuni termini nuovi, piuttosto tecnici:

**Total Unique Sites** è il numero di tutti gli indirizzi IP univoci che hanno visitato almeno una volta il sito nel periodo di riferimento. È un dato da valutare con prudenza per le avvertenze già esplicitate nelle note precedenti riguardo agli indirizzi IP.

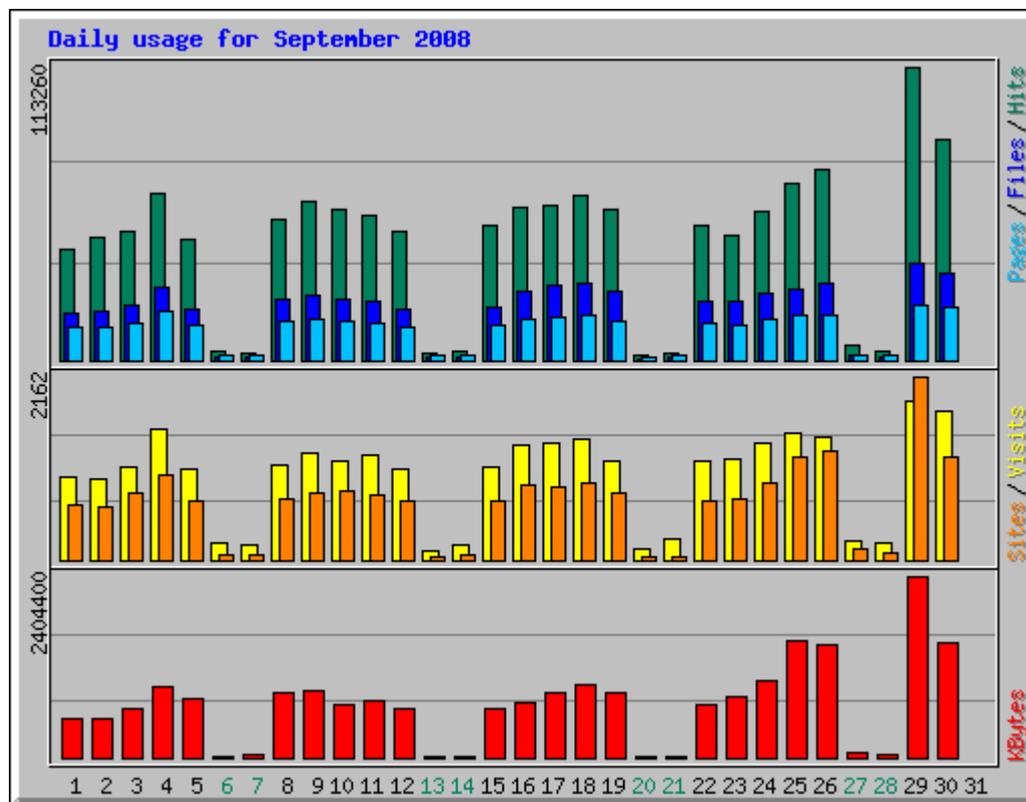
**Total Unique URLs** è il numero delle URL univoche, appartenenti al sito, richieste dai browser dei visitatori. Ricordiamo che ciascuna pagina di *indata* ha un proprio specifico URL:

<https://indata.istat.it/vela/registra.php> è la pagina per la registrazione iniziale di un rispondente dell'Indagine VELA; <https://indata.istat.it/turismo/modifica.php> è la pagina della Rilevazione sull'attività alberghiera per eseguire la modifica della password personale del rispondente.

**Total Unique Referrers** è il numero delle URL univoche delle pagine Html di provenienza che fanno richieste al sito. È da tenere presente che la maggior parte delle pagine del sito contiene link ad altri oggetti (immagini ecc.) sempre contenuti nel sito.

**Total Unique User Agents** è il numero complessivo dei browser (Explorer, Opera, Firefox, ecc. nelle loro varie versioni) e dei *crawler* (motori di ricerca che esaminano e catalogano ciclicamente tutte le pagine disponibili sul Web) con i quali *indata* è acceduto. Ciascuno di loro lascia una specifica traccia sul server ma alcuni browser consentono all'utente di alterare tale traccia. Se tutti i visitatori di *indata* utilizzassero uno stesso browser di una stessa versione e nessun crawler esaminasse le nostre pagine il valore di questo dato sarebbe 1.

**Grafico 2:** *Utilizzo giornaliero di indata nel mese di settembre 2008*



Gli stessi dati del grafico precedente sono forniti anche in forma tabellare:

Tabella 8: Utilizzo giornaliero di indata nel mese di settembre 2008

Daily Statistics for September 2008												
Day	Hits		Files		Pages		Visits		Sites		KBytes	
1	42508	3.16%	18457	3.30%	12335	3.46%	964	3.26%	646	5.12%	527670	2.51%
2	47140	3.50%	18955	3.39%	12759	3.58%	957	3.24%	629	4.98%	525635	2.50%
3	49524	3.68%	21298	3.80%	14413	4.04%	1098	3.72%	779	6.17%	647254	3.08%
4	64548	4.80%	27936	4.99%	18901	5.30%	1546	5.23%	999	7.92%	938500	4.46%
5	46844	3.48%	19899	3.55%	13393	3.76%	1069	3.62%	698	5.53%	787346	3.75%
6	3716	0.28%	1509	0.27%	2253	0.63%	202	0.68%	68	0.54%	20685	0.10%
7	2436	0.18%	899	0.16%	1946	0.55%	172	0.58%	53	0.42%	49084	0.23%
8	54610	4.06%	23533	4.20%	14669	4.11%	1125	3.81%	728	5.77%	849373	4.04%
9	61135	4.55%	25066	4.48%	15566	4.36%	1265	4.28%	796	6.31%	878549	4.18%
10	57973	4.31%	23179	4.14%	14969	4.20%	1175	3.98%	810	6.42%	706399	3.36%
11	56173	4.18%	22570	4.03%	14111	3.96%	1238	4.19%	774	6.13%	761607	3.62%
12	49797	3.70%	19767	3.53%	12802	3.59%	1081	3.66%	702	5.56%	648550	3.09%
13	2963	0.22%	813	0.15%	1735	0.49%	99	0.34%	45	0.36%	13231	0.06%
14	3335	0.25%	1097	0.20%	2026	0.57%	171	0.58%	58	0.46%	16407	0.08%
15	51629	3.84%	20182	3.61%	13646	3.83%	1088	3.68%	705	5.59%	641976	3.05%
16	58819	4.37%	26849	4.80%	15499	4.35%	1349	4.57%	874	6.93%	738348	3.51%
17	59388	4.42%	28524	5.10%	16647	4.67%	1384	4.69%	859	6.81%	847121	4.03%
18	63575	4.73%	29531	5.28%	17588	4.93%	1425	4.82%	914	7.24%	972348	4.63%
19	58525	4.35%	26447	4.72%	14712	4.13%	1170	3.96%	779	6.17%	851994	4.05%
20	1765	0.13%	397	0.07%	1277	0.36%	126	0.43%	27	0.21%	5660	0.03%
21	2507	0.19%	865	0.15%	1741	0.49%	245	0.83%	46	0.36%	12400	0.06%
22	52152	3.88%	22967	4.10%	13958	3.91%	1167	3.95%	701	5.56%	690678	3.29%
23	48092	3.58%	22763	4.07%	13236	3.71%	1180	4.00%	723	5.73%	802770	3.82%
24	57432	4.27%	25502	4.56%	15810	4.43%	1378	4.67%	896	7.10%	1007273	4.79%
25	68043	5.06%	27397	4.89%	17361	4.87%	1485	5.03%	1209	9.58%	1551243	7.38%
26	73356	5.45%	29501	5.27%	17189	4.82%	1436	4.86%	1271	10.07%	1495569	7.12%
27	5752	0.43%	2036	0.36%	2308	0.65%	218	0.74%	134	1.06%	79012	0.38%
28	3129	0.23%	903	0.16%	1855	0.52%	194	0.66%	81	0.64%	30173	0.14%
29	113260	8.42%	37674	6.73%	21302	5.97%	1860	6.30%	2162	17.13%	2404400	11.44%
30	84970	6.32%	33259	5.94%	20613	5.78%	1752	5.93%	1212	9.60%	1518624	7.22%

Sono informazioni di più facile interpretazione. È chiaramente visibile la brusca caduta per ogni fine settimana (sabato 6 e domenica 7, 13 e 14, ecc.). Gli accessi, comunque, sono sempre presenti, anche se in numero ridotto. A tali accessi corrispondono anche invii effettivi di questionari come si evince dai log di acquisizione delle indagini.

E' disponibile anche la distribuzione oraria degli accessi, sia in grafico che in tabella

Grafico 3: Utilizzo orario di indata nel mese di settembre 2008

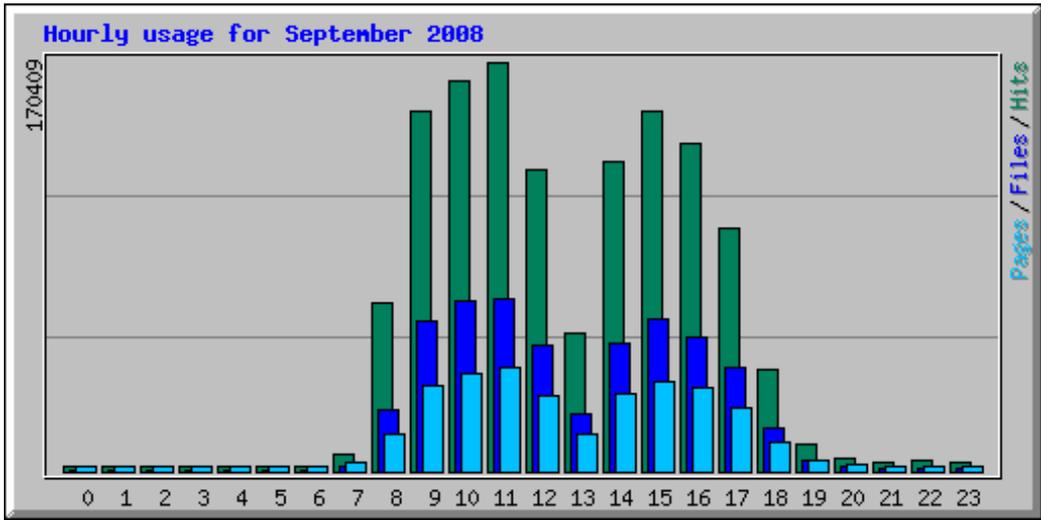


Tabella 9: Utilizzo orario di indata nel mese di settembre 2008

Hourly Statistics for September 2008												
Hour	Hits			Files		Pages		KBytes				
	Avg	Total		Avg	Total	Avg	Total	Avg	Total			
0	79	2375	0.18%	23	699	0.12%	67	2011	0.56%	266	7976	0.04%
1	78	2342	0.17%	22	666	0.12%	65	1952	0.55%	286	8575	0.04%
2	76	2289	0.17%	22	670	0.12%	64	1935	0.54%	280	8414	0.04%
3	67	2023	0.15%	16	490	0.09%	63	1890	0.53%	252	7575	0.04%
4	67	2028	0.15%	16	502	0.09%	62	1887	0.53%	203	6094	0.03%
5	67	2034	0.15%	16	499	0.09%	63	1899	0.53%	269	8065	0.04%
6	82	2462	0.18%	24	729	0.13%	66	1987	0.56%	361	10821	0.05%
7	249	7493	0.56%	71	2149	0.38%	115	3475	0.97%	2148	64445	0.31%
8	2329	69886	5.20%	860	25803	4.61%	520	15611	4.38%	35693	1070799	5.09%
9	4984	149537	11.12%	2088	62656	11.19%	1200	36004	10.10%	76007	2280205	10.85%
10	5425	162753	12.10%	2376	71296	12.74%	1350	40521	11.36%	89066	2671977	12.71%
11	5680	170409	12.67%	2381	71451	12.76%	1432	42984	12.05%	94140	2824207	13.44%
12	4188	125644	9.34%	1749	52479	9.38%	1037	31122	8.73%	61860	1855810	8.83%
13	1919	57599	4.28%	801	24057	4.30%	506	15194	4.26%	32788	983646	4.68%
14	4282	128481	9.55%	1777	53319	9.53%	1075	32277	9.05%	66928	2007853	9.55%
15	4988	149650	11.13%	2126	63797	11.40%	1249	37472	10.51%	76396	2291882	10.90%
16	4534	136024	10.11%	1862	55877	9.98%	1174	35238	9.88%	73192	2195748	10.45%
17	3374	101231	7.53%	1453	43598	7.79%	889	26693	7.48%	52411	1572321	7.48%
18	1409	42277	3.14%	602	18062	3.23%	403	12108	3.40%	25884	776514	3.69%
19	383	11497	0.85%	163	4904	0.88%	156	4684	1.31%	5046	151366	0.72%
20	177	5324	0.40%	72	2161	0.39%	99	2972	0.83%	1432	42964	0.20%
21	122	3682	0.27%	42	1275	0.23%	73	2210	0.62%	2017	60501	0.29%
22	149	4478	0.33%	49	1470	0.26%	77	2321	0.65%	2007	60218	0.29%
23	119	3578	0.27%	38	1166	0.21%	72	2173	0.61%	1730	51904	0.25%

I valori hanno un andamento chiaramente conforme all'usuale orario di lavoro, compresa la pausa per il pranzo. È comunque sempre presente una piccola parte di questionari inviati in ore inconsuete. A questo riguardo sono tipici, anche se ridotti, gli invii eseguiti durante i turni notturni degli alberghi per la rilevazione sull'attività alberghiera. L'esame congiunto della distribuzione degli accessi per giorno del mese e per ora, dimostra che il servizio di acquisizione dati è sempre operativo (7x7 h24).

Di un certo interesse è la tabella che segue, sempre riferita a Settembre 2008. Sono le pagine iniziali di entrata sul sito ordinate per numero di visite. Nella colonna di destra (URL) sono elencati gli indirizzi Web appartenenti a varie indagini (SIPI = Indagine mensile sulla produzione industriale, FATT = Indagine mensile sul fatturato e ordinativi, PRODCOM = Rilevazione annuale della produzione industriale, SCI = Rilevazione annuale del sistema dei conti delle imprese, PMI = Rilevazione sulle medie e piccole imprese, VELA = Rilevazione trimestrale sui posti vacanti e ore lavorate, IDRIG = Rilevazione sui servizi idrici (Grandi gestori). La prima URL della lista (/) si riferisce alla home page generale del sito di acquisizione – <https://indata.istat.it> – dal quale, in alternativa all'URL specifica, è possibile selezionare una delle indagini disponibili.

PROD07GES è la pagina di gestione di PRODCOM utilizzata dal personale interno per monitorare l'indagine, ripristinare le password iniziali dei rispondenti, ecc.

Nel caso di IDRIG le pagine iniziali sembrerebbero due, ma di fatto è sempre la stessa pagina referenziata in due modi diversi.

**Tabella 10:** *Lista delle dieci pagine iniziali maggiormente accedute nel mese di settembre 2008*

Top 10 of 415 Total Entry Pages					
#	Hits		Visits		URL
1	7313	0.54%	3625	21.70%	<a href="#">/</a>
2	5959	0.44%	1449	8.67%	<a href="#">/sipi/</a>
3	7251	0.54%	1041	6.23%	<a href="#">/fatt/</a>
4	37605	2.80%	943	5.65%	<a href="#">/prodc07/</a>
5	33388	2.48%	939	5.62%	<a href="#">/sci/</a>
6	8700	0.65%	601	3.60%	<a href="#">/pmi/</a>
7	863	0.06%	471	2.82%	<a href="#">/vela/formreg.php</a>
8	29	0.00%	432	2.59%	<a href="#">/idrig</a>
9	693	0.05%	373	2.23%	<a href="#">/prod07Ges/</a>
10	7147	0.53%	337	2.02%	<a href="#">/idrig/</a>

Abbiamo poi la tabella dei Referrers, ovvero, come già detto, le prime 30 URL delle pagine da cui provengono le richieste verso altre pagine del sito *indata*. Al primo posto (Direct Request) troviamo le richieste eseguite dagli utenti scrivendo direttamente, nella barra degli indirizzi del browser, il nome della pagina che vogliono visualizzare. Seguono poi le richieste eseguite partendo dalla home page del sito da cui è possibile entrare in tutte le home page delle singole rilevazioni.

È da tener presente che nella lettera inviata al rispondente, chiamato a fornire dati per una determinata rilevazione, viene sempre indicato l'URL specifico dell'indagine.

**Tabella 11:** *Lista dei primi trenta Referrers nel mese di settembre 2008*

Top 30 of 1228 Total Referrers			
#	Hits		Referrer
1	276027	20.52%	- (Direct Request)
2	113015	8.40%	<a href="https://indata.istat.it/">https://indata.istat.it/</a>
3	107439	7.99%	<a href="https://indata.istat.it/prodcom07/index.php">https://indata.istat.it/prodcom07/index.php</a>
4	76624	5.70%	<a href="https://indata.istat.it/">https://indata.istat.it/</a>
5	63889	4.75%	<a href="https://indata.istat.it/sci/index.php">https://indata.istat.it/sci/index.php</a>
6	51292	3.81%	<a href="https://indata.istat.it/idrig/index.php">https://indata.istat.it/idrig/index.php</a>
7	42010	3.12%	<a href="https://indata.istat.it/sci/index.php">https://indata.istat.it/sci/index.php</a>
8	37529	2.79%	<a href="https://indata.istat.it/pmi/index.php">https://indata.istat.it/pmi/index.php</a>
9	31272	2.32%	<a href="https://indata.istat.it/sci/">https://indata.istat.it/sci/</a>
10	30004	2.23%	<a href="https://indata.istat.it/prodcom07/">https://indata.istat.it/prodcom07/</a>
11	21647	1.61%	<a href="https://indata.istat.it/fatt/index.php">https://indata.istat.it/fatt/index.php</a>
12	20774	1.54%	<a href="https://indata.istat.it/vela/">https://indata.istat.it/vela/</a>
13	16689	1.24%	<a href="https://indata.istat.it/idrig/">https://indata.istat.it/idrig/</a>
14	16615	1.24%	<a href="https://indata.istat.it">https://indata.istat.it</a>
15	15923	1.18%	<a href="https://indata.istat.it/sipi/">https://indata.istat.it/sipi/</a>
16	15642	1.16%	<a href="https://indata.istat.it/vela/accesso.php">https://indata.istat.it/vela/accesso.php</a>
17	13964	1.04%	<a href="https://indata.istat.it/sipi/index.php">https://indata.istat.it/sipi/index.php</a>
18	12879	0.96%	<a href="https://indata.istat.it/bip/index.php">https://indata.istat.it/bip/index.php</a>
19	10431	0.78%	<a href="https://indata.istat.it/indata_file/terminati.php">https://indata.istat.it/indata_file/terminati.php</a>
20	10192	0.76%	<a href="https://indata.istat.it/fatt/">https://indata.istat.it/fatt/</a>
21	10182	0.76%	<a href="https://indata.istat.it/occ1/accesso.php">https://indata.istat.it/occ1/accesso.php</a>
22	9672	0.72%	<a href="https://indata.istat.it/prod07Ges/accessoAmmin.php">https://indata.istat.it/prod07Ges/accessoAmmin.php</a>
23	9557	0.71%	<a href="https://indata.istat.it/pmi/">https://indata.istat.it/pmi/</a>
24	9363	0.70%	<a href="https://indata.istat.it/occ1/">https://indata.istat.it/occ1/</a>
25	8691	0.65%	<a href="https://indata.istat.it/bipGes/vista.php">https://indata.istat.it/bipGes/vista.php</a>
26	8249	0.61%	<a href="https://indata.istat.it/ambint/index.php">https://indata.istat.it/ambint/index.php</a>
27	7602	0.57%	<a href="https://indata.istat.it/sciGes/stato_cod.php">https://indata.istat.it/sciGes/stato_cod.php</a>
28	7394	0.55%	<a href="https://indata.istat.it/fasweb/sito.php">https://indata.istat.it/fasweb/sito.php</a>
29	6118	0.45%	<a href="https://indata.istat.it/clag/accesso.php">https://indata.istat.it/clag/accesso.php</a>
30	5903	0.44%	<a href="https://indata.istat.it/datistat/">https://indata.istat.it/datistat/</a>

La tabella seguente si riferisce alle prime 20 stringhe utilizzate dai rispondenti nei motori di ricerca (Google, Yahoo, ecc.) per raggiungere i siti di acquisizione delle varie indagini. Tale tabella ha permesso di individuare un aspetto non previsto del comportamento dei rispondenti: alcuni di loro non utilizzano l'URL che gli viene comunicata nella lettera di invito a partecipare all'indagine per raggiungere direttamente il sito dell'indagine (ad esempio per l'Indagine PRODCOM <https://indata.istat.it/prodcom07>) ma utilizzano i motori di ricerca fornendo stringhe le più varie: <https://indata.istat.it/prodcom07>, <indata.istat.it/prodcom07>, <http://indata.istat.it/prodcom07>, ecc.

Tabella 12: Lista delle prime venti stringhe di ricerca nel mese di settembre 2008

Top 20 of 77 Total Search Strings			
#	Hits		Search String
1	129	18.19%	www.indata.istat.it/vela
2	75	10.58%	indata.istat.it/sciweb
3	70	9.87%	www.indata.istat.it
4	40	5.64%	https://indata.istat.it/sci
5	40	5.64%	indata
6	38	5.36%	indata.istat.it/prodcom07
7	28	3.95%	www.indata.istat.it/pmi
8	25	3.53%	www.indata.istat.it/sci
9	24	3.39%	www.indata.istat.it/vela
10	20	2.82%	indata.istat.it
11	12	1.69%	//indata.istat.it/prodcom07
12	12	1.69%	http://indata.istat.it/sciweb
13	12	1.69%	indata istat
14	9	1.27%	https://indata.istat.it/vela
15	9	1.27%	https://indata.istat.it/prodcom07
16	8	1.13%	https://indata.istat.it/pmi
17	7	0.99%	http://indata.istat.it/prodcom07
18	7	0.99%	https://indata.istat.it/prodcom07
19	7	0.99%	istat vela
20	7	0.99%	prodcom07

Segue, infine, la tabella degli User agent più utilizzati. L'informazione fornita conferma un esito prevedibile: circa l'85% dei rispondenti utilizza il browser della Microsoft Internet Explorer. È da considerare comunque che Webalizer non è in grado di rilevare alcuni browser di ultima generazione, di rilascio successivo all'ultimo aggiornamento del prodotto.

Nell'elenco compare anche Nagios: è un software utilizzato in Istat per controllare ciclicamente lo stato dei server dell'Istituto ed il loro corretto funzionamento. Tra questi è presente, ovviamente, anche *indata*.

**Tabella 13:** *Lista degli User agent maggiormente utilizzati nel mese di settembre 2008*

Top 15 of 685 Total User Agents			
#	Hits		User Agent
1	1142629	84.95%	Microsoft Internet Explorer
2	118326	8.80%	Netscape
3	24507	1.82%	Lynx
4	8247	0.61%	check_http/1.81 (nagios-plugins 1.4)
5	8173	0.61%	check_http/1.96 (nagios-plugins 1.4.5)
6	2583	0.19%	ha
7	563	0.04%	msnbot/1.1 (+http://search.msn.com/msnbot.htm)
8	358	0.03%	Microsoft Data Access Internet Publishing Provider Protocol D
9	289	0.02%	DoCoMo/1.0/N505i/c20/TB/W20H10 (compatible; Googlebot-Mobile/
10	223	0.02%	Opera/9.50 (Windows NT 5.1; U; it)
11	204	0.02%	Googlebot-Image/1.0
12	147	0.01%	Microsoft Office Protocol Discovery
13	135	0.01%	msnbot-media/1.1 (+http://search.msn.com/msnbot.htm)
14	103	0.01%	MSFrontPage/4.0
15	103	0.01%	Microsoft Data Access Internet Publishing Provider Cache Mana

## 5. PROSPETTIVE

Il potenziale della cattura dati potrebbe essere esteso a tutte le rilevazioni con vantaggi certi e significativi in termini di costi (alcune indagini sulle imprese hanno consentito risparmi sulle operazioni di stampa, scegliendo di inviare ai rispondenti i questionari cartacei solo su richiesta, e altre sulle istituzioni sono andate in onda solo su web) e tempestività (alcune indagini congiunturali sono riuscite ad anticipare la consegna delle stime ad Eurostat).

Presso i settori di produzione statistica, in relazione alle esperienze già maturate, aumentano sensibilità e aspettative, e conseguentemente cresce la domanda di interventi di innovazione delle fasi di raccolta. Una maggiore determinazione nel proporre l'innovazione al momento dell'invito ai rispondenti è necessaria da parte degli amministratori delle rilevazioni: deve essere valutata la possibilità di fornire il modello cartaceo solo su richiesta e in ogni caso la modalità web deve essere presentata come primaria. Ai rispondenti vanno presentati i vantaggi dell'acquisizione telematica e va data la possibilità di esprimere suggerimenti e critiche. Il rispondente deve sentirsi parte attiva del processo di innovazione.

L'attività di sviluppo sarà sempre più complessa (molte rilevazioni richiedono per esempio funzioni online di monitoraggio e gestione da parte di organi intermedi, siano essi i nostri uffici regionali o enti territoriali).

Anche l'attività di gestione si incrementa costantemente: per ogni indagine che adotta la cattura dati le successive ripetizioni richiedono comunque un'attività di aggiornamento, assistenza e monitoraggio; essendo numerose decine le indagini attive, assicurare il loro regolare svolgimento periodico è sempre più impegnativo.

Si configurano dunque prospettive di sviluppo sia quantitativo (numero delle rilevazioni trattate) sia qualitativo (ampliamento funzionale) che implicheranno interventi in termini di:

- Adeguamento risorse fisiche (le risorse server attualmente non sono dedicate ma condivise con altri sistemi applicativi – condizione diversa da quella prevista in fase di progettazione e che non opera a favore della stabilità dell’ambiente)
- Adeguamento architettura fisica (l’alta affidabilità non è completamente realizzata; recenti fermi dei servizi di rete e dei server non hanno prodotto perdita di informazione solo perché le applicazioni di cattura dati sono state realizzate prevedendo scritture multiple su server diversi)
- Analisi dell’impatto sul processo di produzione della rilevazione e sulle risorse locali (i settori che vogliono introdurre innovazione devono farsi carico del cambiamento organizzativo)
- Analisi dell’impatto sui contratti in essere per fornitura di servizi di stampa e spedizione modelli e registrazione dati
- Definizione di procedure di monitoraggio della raccolta più efficaci (collaborazione con i settori di produzione)
- Misurazione della qualità dei dati nella raccolta mix-mode per verificare se la tecnica influenza i parametri di qualità (collaborazione con i settori di produzione)
- Mantenimento di uno spazio di sperimentazione che preveda anche la possibilità di coinvolgere società di mercato (non si può rischiare l’obsolescenza se si propone innovazione)
- Approfondimento del requisito della usabilità e realizzazione di usability test
- Predisposizione di benefit informativi per i rispondenti (accesso diretto a informazione statistica prodotto della rilevazione, accesso ai propri dati precedentemente forniti)
- Predisposizione di altri benefit per i rispondenti (guide animate che illustrino le azioni richieste: dalla registrazione all’invio del questionario; sconto sulle pubblicazioni Istat ai rispondenti virtuosi, più solleciti, più assidui, ecc.)

## 6. ALLEGATO – RILEVAZIONI ATTIVE E TERMINATE

<i>indata</i> – Rilevazioni attive		
N	Rilevazione	Questionario
1	Dati ambientali nelle città	<i>t</i>
2	Rilevazione sui servizi idrici (Grandi gestori)	<i>d</i>
3	Indagine preliminare sistema informativo integrato sul mercato del lavoro	<i>t</i>
4	Indagine triennale sui laureati	<i>d</i>
5	Prezzi settore agricolo (I fase)	<i>d</i>
6	Indagine sulla capacità degli esercizi ricettivi	<i>e</i>
7	Rilevazione mensile sulle vendite al dettaglio	<i>t</i>
8	Rilevazione trimestrale del fatturato e ordinativi	<i>t</i>
9	Indagine mensile su fatturato e ordinativi	<i>e</i>
10	Indagine mensile su produzione industriale	<i>p</i>
11	Rilevazione mensile movimento clienti strutture ricettive	<i>d</i>
12	Rilevazione mensile su occupazione, orari di lavoro e retribuzioni	<i>t</i>
13	Rilevazione annuale del trasporto ferroviario	<i>t</i>
14	Rilevazione periodica sull'attività alberghiera	<i>t</i>
15	Rilevazione trimestrale su posti vacanti e ore lavorate	<i>t</i>
16	Rilevazione dei bilanci consuntivi degli enti pubblici	<i>e</i>
17	Rilevazione sulla caccia	<i>e</i>
18	Rilevazione semestrale estimativa della consistenza del bestiame	<i>t</i>
19	Rilevazione annuale delle tecnologie di informazione delle società finanziarie	<i>t</i>
20	Rilevazione annuale delle tecnologie di informazione della Pubblica Amministrazione	<i>t</i>
21	Rilevazione annuale delle tecnologie di informazione delle società non finanziarie	<i>t</i>

22	International sourcing - Dinamiche e modalità di internazionalizzazione delle imprese italiane	<i>t</i>
23	Rilevazione statistica sulle imprese a controllo estero	<i>t</i>
24	Rilevazione annuale sulla pesca nei laghi e bacini artificiali	<i>t</i>
25	Rilevazione sulle piccole e medie imprese	<i>e</i>
26	Rilevazione annuale della produzione industriale	<i>e</i>
27	Rilevazione annuale sulla struttura del costo del lavoro	<i>e</i>
28	Rilevazione annuale sui risultati economici delle aziende agricole	<i>e</i>
29	Rilevazione su ricerca e sviluppo negli enti e nelle istituzioni pubbliche	<i>e</i>
30	Rilevazione su ricerca e sviluppo nelle istituzioni private no profit	<i>e</i>
31	Rilevazione annuale del sistema dei conti delle imprese	<i>e</i>
32	Struttura retribuzioni 2006	<i>e</i>
33	Rilevazione annuale sulla struttura e produzione delle aziende agricole	<i>b</i>
34	Rilevazione annuale sulle telecomunicazioni	<i>e</i>

### Questionario

- t Teleform
- e Excel
- b Blaise
- p PHP
- s PHP Surveyor
- d Dataset

<i>indata</i> – Rilevazioni terminate		
<b>N</b>	<b>Rilevazione</b>	<b>Questionario</b>
1	Rilevazione annuale stima provvisoria del valore aggiunto	<i>e</i>
2	Rilevazione sugli archivi dei numeri civici	<i>t</i>
3	Rilevazione dei costi censuari dei Comuni	<i>t</i>
4	Rilevazione sull'uso dei dati censuari	<i>t</i>

### **Questionario**

- t Teleform
- e Excel
- b Blaise
- p PHP
- s PHP Surveyor
- d Dataset



## Bibliografia

- Accoto C. e Pavone L., *Misurare le audience in internet. Teorie, tecniche e metriche per la misurazione degli utenti in rete*, Franco Angeli, 2007
- Balestrino R. e Fanfoni L., “The quality of production processes and the electronic survey system: the experience of Istat Structural Business Statistics”, European Conference on Quality in Survey Statistics, 24-26 April 2006, Cardiff, UK
- Balestrino R., Macchia S. e Murgia M. “Data capturing strategies used in Istat to improve quality”, ECE/Eurostat/OECD Work Session on Statistical Editing, Bonn, 2006
- Conrad F.G. e Schober M.F., *Envisioning the Survey Interview of the Future*, John Wiley & Sons Inc, Wiley Series in Survey, 2007
- Visciola M., *Usabilità dei siti web*, Apogeo Editore, 2007



## Documenti ISTAT(\*)

- 1/2005 – Francesco Cuccia, Simone De Angelis, Antonio Laureti Palma, Stefania Macchia, Simona Mastroluca e Domenico Perrone – *La codifica delle variabili testuali nel 14° Censimento Generale della Popolazione*
- 2/2005 – Marina Peci – *La statistica per i Comuni: sviluppo e prospettive del progetto Sisco.T (Servizio Informativo Statistico Comunale. Tavole)*
- 3/2005 – Massimiliano Renzetti e Annamaria Urbano – *Sistema Informativo sulla Giustizia: strumenti di gestione e manutenzione*
- 4/2005 – Marco Broccoli, Roberto Di Giuseppe e Daniela Pagliuca – *Progettazione di una procedura informatica generalizzata per la sperimentazione del metodo Microstrat di coordinamento della selezione delle imprese soggette a rilevazioni nella realtà Istat*
- 5/2005 – Mauro Albani e Francesca Pagliara – *La ristrutturazione della rilevazione Istat sulla criminalità minorile*
- 6/2005 – Francesco Altarocca e Gaetano Sberno – *Progettazione e sviluppo di un "Catalogo dei File Grezzi con meta-dati di base" (CFG) in tecnologia Web*
- 7/2005 – Salvatore F. Allegra e Barbara Baldazzi – *Data editing and quality of daily diaries in the Italian Time Use Survey*
- 8/2005 – Alessandra Capobianchi – *Alcune esperienze in ambito internazionale per l'accesso ai dati elementari*
- 9/2005 – Francesco Rizzo, Laura Vignola, Dario Camol e Mauro Bianchi – *Il progetto "banca dati della diffusione congiunturale"*
- 10/2005 – Ennio Fortunato e Nadia Mignolli – *I sistemi informativi Istat per la diffusione via web*
- 11/2005 – Ennio Fortunato e Nadia Mignolli – *Sistemi di indicatori per l'attività di governo: l'offerta informativa dell'Istat*
- 12/2005 – Carlo De Gregorio e Stefania Fatello – *L'indice dei prezzi al consumo dei testi scolastici nel 2004*
- 13/2005 – Francesco Rizzo e Laura Vignola – *RSS: uno standard per diffondere informazioni*
- 14/2005 – Ciro Baldi, Diego Bellisai, Stefania Fivizzani, Annalisa Lucarelli e Marina Sorrentino – *Launching and implementing the job vacancy statistics*
- 15/2005 – Stefano De Francisci, Massimiliano Renzetti, Giuseppe Sindoni e Leonardo Tinolini – *La modellazione dei processi nel Sistema Informativo Generalizzato di Diffusione dell'ISTAT*
- 16/2005 – Ennio Fortunato e Nadia Mignolli – *Verso il Sistema di Indicatori Territoriali: rilevazione e analisi della produzione Istat*
- 17/2005 – Raffaella Cianchetta e Daniela Pagliuca – *Soluzioni Open Source per il software generalizzato in Istat: il caso di PHPSurveyor*
- 18/2005 – Gianluca Giuliani e Barbara Boschetto – *Gli indicatori di qualità dell'Indagine continua sulle Forze di Lavoro dell'Istat*
- 19/2005 – Rossana Balestrino, Franco Garritano, Carlo Cipriano e Luciano Fanfoni – *Metodi e aspetti tecnologici di raccolta dei dati sulle imprese*
- 1/2006 – Roberta Roncati – [www.istat.it](http://www.istat.it) (versione 3.0) *Il nuovo piano di navigazione*
- 2/2006 – Maura Seri e Annamaria Urbano – *Sistema Informativo Territoriale sulla Giustizia: la sezione sui confronti internazionali*
- 3/2006 – Giovanna Brancato, Riccardo Carbini e Concetta Pellegrini – *SIQual: il sistema informativo sulla qualità per gli utenti esterni*
- 4/2006 – Concetta Pellegrini – *Soluzioni tecnologiche a supporto dello sviluppo di sistemi informativi sulla qualità: l'esperienza SIDI*
- 5/2006 – Maurizio Lucarelli – *Una valutazione critica dei modelli di accesso remoto nella comunicazione di informazione statistica*
- 6/2006 – Natale Renato Fazio – *La ricostruzione storica delle statistiche del commercio con l'estero per gli anni 1970-1990*
- 7/2006 – Emilia D'Acunto – *L'evoluzione delle statistiche ufficiali sugli indici dei prezzi al consumo*
- 8/2006 – Ugo Guarnera, Orietta Luzi e Stefano Salvi – *Indagine struttura e produzioni delle aziende agricole: la nuova procedura di controllo e correzione automatica per le variabili su superfici aziendali e consistenza degli allevamenti*
- 9/2006 – Maurizio Lucarelli – *La regionalizzazione del Laboratorio ADELE: un'ipotesi di sistema distribuito per l'accesso ai dati elementari*
- 10/2006 – Alessandra Bugio, Claudia De Vitiis, Stefano Falorsi, Lidia Gargiulo, Emilio Gianicolo e Alessandro Pallara – *La stima di indicatori per domini sub-regionali con i dati dell'indagine: condizioni di salute e ricorso ai servizi sanitari*
- 11/2006 – Sonia Vittozzi, Paola Giacchè, Achille Zuchegna, Piero Crivelli, Patrizia Collesi, Valerio Tiberi, Alexia Sasso, Maurizio Bonsignori, Giuseppe Stassi e Giovanni A. Barbieri – *Progetto di articolazione della produzione editoriale in collane e settori*
- 12/2006 – Alessandra Coli, Francesca Tartamella, Giuseppe Sacco, Ivan Faiella, Marcello D'Orazio, Marco Di Zio, Mauro Scanu, Isabella Siciliani, Sara Colombini e Alessandra Masi – *La costruzione di un Archivio di microdati sulle famiglie italiane ottenuto integrando l'indagine ISTAT sui consumi delle famiglie italiane e l'Indagine Banca d'Italia sui bilanci delle famiglie italiane*
- 13/2006 – Ersilia Di Pietro – *Le statistiche del commercio estero dell'Istat: rilevazione Intrastat*
- 14/2006 – Ersilia Di Pietro – *Le statistiche del commercio estero dell'Istat: rilevazione Extrastat*
- 15/2006 – Ersilia Di Pietro – *Le statistiche del commercio estero dell'Istat: comparazione tra rilevazione Intrastat ed Extrastat*
- 16/2006 – Fabio M. Rapiti – *Short term statistics quality Reporting: the LCI National Quality Report 2004*
- 17/2006 – Giampiero Siesto, Franco Branchi, Cristina Casciano, Tiziana Di Francescantonio, Piero Demetrio Falorsi, Salvatore Filiberti, Gianfranco Marsigliesi, Umberto Sansone, Ennio Santi, Roberto Sanzo e Alessandro Zeli – *Valutazione delle possibilità di uso di dati fiscali a supporto della rilevazione PMI*
- 18/2006 – Mauro Albani – *La nuova procedura per il trattamento dei dati dell'indagine Istat sulla criminalità*
- 19/2006 – Alessandra Capobianchi – *Review dei sistemi di accesso remoto: schematizzazione e analisi comparativa*
- 20/2006 – Francesco Altarocca – *Gli strumenti informatici nella raccolta dei dati di indagini statistiche: il caso della Rilevazione sperimentale delle tecnologie informatiche e della comunicazione nelle Pubbliche Amministrazioni locali*
- 1/2007 – Giuseppe Stassi – *La politica editoriale dell'Istat nel periodo 1996-2004: collane, settori, modalità di diffusione*
- 2/2007 – Daniela Ichim – *Microdata anonymisation of the Community Innovation Survey data: a density based clustering approach for risk assessment*
- 3/2007 – Ugo Guarnera, Orietta Luzi e Irene Tommasi – *La nuova procedura di controllo e correzione degli errori e delle mancate risposte parziali nell'indagine sui Risultati Economici delle Aziende Agricole (REA)*

- 4/2007 – Vincenzo Spinelli – *Processo di Acquisizione e Trattamento Informatico degli Archivi relativi al Modello di Dichiarazione 770*
- 5/2007 – Anna Di Carlo, Maria Picci, Laura Posta, Michaela Raffone, Giuseppe Stassi e Fiorella Tortora – *La progettazione dei Censimenti generali 2010-2011: 1 - Analisi, valutazione e proposte in merito ad atti di normazione e finanziamento*
- 6/2007 – Silvia Bruzzone, Atonia Manzari, Marilena Pappagallo e Alessandra Reale – *Indagine sulle Cause di Morte: Nuova procedura automatica per il controllo e la correzione delle variabili demo-sociali*
- 7/2007 – Maura Giacommo, Carlo Vaccari e Monica Scannapieco – *Indagine sulle Scelte Tecnologiche degli Istituti Nazionali di Statistica*
- 8/2007 – Lamberto Pizzicannella – *Sviluppo del processo di acquisizione e trattamento informatico degli archivi relativi al modello di dichiarazione 770. Anni 2004 – 2005*
- 9/2007 – Damiano Abbatini, Lorenzo Cassata, Fabrizio Martire, Alessandra Reale, Giuseppina Ruocco e Donatella Zindato – *La progettazione dei Censimenti generali 2010-2011 2 - Analisi comparativa di esperienze censuarie estere e valutazione di applicabilità di metodi e tecniche ai censimenti italiani*
- 10/2007 – Marco Fortini, Gerardo Gallo, Evelina Paluzzi, Alessandra Reale e Angela Silvestrini – *La progettazione dei censimenti generali 2010 – 2011 3 – Criticità di processo e di prodotto nel 14° Censimento generale della popolazione e delle abitazioni: aspetti rilevanti per la progettazione del 15° Censimento*
- 11/2007 – Domenico Adamo, Damiana Cardoni, Valeria Greco, Silvia Montecolle, Sante Orsini, Alessandro Ortensi e Miria Savioli – *Strategie di correzione del questionario sulla qualità della vita dell'infanzia e dell'adolescenza. Indagine multiscopo sulle famiglie. Aspetti della vita quotidiana 2005*
- 12/2007 – Carlo Nappi – *Manuale per la preparazione di originali "ready to print"*
- 1/2008 – Franco Lorenzini – *Indagine sulle unità locali delle imprese: la flessibilità organizzativa e il ruolo degli uffici regionali come strategia per la riduzione del disturbo statistico e il raggiungimento di elevati tassi di risposta*
- 2/2008 – Elisa Berntsen, Simone De Angelis, Simona Mastroluca – *La progettazione dei Censimenti generali 2010-2011 4-L'uso dei dati censuari del 2000-2001: alcune evidenze empiriche*
- 3/2008 – Marina Peci – *Progetto SCQ -Scuola Conoscenza Qualità-Statistica e Studenti*
- 4/2008 – Giampiero Siesto, Franco Branchi, Cristina Casciano, Tiziana Di Francescantonio, Piero Demetrio Falorsi, Salvatore Filiberti, Gianfranco Marsigliesi, Umberto Sansone, Ennio Santi, Roberto Sanzo e Alessandro Zel – *Messa a regime dell'uso dei dati fiscali (Modelli UNICO) per l'integrazione delle mancate risposte e la riduzione del numero delle unità campione della rilevazione PMI*
- 5/2008 – Giovanni Seri e Maurizio Lucarelli – *A.D.ELE. Il laboratorio per l'Analisi dei Dati ELEmentari. Monitoraggio dell'attività Anni 2004-2007*
- 6/2008 – Francesco Altarocca – *Strumenti informatici innovativi nella conduzione di indagini statistiche*
- 1/2009 – Silvia Dardanelli, Simona Mastroluca, Alessandro Sasso e Mariangela Verrascina – *La progettazione dei censimenti generali 2010 – 2011 5 - Novità di regolamentazione internazionale per il 15° Censimento generale della popolazione e delle abitazioni*
- 2/2009 – Rossana Balestrino e Alberto Gaucci – *Tecniche di cattura dati nei processi di produzione statistica*