

Note Informative

La società

Audiweb SRL è il soggetto realizzatore e distributore dei dati sulla audience online.

Obiettivo primario di Audiweb è fornire informazioni oggettive e imparziali al mercato, di carattere quantitativo e qualitativo, sulla fruizione del mezzo Internet e su ogni altra rete o sistema online (quali ad esempio la telefonia cellulare o l'IPTV) utilizzando opportuni strumenti di rilevazione

Governance

Audiweb è una società partecipata da Fedoweb (50%), associazione degli editori online, da UPA Utenti Pubblicità Associati (25%), che rappresenta le aziende nazionali e multinazionali che investono in pubblicità e da Assap Servizi s.r.l., l'azienda servizi di proprietà di AssoComunicazione (25%), associazione delle agenzie e centri media operanti in Italia. Si configura quindi come una Joint Industry Committee con la partecipazione delle associazioni di categoria di tutti gli operatori del mercato. È gestita da un Consiglio di Amministrazione affiancato da un Comitato Tecnico che ha funzioni propositive e consultive in relazione all'impostazione delle rilevazioni, delle ricerche e della diffusione dei dati ottenuti.

Dati anagrafici

Audiweb srl Via Larga 13, 20122 Milano
Tel. +39. 02.58315141, Fax +39. 02.58318705
E-mail: audiweb@audiweb.it - www.audiweb.it
Partita Iva: 12521160155, REA: 1563606
Capitale Sociale Euro 30.000
Presidente e Legale Rappresentante: Enrico Gasperini

1. I Servizi di rilevazione

Audiweb fornisce al mercato, attraverso un sistema integrato di servizi, informazioni statisticamente attendibili sulla audience online e una base di dati elementari di navigazione del panel per rendere possibile ogni tipo di analisi da parte degli utenti

1.1 Struttura dei servizi di rilevazione

Audiweb realizza e fornisce al mercato un sistema integrato di servizi per la rilevazione e la diffusione di tutte le fonti di dati sulla fruizione dei media online (*Online Audience Measurement Service*), con il supporto metodologico di organizzazioni internazionali come IAB e di propri partner di ricerca.

Il sistema integra diverse fonti:

- **Ricerca di base.** Ricerca quantitativa costituita da una serie (*waves*) di interviste, su un campione della popolazione italiana residente (individui di 11-74 anni), effettuata mediante questionari strutturati con metodologia CAPI (Computer Aided Personal Interview) e finalizzata alla definizione dell'universo degli utenti Internet e alla descrizione delle loro caratteristiche in termini di profilo socio-demografico e attitudinale.
- **Catalogo.** Informazioni su *tutta* l'offerta editoriale disponibile su Internet, strutturata per gerarchia di navigazione (*Property, Brand, Domain, Channel...*), organizzata per categorie di contenuti editoriali e per macro-aggregazioni.
- **Sistema censuario.** Rilevazione oggettiva e completa (censuaria) di *tutta* l'attività di Internet sui siti del catalogo, tramite feedback tecnici verso i server di raccolta dati attivati dalla apertura di pagine web e altre azioni.
- **Panel.** Rilevazione oggettiva della fruizione di Internet sui siti del catalogo da parte di campioni statisticamente rappresentativi di popolazione, tramite l'installazione nei loro computer di *software meter*. L'operazione avviene nel pieno rispetto delle norme sulla privacy.

Le due fonti *Panel* e *Sistema censuario* utilizzano differenti metodologie di rilevazione e le stime dell'audience prodotte, per quanto sia grande il campione di utenza, possono essere differenti.

Ciò è dovuto alla estrema complessità di Internet, dove allo stesso contenuto editoriale si può accedere attraverso *device* diversi non tutti rilevati dal *software meter*. Ad esempio si può navigare su Internet con un telefono cellulare e non solo con un computer o collegarsi alla rete dall'estero.

Inoltre l'utilizzo di rilevazioni basate sul solo panel *non è da solo sufficiente per permettere la rilevazione anche dei siti di minori dimensioni o di sezioni meno frequentate di siti complessivamente grandi per quanto riguarda l'audience e il traffico*, causa l'esistenza di una soglia di rappresentatività al di sotto della quale non è possibile garantire l'adeguata affidabilità statistica dei dati, ma la cui somma di audience è rilevante data l'estrema frammentazione del mezzo Internet.

Il sistema censuario consente in aggiunta la *pubblicazione dei dati quotidiani nel giorno successivo*, rispondendo ad un'esigenza di mercato. Il sistema panel, invece, prevede una pubblicazione con cadenza mensile, per i limiti oggettivi dei costi del sistema di rilevazione in funzione delle risorse del sistema nazionale della pubblicità online.

Per i motivi suddetti e dato l'obiettivo di Audiweb di garantire la completezza di informazioni e l'accesso a tutti gli attori del mercato è indispensabile quindi utilizzare la doppia rilevazione panel e censuaria.

Per assolvere il proprio obiettivo di garantire una larga diffusione delle informazioni Audiweb *distribuisce le proprie informazioni fino al dato elementare del panel*. In tale modo si rende possibile per un sottoscrittore qualsiasi tipo di analisi, applicando sul database software applicativi per la costruzione, valutazione e ottimizzazione di piani media realizzati da operatori specializzati.

Per descrivere i servizi di rilevazione, i loro criteri di distribuzione, i relativi costi e le modalità di pubblicazione dei dati occorre distinguere tra varie tipologie di aziende richiedenti:

- **Publisher iscritti.** Sono gli editori, le concessionarie, gli aggregatori o i service provider che richiedono la rilevazione e la pubblicazione dei dati analitici di navigazione e di audience delle proprie *Property* nel sistema censuario e nel database dei dati elementari da panel. Qualora la *Property* non abbia installato un sistema censuario conforme alle specifiche Audiweb la pubblicazione avviene per il solo database panel. Se la dimensione di audience di una *Property* è inferiore alla soglia di rappresentatività del panel la pubblicazione avviene per il solo sistema censuario.
- **Utilizzatori professionali**, cioè i soggetti interessati ai dati per analisi di carattere pubblicitario (agenzie pubblicitarie o web, agenzie media, aziende investitrici, società di consulenza etc.).
- **Utilizzatori**, cioè una molteplicità di soggetti interessati ai dati per analisi di carattere scientifico, commerciale, divulgativo (aziende investitrici, società di consulenza, operatori dell'informazione, aziende produttrici di contenuti o servizi non censiti etc.).

Tutte le categorie delle aziende richiedenti possono avere accesso *a tutti i servizi e tutti i dati* seppur con un sistema di costi differenti, esclusivamente sottoscrivendo un contratto con Audiweb che consente la registrazione elettronica al sistema. Non sono previsti sconti per nessuna categoria di aziende richiedenti.

I servizi forniti sono:

- **AW Report**
Report sintetico quotidiano distribuito elettronicamente nella giornata successiva alla rilevazione
 - Riporta i dati principali di audience delle Property codificate nel catalogo derivati dal *sistema censuario* (ad es. browser unici, pagine viste, visite, tempo medio per pagina, durata della visita etc)

- *I dati si riferiscono solo alle Property dei Publisher iscritti presenti nel sistema censuario*
 - La fonte è la rilevazione AW Census
- **AW View**
Interfaccia che consente di visualizzare report sui dati di navigazione dell'intera offerta editoriale presente su Internet, aggregati per mese
 - Riporta i principali dati sintetici di navigazione delle Property codificate nel catalogo derivati dal *panel* (ad es. utenti unici, penetrazione, pagine viste, visite, tempo medio per pagina etc), divisi per *device e luogo* di utilizzo (ad es. PC a CASA o PC in UFFICIO).
 - *I dati si riferiscono a tutte le Property Internet, ovvero sia alle Property dei Publisher iscritti che a tutte le altre che superino una data soglia di audience, indipendentemente dalla loro iscrizione ad Audiweb*
 - L'accesso viene consentito con password personalizzate fornite ai sottoscrittori del servizio
 - La fonte è la rilevazione **AW Panel**
 - **AW Database**
File distribuiti elettronicamente con cadenza mensile contenenti:
 - AW Database RL (Respondent Level), ovvero i dati elementari del panel costituiti dai clickstream dei singoli componenti, normalizzati, ed aggregati sulla base di fasce orarie
 - l'archivio delle rilevazioni quotidiane del sistema censuario
 - il catalogo con le codifiche delle Property
 - *I dati di AW Database RL si riferiscono solo alle Property dei Publisher iscritti, di qualsiasi dimensione. Per le Property che hanno una audience inferiore alla soglia di rappresentatività del panel viene distribuito il solo dato censuario*
 - Il dato viene distribuito a software house abilitate da Audiweb a cui viene fornita la lista dei sottoscrittori del servizio.
 - Le fonti sono le rilevazioni **AW Census** e **AW Panel**
 - **AW Trends**
Report trimestrale distribuito elettronicamente
 - Contiene la ricerca sull'Universo degli utenti attuali e potenziali e sul modello di utilizzo di Internet anche in relazione con gli altri mezzi.
 - La fonte è la rilevazione **AW Ricerca di Base**

** Property: è il più alto livello di aggregazione di un insieme di siti, domini, brand e channel. La property fa generalmente capo ad un'azienda, organizzazione, agenzia di governo, gruppo privato o altra istituzione che abbia la proprietà del 50% + 1 di ciascun dominio e/url al proprio interno. Un dominio o un sito unico può essere iscritto ad un'unica property. Di conseguenza, un sito che è co-posseduto da due gruppi societari sarà inserito in una sola delle due property.*

**Publisher iscritto: è un editore online che si "iscrive" al servizio rappresentando una o più Property e consente ad Audiweb la pubblicazione dei dati analitici.*

1.2 Costo di accesso ai servizi di rilevazione

L'accesso ai servizi di rilevazione, comporta la sottoscrizione di un contratto con Audiweb. I costi sono diversi a seconda della tipologia dell'azienda richiedente. Tutti i corrispettivi sono indicati al netto dell'IVA.

Per gli *Utilizzatori* i costi dei servizi sono:

- **AW Report**, gratuito, previa registrazione elettronica sul sito www.audiweb.it
- **AW View, AW Trends, AW Database**, Euro 50.000

Se l'azienda che richiede il servizio non è una media company, ovvero non percepisce né direttamente né indirettamente ricavi di qualsiasi natura dal mezzo internet, allora può accedere al servizio al costo di 25.000 Euro.

Per le altre aziende richiedenti, per motivi di equità, i costi variano, da uno minimo per l'accesso fino a un tetto massimo, in funzione del fatturato pubblicitario gestito o della audience della *Property* iscritta.

Dati gli alti tassi di crescita del mercato pubblicitario online è prevista una revisione annuale sia delle classi di appartenenza che della componente variabile del costo, che tiene conto dei dati di crescita del mercato e degli investimenti sostenuti da Audiweb per lo sviluppo dei servizi.

Per gli *Utilizzatori professionali* i costi sono calcolati sulla base delle dimensioni del *online billing* dell'azienda, ovvero del fatturato pubblicitario online annuale gestito dal richiedente per conto dei suoi clienti, dichiarato alla sottoscrizione del contratto:

- con *online billing* < 2 milioni di Euro
Costo: 2.500 Euro + (*online billing* / 1.000) Euro
Accesso a AW Report e AW Database
L'Accesso a AW View e AW Trends comporta un costo aggiuntivo di 22.500 Euro
- con *online billing* > 2 milioni di Euro
Costo: 25.000 Euro + (*online billing* / 1.000) Euro, con un tetto massimo di 100.000 Euro.
Accesso a AW Report, AW View, AW Trends e AW Database

Per i **Publisher iscritti** i costi sono calcolati sulla base del numero delle **Property** iscritte e delle loro misure in termini di audience e traffico mensile. Un **Publisher iscritto** può definire una **Custom Property** ovvero l'aggregazione di una serie di **Property**, iscritte come un'unica entità seppur rilevate separatamente. Può definire anche una **Custom Property tematica**, ovvero l'aggregazione di una serie di **Brand** e **Channel** con tematiche affini, che vengono rappresentati unitariamente per scopi commerciali.

Per determinare il costo minimo di accesso, ogni *Property* iscritta viene divisa per classi, in funzione del suo *Active Reach*. Per la componente variabile del costo si utilizza una formula che misura la audience e il traffico ponderando in ugual misura i parametri di *Unique Audience* e *Total Minutes* calcolati come media mensile dell'ultimo semestre intero disponibile dalla rilevazione AW View. Per semestre intero si intende, per ciascun anno, rispettivamente gennaio–giugno (1° semestre) e luglio-dicembre (2° semestre):

- $Property\ W = (Unique\ Audience + Total\ Minutes / (1.000 * 47))$

La sottoscrizione del servizio dà diritto alla rilevazione della *Property* in AW Database e, qualora il publisher abbia installato un sistema di rilevazione censuaria conforme alle specifiche Audiweb e ne faccia esplicita richiesta, in AW Census, con i seguenti costi:

- con *Active Reach* < 1%
Costo: 10.000 Euro.
Accesso a AW Report, AW Trends e AW Database. L'accesso ad AW View comporta un costo aggiuntivo di 20.000 Euro
- con *Active Reach* compreso tra 1% e 3%
Costo: 10.000 Euro + 5,5 * *Property W*.
Accesso a AW Report, AW Trends e AW Database. L'accesso ad AW View comporta un costo aggiuntivo di 20.000 Euro
- con *Active Reach* compreso fra il 3% e il 10%
Costo: 40.000 Euro + 5,5 * *Property W*.
Accesso a AW Report, AW View, AW Trends e AW Database
- con *Active Reach* >= 10%
Costo: 60.000 Euro + 5,5 * *Property W*, con un tetto massimo di 250.000 Euro.
Accesso a AW Report, AW View, AW Trends e AW Database

Custom Property,

Viene trattata come se fosse una property unica. Per determinare i costi si prendono in considerazione la somma di Active Reach e Unique Audience deduplicate.

Accesso a AW Report e AW Database. Se il costo totale è superiore a 50.000 Euro, si accede anche ad AW View e AW Trends.

Per un publisher all'interno di una custom property che vuole accedere ad AW View, c'è un costo aggiuntivo di 20.000 Euro.

Custom Property tematiche, costo: 5.000 Euro, con la possibilità di realizzarne fino ad un massimo di 20.

1.3 Metodologia generale

Di seguito alcuni criteri generali e le metriche che sottendono l'intero sistema (*Parametri Audiweb*), che sono continuamente allineati agli standard internazionali di IAB (Internet Advertising Bureau), rimandando l'approfondimento metodologico alla descrizione delle varie rilevazioni.

Parametri Audiweb

Parametro	Riferito a	Fonte
Browsers Unici (Unique Browsers)		Censuario
Pagine Viste (Page Views)	Browser	Censuario
Visite (Sessions)	Browser	Censuario
Tempo medio per pagina (Time per Page)	Browser	Censuario

Durata Visita (Time per Session)	Browser	Censuario
Utenti Unici (Unique Audience)		Panel
Penetrazione Universo Attivi (Active Reach)	Utenti	Panel
Penetrazione Universo (Universe Reach)	Utenti	Panel
Pagine Viste (Page Views)	Utenti	Panel
Visite (Sessions)	Utenti	Panel
Tempo totale (Total minutes)	Utenti	Panel
Tempo medio per pagina (Time per Page)	Utenti	Panel
Durata Visita (Time per Session)	Utenti	Panel

Tabella - Parametri Audiweb

Per la fonte censuaria le metriche si intendono per il traffico depurato dalle attività di robot e spider e da eventuali errori.

VISITA - una o più richieste consecutive fatte dallo stesso visitatore all'interno di un sito con un tempo limite di inattività di 30 minuti. La ripresa dell'attività dopo 30 minuti sarà conteggiata come una seconda visita. Non devono essere considerate le attività di robot e spider.

BROWSER UNICI – il numero di browser diversi che, in un determinato arco temporale, effettuano una o più visite a un sito. Sono convenzionalmente identificati da un cookie unico nel caso dei sistemi browser based, o nel caso dei sistemi server based da una combinazione unica di IP address + un altro indicatore che può essere uno user agent, un cookie, un registration_ID. Deve essere comunque sempre disponibile, per eventuali controlli, il dato relativo al numero degli IP univoci nel medesimo arco temporale.

PAGINE VISTE - si intende per pagina un documento che contiene testo, immagini, suoni o altri oggetti. Si intende per PAGINA VISTA una richiesta esplicita, fatta da un utente, depurata dalle attività di robot, spider, e da eventuali codici di errore. Nel caso il sistema lo consenta, il tag di misurazione deve essere posizionato in fondo alla pagina e quindi sarà considerata pagina vista quella interamente scaricata e visualizzata sullo schermo dell'utente.

- Si considera PAGINA VISTA anche quella proveniente da cache e da proxy, se puntualmente misurabile.

- Pop-up, pop-under e splash pages = misurate solo se richieste dall'utente.

- Interstitial e jump pages = non misurate

- Sondaggi (Surveys) = misurate se richiesti esplicitamente dall'utente

- Autorefreshed pages = sono misurate tutte. Quelle con un tempo di refresh inferiore ai 5' dovranno essere esposte in modo separato e identificabile rispetto alle altre. Dovrà in ogni caso essere dichiarato dall'editore il tempo di refresh applicato.

- *Pagine a frame = per ogni pagina non può essere conteggiato più di un frame. Anche in questo caso sarà importante la corretta dichiarazione dell'editore su quale è il frame considerato più rilevante e che quindi deve essere conteggiato.*

- *Le pagine generate da sistemi chat, forum, instant message, personal pages sono misurate ma devono poter essere identificabili anche separatamente, se necessario.*

TEMPO MEDIO PER PAGINA - *il tempo in minuti e secondi trascorso sulla pagina.*

TEMPO TOTALE - *il tempo complessivo in minuti e secondi trascorso sulle pagine prese in esame.*

DURATA DELLA VISITA - *il tempo in minuti e secondi della visita.*

UTENTI UNICI - *sono i singoli individui che si sono collegati, in un determinato arco temporale, ad un sito e/o a pagine di esso effettuando una o più visite. Differiscono dai browser unici in quanto rappresentano persone fisiche.*

PENETRAZIONE UTENTI ATTIVI (ACTIVE REACH) – *percentuale di utenti, rispetto ad un universo di riferimento, la popolazione italiana attiva nel periodo di osservazione, che in un determinato arco temporale ha visitato un sito o parte di esso. Esempio: popolazione Internet italiana attiva : 20 milioni di persone, Utenti di un sito XXX: 10 milioni => Active Reach = 50%*

PENETRAZIONE UNIVERSO (UNIVERSE REACH) – *percentuale di utenti, rispetto all'universo totale degli utenti potenziali, cioè coloro che hanno la possibilità di collegarsi anche se non lo hanno necessariamente fatto nel periodo di osservazione, che in un determinato arco temporale ha visitato un sito o parte di esso.*

1.4 Partner per i servizi di rilevazione

Per la realizzazione delle rilevazioni, Audiweb si avvale della partnership con due Istituti di Ricerca:

- Doxa - per l'esecuzione su commissione della Ricerca di Base con l'obiettivo dell'alimentazione del panel, la stima degli Universi di riferimento e la rilevazione della diffusione dei diversi media (**AW Ricerca di Base**)
- Nielsen Online Italia - per la fornitura secondo le specifiche fornite da Audiweb e con un contratto di distribuzione esclusiva, dei dati da panel (**AW Panel**). Il sistema di gestione del panel, il relativo software meter e i dati sono di proprietà di Nielsen Online Italia

Inoltre Audiweb si occupa, mediante risorse interne o apposite società esterne, di sviluppare le attività e di gestire i processi relativi alla pubblicazione dei dati censuari giornalieri, tramite AW Report, dell'AW Database e della pubblicazione di tutte le informazioni sul sito www.audiweb.it

2. AW Ricerca di Base

Dati anagrafici società di ricerca partner

Doxa - ISTITUTO PER LE RICERCHE STATISTICHE E L'ANALISI
DELL'OPINIONE PUBBLICA - MILANO - Via B. Panizza, 7 - Tel. (02) 48.19.32.76

2.1 Struttura del servizio

Per il raggiungimento degli obiettivi di misurazione della Audience è necessario individuare e quantificare quale parte della popolazione Italiana 2-74 anni di età abbia accesso a Internet, con quali modalità specifiche acceda, da quali luoghi, con quali *device*, con quale frequenza lo utilizzi.

Sarà necessario non solo quantificare tale insieme di popolazione ma poterlo anche descrivere in termini di profilo sociodemografico e di eventuali altri parametri di comportamento verso il mezzo Internet.

Tali informazioni possono essere raccolte tramite interviste quantitative su questionario strutturato presso un campione probabilistico della popolazione italiana 11-74 anni di età.

Definiamo tale tipo di attività come Ricerca di Base (RdB) o Establishment Survey.

Le informazioni raccolte con la ricerca di base alimenteranno i seguenti flussi informativi o operativi

- Produzione di reportistica RdB sulle informazioni raccolte
- Definizione secondo parametri sociodemografici e comportamentali della griglia di stratificazione sulla quale si dovrà modellare il Panel di rilevazione
- Utilizzo dei nominativi degli intervistati nella ricerca di base per alimentare gli inserimenti degli individui e delle famiglie nel Panel di rilevazione

2.2 Metodologia utilizzata

La collettività a cui si fa riferimento per l'indagine Ricerca di Base Audiweb è costituita dalla popolazione italiana adulta, così definita: tutti i cittadini italiani di ambo i sessi, di età compresa fra i 11 e i 74 anni compiuti che risiedono nell'intero territorio nazionale. In particolare verranno comunque raccolte le informazioni relative anche alla popolazione di età tra i 2 e i 10 anni di età e alla popolazione di età superiore ai 74 anni qualora residenti in famiglie il cui intervistato fa parte del campione di riferimento (11-74 anni)

Vengono realizzate 7.200 interviste "face to face" a domicilio con riferimento all'individuo da effettuarsi, ogni anno, in quattro tranches stagionali, trimestrali, di 1.800 casi ciascuna, con tecnica di raccolta dei dati CAPI (Computer Aided Personal Interview).

La numerosità di 7.200 interviste / anno è riferita al totale individui, indipendentemente dal fatto che l'individuo o la famiglia abbiano o meno accesso a Internet.

Essendo d'altra parte uno degli scopi della Ricerca di Base quantificare e descrivere la parte di popolazione che ha accesso a Internet rispetto a quella che non ha accesso, occorre appunto che il campione sia rappresentativo della popolazione nel suo complesso, con o senza accesso Internet.

2.3 La consistenza del campione oggetto dell'indagine

Il campione casuale probabilistico è del tipo a tre stadi, cioè rispettivamente:

- primo stadio: scelta dei comuni
- secondo stadio: scelta di ulteriori unità di definizione territoriale (le “sezioni elettorali”)
- terzo stadio: unità “elettori”, che possono essere utilizzati come unità finali del campionamento diretto a “individui adulti” (viene intervistato l'individuo estratto).

Considerando la popolazione italiana divisa in celle di pre stratificazione in base a criteri geografici il campione sarà di tipo proporzionale. All'interno delle celle geografiche di pre stratificazione, la proporzionalità del campione all'universo sarà conseguita grazie alla scelta casuale degli intervistati secondo lo schema di campione a tre stadi illustrato.

Nel terzo stadio di selezione sarà estratto da ogni sezione elettorale un numero di nominativi pari ad un multiplo del numero di interviste da effettuare nella sezione elettorale.

Un numero di nominativi pari al numero di individui da intervistare sarà contrassegnato con la lettera “O” (nominativo “ordinario”), e cioè da intervistare senz'altro, mentre i rimanenti saranno contrassegnati con la lettera “R” (nominativo “di riserva”), da intervistare soltanto se – per un motivo indipendente dalla buona volontà dell'intervistatore – un nominativo “O” non potesse essere intervistato.

L'intervistatore effettuerà la selezione finale degli individui: in situazione ottimale potrà intervistare tutti gli individui assegnati, per un punto di campionamento (sezione elettorale), come nominativi “O”, ma in altri casi incontrerà rifiuti ed altri impedimenti nei tentativi di intervistare una parte dei nominativi “O”, e dovrà ripiegare su nominativi “R” di riserva. Dovrà comunque, in ogni punto di campionamento, completare il numero di interviste assegnato, osservando tutta una serie di regole e vincoli per l'utilizzazione dei nominativi e per la eventuale sostituzione con nominativi “R” dei nominativi “O” non potuti intervistare. Fra gli altri obblighi avrà anche quello di non lasciar cadere l'intervista ad un nominativo se non dopo aver effettuato un numero minimo di tentativi (almeno 4), in almeno 2 giorni diversi e in 4 orari diversi, oppure almeno 3 tentativi, di cui almeno uno dopo le ore 18 nei giorni feriali o di sabato e domenica.

Qualora anche i nominativi R vengano esauriti, l'intervistatore procede seguendo il criterio di sostituzione della famiglia più vicina.

I nominativi estratti dalle sezioni elettorali si dovranno ovviamente riferire alla sola fascia di popolazione 18 – 74 anni.

Data la necessità di rappresentare nella Ricerca di Base anche la fascia 11-17 anni per poter intervistare i minorenni si procederà nel modo seguente.

Per ciascuna sezione elettorale in cui dal piano di campionamento è previsto l'intervista ad un minorenne (11 – 17 anni), l'intervistatore incaricato di estrarre i nominati disporrà oltre ai nominativi “O” ed “R” anche un nominativo aggiuntivo indicato con “G”.

A partire da tale nominativo dovrà verificare se all'interno della famiglia legata al nominativo “G” esiste un giovane (11 - 17 anni). In caso positivo dovrà intervistarlo, in caso negativo utilizzerà il criterio della famiglia più vicina.

A posteriori, per meglio bilanciare il campione, sarà predisposto un accurato piano di ponderazione secondo più griglie di caratteri. La rete di caratteri tenuti sotto controllo dalla ponderazione sarà più ricca e completa di quella utilizzata per la pre-stratificazione

geografica, comprenderà infatti anche caratteri demografici e strutturali degli individui (responsabile acquisti, capofamiglia, numero di componenti della famiglia...), e, possibilmente, anche dell'abitazione (p. es. dotazione linee telefoniche di connessione ad Internet), e dovrà utilizzare modelli-universo sicuri, in quanto provenienti da statistiche ufficiali Istat o da altri fonte ritenute affidabili.

2.4 Le modalità di rilevazione

La rilevazione avviene mediante la proposizione agli individui del campione di Interviste quantitative tramite questionari strutturati gestiti in modalità CAPI (Computer Aided Personal Interview).

Il questionario analizza le seguenti aree :

- sussistenza in famiglia di un accesso a Internet e analisi della tipologia di accesso
- dotazioni tecnologiche della famiglia in generale
- accesso da parte dell' intervistato a Internet, o tramite la modalità identificata come disponibile alla famiglia, PC, o tramite altre modalità
- se l' intervistato lavora, accesso a Internet dal luogo di lavoro
- informazioni sulla fruizione da parte dell' intervistato di altri media (Radio, Tv etc)
- quadro generale dei componenti della famiglia

Al termine della intervista , relativamente agli individui utili per completare la numerosità necessaria dei Panel HOME_PROB e WORK_PROB l' intervistatore propone la adesione al panel, lascia all'intervistato il materiale illustrativo e le istruzioni per l'adesione al panel e se possibile esegue già con l'intervistato la procedura di adesione al panel Home con relativa installazione del meter.

2.5 distribuzione dei dati

I dati raccolti dalla Ricerca di Base vengono inviati Nielsen per:

- le necessità relative all'inclusione dei panelisti reclutati in AW Panel
- la ponderazione e l'espansione all'universo dei dati AW Panel
- La estensione ALNE e la clonazione ALE (vedi paragrafi 5.4 e 5.5)

I dati raccolti dalla Ricerca di Base vengono inoltre resi disponibili ad Audiweb secondo i seguenti formati o elaborazioni:

- Tavole statistiche relative al campione intervistato
- Tavole statistiche relative alle domande del questionario
- File con dati elementari (rigorosamente anonimi) relativo a tutte le risposte del questionario.

2.6 Il periodo di rilevazione

Continuativo con ciclo di produzione dei dati trimestrale.

2.7 Indirizzo Web di pubblicazione della metodologia

www.audiweb.it

3. AW Census

3.1 Struttura del servizio

AW Census è il sistema di rilevazione censuaria di Audiweb, che consente la pubblicazione dei dati *complessivi di traffico sui siti proveniente da utenti indistinti (browser)*, anche per le Property (o parti di esse) di minori dimensioni di audience, data l'esistenza di una soglia di rappresentatività al di sotto della quale non è possibile garantire l'adeguata affidabilità statistica dei dati da panel.

La rilevazione consente in aggiunta di dare una misura del traffico totale, incluso quello da device diversi non rilevati dal panel e quello dei residenti al di fuori del territorio nazionale

Un ulteriore importante valore aggiunto è la *pubblicazione dei dati quotidiani nel giorno successivo*, rispondendo ad un'esigenza di mercato.

Con AW Census sono prodotti, con frequenza **giornaliera** e relativamente ai domini e ai canali di *Property* iscritte al servizio, i dati relativi alle metriche di traffico e audience quali: pagine viste, browser unici, visite, tempo medio per pagina, durata della visita (parametri descritti nell'apposita sezione).

In una prima fase, già operativa da alcuni anni, i dati suddetti sono stati pubblicati sul sito Audiweb, solo su base mensile.

3.2 Metodologia utilizzata

Le informazioni di traffico sopra descritte sono raccolte mediante l'strumentazione di un apposito codice "Tag" nella pagine dei siti oggetto di rilevazione. Le elaborazioni vengono effettuate da server appositamente predisposti per la raccolta, la catalogazione e l'elaborazione delle informazioni secondo specifiche reportistiche e formati.

I server e i sistemi di tagging sono di terze parti. Audiweb provvede a:

- Regolare la gestione del tagging
- Certificare l'adeguatezza dei sistemi delle terze parti (rispetto alle indicazioni riportate nell'allegato C delle presenti Note, "Modalità di utilizzo dei dati censuari")
- Rilevare automaticamente i dati dalle terze parti
- Sottoporre a controlli campionari le Property per verificare la corretta applicazione

I sistemi certificati sono di tipo *browser based*, cioè sistemi basati sulla misurazione dei parametri Audiweb (page views, unique browser, ecc.) tramite l'impiego di server esterni che rilevano l'apertura completa delle pagine, mediante apposito tag inserito all'interno di tutte le pagine del sito/canale oggetto della rilevazione.

Sui dati censuari erogati da terze parti vengono effettuati, da parte di Audiweb, controlli specifici per garantire la corretta rilevazione, produzione e distribuzione degli stessi.

In particolare viene controllato che nel conteggio delle metriche (parametri Audiweb di cui sopra) vengano rimossi i contributi provenienti da utenze artificiali quali spider, robots, programmi di crawling dei motori di ricerca, ecc. Vengono inoltre controllate: la metodologia con la quale vengono effettuati i calcoli degli unique browser e identificate le pagine viste ossia se completamente scaricate dal browser.

I valori di pagine viste e tempo ottenute dal sistema censuario saranno utilizzati in modo cumulativo (somma dei valori censuari di gruppi di siti) nelle fasi di ponderazione di AW Panel e AW Database RL (vedi), questo soggetto ai seguenti punti :

- a) I valori di traffico censuario dovranno essere depurate dalle parti di traffico non rappresentate da AW Panel (navigazione da device diversi dai PC, quali cellulari etc, navigazione dall' estero, spider, automi).
- b) Tali dati censuari dovranno essere stati rilevati solo tramite sistemi di Tagging certificati da Audiweb come adeguati alla produzione di stime da utilizzarsi nelle procedure di ponderazione del panel
- c) Su tali dati censuari Audiweb potrà' effettuare un ulteriore livello di controlli di maggior dettaglio rispetto a quelli già esposti
- d) Rispetto ai dati pubblicati in AW Census i valori da usare per la ponderazione di AW Panel e AW Database RL dovranno essere resi coerenti con le modalità di conteggio del “meter” installato sui PC dei panelisti (es. eliminazione dai conteggi delle attività svolte su finestre diverse da quella “in focus”).

3.3 Il periodo di rilevazione

Continuativo, pubblicazione su base quotidiana.

3.4 Indirizzo Web di pubblicazione della metodologia

www.audiweb.it

4. AW Panel

Dati anagrafici società di ricerca partner

Nielsen Online Italia srl - Via Tortona, 35 20144 Milano - Telefono: 02/36578450, Fax: 02/54102323, Email: care-it@netratings.com
Capitale sociale: Cap. Soc. Euro 10.400 – Reg. Imprese 109515/2000 – REA 1615204
Partita Iva e Codice Fiscale: 13080470159

4.1 Struttura del servizio

La Ricerca di Base, utilizzando la metodologia del questionario quantitativo, può analizzare fenomeni “macro” per i quali l’intervistato possa dare un’affidabile risposta. La rilevazione dei dettagli dell’utilizzo di Internet (siti e sezioni di siti visitati, tempo speso etc) non può perciò essere rilevata tramite dichiarazione dell’intervistato.

Per ottenere tali informazioni occorre disporre di un campione continuativo, panel, di individui sui quali sia possibile effettuare una rilevazione tecnica dell’effettivo comportamento di navigazione.

La quantificazione della audience Internet e l’operatività del Database sono relative alla popolazione Italiana 2 - 74 anni di età e a target specifici descritti in termini di segmenti sociodemografici della popolazione suddetta. L’audience di Internet viene misurata, tramite apposito software meter con riferimento ai luoghi di fruizione del mezzo quali Casa e Ufficio e relativamente alla sola navigazione effettuata da computer.

La distribuzione delle informazioni rilevate da AW Panel avviene tramite due differenti servizi:

- **AW View:** è formato da un insieme di report predefiniti contenenti dati già elaborati e aggregati relativi ai parametri Audiweb di tutte le *Property* Internet misurate dal sistema. La frequenza di aggiornamento è mensile. Gli utenti accedono a AW View mediante browser Internet. La soglia di rappresentatività dei dati relativi alle singole Property è di 40 panelisti che hanno navigato sulla property almeno un volta nel mese oggetto di rilevazione
- **AW Database RL (Respondent Level):** è la componente del servizio AWDatabase che rappresenta l’insieme dei dati elementari analitici della *meter panel*, riferiti ai dati dei soli *Publisher iscritti*. I dati si riferiscono ad ogni singolo individuo partecipante al panel e fanno riferimento al profilo sociodemografico e ai dettagli analitici (a livello di clickstream secondo la categorizzazione delle url all’interno del catalogo AW View) della navigazione effettuata dall’individuo nel periodo temporale relativo allo specifico rilascio di AW Database. AW Database è rilasciato con frequenza mensile sotto forma di un insieme di file elaborabili con software di terze parti specifici per l’Analisi delle Audience e la Valutazione dei Piani, così come oggi già fatto per gli altri mezzi (radio, stampa, tv).

4.2 Metodologia utilizzata

Obiettivo del Panel è misurare in modo dettagliato il consumo di Internet effettuato da un campione di individui sui computer della famiglia (campione Home) e sul computer che eventualmente l’individuo utilizza sul lavoro (campione Work).

In funzione delle evidenze della Ricerca di Base, che invece, anche se su sola dichiarazione, analizza il consumo di Internet in ogni posto e con ogni tipo di device, la rilevazione effettuata sul Panel potrà essere in futuro estesa oltre gli attuali limiti.

La misurazione dettagliata del consumo Internet avviene installando sul/sui computer dell'individuo, a casa e al lavoro, un particolare software meter che registra automaticamente tutti i dettagli dell'attività effettuata sul computer. Tali informazioni sono poi trasmesse automaticamente dal software meter ad un'unità centrale.

Nel caso di PC familiari usati da più membri della famiglia sia all'accensione che dopo ogni periodo di inattività appare una schermata che richiede all'utente del computer la sua identificazione, in modo da poter associare allo specifico individuo della famiglia la sua effettiva navigazione.

Si sta vagliando la possibilità di introdurre un diverso sistema di identificazione automatica del singolo panelista, in grado di evitare/ridurre l'intrusività del sistema eliminando la comparsa della schermata iniziale di identificazione.

4.3 La consistenza del campione oggetto dell'indagine

4.3.1 I panel Home e Work

Da un punto di vista teorico risulterebbe necessario conoscere per un individuo del panel che abbia accesso a Internet sia da casa che dal lavoro il consumo effettuato dall'individuo nelle due situazioni.

Risulta, da evidenze passate, molto difficile ed economicamente oneroso, ottenere l'assenso alla installazione da parte dello stesso individuo dei due "meter", uno sul computer di casa e uno sul computer di ufficio.

Il campione del panel è perciò sdoppiato in due panel distinti.

Un primo panel è denominato "Home" e su questo campione di individui si misura il solo consumo effettuato dal computer di casa (anche se l'individuo ha dichiarato di poter accedere a Internet anche dal lavoro).

Un secondo panel è denominato "Work" e su questo campione di individui si misura il solo consumo effettuato dal computer di lavoro (anche se l'individuo ha dichiarato di poter accedere a Internet anche da casa).

Nel caso del panel "Home" sono considerati partecipanti al Panel tutti i componenti della famiglia presso la quale è presente un computer collegato a Internet, questo indipendentemente dal fatto che lo specifico individuo utilizzi o meno Internet (in effetti l'individuo ha sempre la possibilità di poterlo fare in quanto appunto nella sua famiglia è presente un computer collegato a Internet).

Il panel "Work" comprende invece i soli individui 16 -74 anni che hanno un accesso a Internet dal posto di lavoro da un computer utilizzato prevalentemente dall'individuo stesso (sono perciò esclusi gli accessi da computer normalmente condivisi da più persone).

Ciascuno dei due panel "Home" e "Work" è poi suddiviso in due segmenti, un segmento è relativo agli individui reclutati da un campione probabilistico e perciò in ipotesi effettivamente rappresentativi della popolazione indagata, l'altro segmento di individui componenti il panel è invece reclutato online.

La componente "ONLINE" dei panel e' formata da individui che hanno risposto positivamente ad una richiesta di collaborazione loro rivolta tramite:

- a) una email di invito ricevuta

- b) hanno selezionato in modo spontaneo un banner o un link incontrato durante la navigazione in Internet
- c) hanno ricevuto il questionario online di invito mediante comparsa casuale dello stesso durante la navigazione dell'utente.

Gli specifici meccanismi di attivazione dei questionari online, di selezione delle liste email alle quali inviare gli inviti all'adesione e le singole fonti di provenienza banner e link sono validati con opportune procedure atte a dimostrare la non eccessiva distorsione (fatto salvo il solo maggior volume di uso di Internet) della specifica fonte di reclutamento rispetto ai dati di navigazione ottenuti dalle componenti reclutate in modo probabilistico (HOME_PROB e WORK_PROB) del panel. Solo successivamente a tale validazione sono utilizzati nella effettiva produzione gli individui reclutati dalla specifica fonte nel panel online.

Tale segmento di campione online non può ritenersi in sé rappresentativo della popolazione indagata in quanto, proprio per la modalità di reclutamento, è prevalentemente composto da individui con un consistente livello di tempo speso su Internet (aventi perciò maggior probabilità di essere esposti ai questionari online e ai banner/link di reclutamento) e in certi casi con una tipologia particolare di navigazione funzione delle specifiche sorgenti di reclutamento. Per tale fatto è necessario riponderare i panel reclutati online rispetto a parametri derivati dai panel HOME_PROB e WORK_PROB.

Sono date le seguente definizioni:

- **CP = Current Panelist:** è tutto il campione dei panelisti reclutati e sotto misurazione mediante meter installato sul computer del panelista, indipendentemente dall'essere validati ed inseriti nel processo di produzione.
- **DMS = Digital Media Sample:** è il sottoinsieme dei CP, che, nel periodo di osservazione, sono stati validati (indipendentemente dall'aver navigato o meno nel periodo stesso) e quindi concorrono a quantificare l'universo complessivo di riferimento (“Current Digital Media Universe Estimate”)

In sostanza la struttura del Panel è suddivisa nei quattro sottoinsiemi di DMS:

HOME_PROB = campione probabilistico individui rilevati sul computer di casa

HOME_ONLINE = campione reclutato online di individui rilevati sul computer di casa

WORK_PROB = campione probabilistico individui rilevati sul computer di lavoro

WORK_ONLINE = campione reclutato online di individui rilevati sul computer di lavoro

Il panel è composto attualmente di circa 18.000 CP e circa 14.600 DMS, con un obiettivo di crescita fino a circa 20.000 DMS, con il vincolo che le proporzioni delle componenti “probabilistica” e “online” saranno rispettivamente il 22,5% (4.500 individui) ed il 77,5% (15.500 individui). Tale campione sarà reso coerente con i fattori di efficienza statistica riportati di seguito (“Ponderazione del panel”).

Considerando i panel HOME_PROB e WORK_PROB viene mensilmente tenuto sotto controllo il criterio fornito dal cosiddetto “fattore di proiezione medio” (con riferimento alle numerosità di utenti in internet, come risultanti dalla Ricerca di Base, da rappresentare con ogni panelista). Qualora la differenza tra i fattori dei 2 panel sia superiore al 30%, viene attuato un intervento teso al riequilibrio di detto parametro incrementando adeguatamente la numerosità campionaria del campione che presenta il più alto valore del “fattore di proiezione medio” (vedi esemplificazione della modalità di calcolo nell'Allegato Note Informative A).

Analoga operazione viene effettuata con riferimento ai campioni HOME_ONLINE e WORK_ONLINE.

Sono inoltre sottoposti a controllo sistematico gli elementi di stratificazione del panel WORK_PROB, (sesso / classi di età / area geografica / settore merceologico e dimensione dell'azienda) e del panel HOME_PROB (sesso / classi di età / area geografica).

Ciascun panel rispetta il mix di allocazione campionaria derivante da AW Ricerca di Base rispetto alle categorie delle variabili sopraesposte. Per la numerosità campionaria di ciascuna categoria è accettata una variabilità massima del 30% rispetto al valore teorico desunto da AW Ricerca di Base (vedi esemplificazione della modalità di calcolo nell' Allegato Note Informativa A) .

I campioni HOME_PROB e WORK_PROB sono stati reclutati fino al 2007 con metodologia RDD, cioè tramite interviste telefoniche effettuate ad un campione casuale di numeri telefonici.

Dal 2008 i nuovi individui inclusi nei campioni HOME_PROB e WORK_PROB sono reclutati prioritariamente nell'ambito della attività della Ricerca di Base, .

Il fatto che i nuovi individui partecipanti al panel derivino dalla Ricerca Di Base, e cioè per ognuno di essi già si conosca un significativo numero di informazioni, permette di verificare la assenza di distorsioni per gli individui che accettano di far parte del panel.

Infatti, confrontando le risposte al questionario Ricerca di Base fra gli individui che hanno accettato di entrare nel panel e quelle invece di chi non ha accettato, si può verificare se il profilo di chi ha accettato sia o meno differente dal profilo di chi non ha accettato.

In caso di differenza del profilo si interviene con fattori correttivi nella fase di ponderazione dei campioni HOME_PROB e WORK_PROB (vedi Ponderazione del panel).

Relativamente al campione WORK_PROB sono inclusi, per la componente relativa alla "dimensione dell'azienda superiore a 300 addetti", oltre agli individui derivati dalla Ricerca di Base, anche individui selezionati in base a procedure di Company Recruitment.

La procedura di Company Recruitment si basa sull'individuazione di un certo numero di aziende superiori a 300 addetti, appartenenti a vari settori merceologici. In tali aziende si ottiene in primo luogo la collaborazione della azienda stessa nel permettere di installare in un certo numero di computer dei propri dipendenti il "meter" per la misurazione del traffico internet. Successivamente viene proposto ad un campione casuale dei dipendenti delle aziende, aventi un computer aziendale collegato a Internet e utilizzato dal dipendente in modo esclusivo o prevalente, l'adesione al campione WORK_PROB.

Il tasso di rotazione complessivo del panel sarà almeno del 20% annuo.

4.3.2 La ponderazione dei panel

Ciascuno dei 4 sottoinsiemi del panel viene sottoposto ad una fase di ponderazione per riequilibrare eventuali distorsioni e non proporzionalità del panel stesso rispetto agli universi rappresentati.

Per quanto riguarda i panel HOME_PROB e WORK_PROB essi derivano è vero da un campionamento probabilistico, perciò in teoria non distorto, ma non sussistendo un tasso di accettazione del 100%, per l'evidente presenza di un numero significativo dei potenziali individui estratti per la inclusione nel panel, ma che in realtà non accettano la proposta di collaborazione, è necessario verificare che il profilo di chi ha accettato corrisponda al profilo della popolazione studiata. Sussistendo inevitabili differenze (ad

esempio il tasso di accettazione può essere più basso per individui di certe fasce di età) occorre riponderare il campione in modo che il suo profilo ponderato corrisponda al profilo della popolazione studiata.

La ponderazione viene effettuata con riferimento ad informazioni note e ritenute affidabili.

I campioni HOME_PROB e WORK_PROB sono perciò ponderati con riferimento ai profili derivati dalla Ricerca di Base.

In particolare si verifica, e si ottiene tramite la ponderazione, il perfetto allineamento dei campioni rispetto ai seguenti parametri:

HOME_PROB : sesso/età/ area geografica

WORK_PROB : sesso/età/dimensione aziendale/settore merceologico aziendale

Per quanto riguarda i panel HOME_ONLINE e WORK_ONLINE essi sono campioni costruiti su un'autoselezione degli individui che hanno risposto positivamente ad un invito alla adesione al panel apparso durante la attività di navigazione su Internet.

Per il meccanismo stesso di attivazione della proposta di collaborazione è ipotizzabile che tali panel sovrarappresentino i "grandi consumatori" di Internet, cioè le persone che dedicano più tempo alla navigazione in Internet.

Dato che dai panel HOME_PROB e WORK_PROB, una volta riponderati sui dati della Ricerca di Base, si ottengono, dai dati raccolti dal meter, le effettive indicazioni della distribuzione della frequenza di consumo di Internet, sui siti visitati etc, tali informazioni derivate dai panel HOME_PROB e WORK_PROB possono essere utilizzate come ulteriori informazioni di riferimento verso le quali bilanciare i panel HOME_ONLINE e WORK_ONLINE.

Si ottiene perciò il perfetto allineamento dei campioni rispetto ai seguenti parametri :

HOME_ONLINE = tutti quelli utilizzati per HOME_PROB e inoltre le classi di volume di navigazione e le classi di tipologie dei siti visitati

WORK_ONLINE = tutti quelli utilizzati per WORK_PROB e inoltre le classi di volume di navigazione e le classi di tipologie dei siti visitati

La ponderazione di ognuno dei quattro panel viene effettuata mensilmente con la metodologia IPF (Iterative Proportional Fitting).

Gli Indicatori Mensili di Efficienza delle quattro ponderazioni sono uguali o superiori ai seguenti valori:

HOME_PROB = 70%

HOME_ONLINE = 35%

WORK_PROB = 70%

WORK_ONLINE = 35%

Gli Indicatori di Efficienza sono calcolati applicando la seguente formula separatamente per gli individui di ciascuno dei quattro panel sopra indicati.

$$\frac{100.0 (\sum_j P_j R_j)^2}{\sum_j P_j \sum_j P_j R_j^2}$$

dove:

P_j è il Prepeso (=1) per l'individuo "j" e R_j è il peso calcolato dalla procedura IPF per l'individuo "j". La sommatoria è estesa a tutti i "j" individui di ciascun panel

L'efficienza quantifica se il campione è ben bilanciato rispetto all'universo di riferimento: ogni singolo peso attribuito ad ogni individuo con valore più basso o più alto rispetto ad un peso "ideale" uguale a 1 abbasserà il valore percentuale della statistica utilizzata. Quindi più la percentuale di efficienza è alta, più il campione sarà bilanciato.

Dopo la fase di ponderazione i dati di navigazione raccolti dal meter sui panel possono essere considerati rappresentativi della navigazione Internet della popolazione rappresentata, cioè la popolazione Italiana 2-74 anni avente accesso a Internet da casa e 16-74 anni avente accesso a Internet dal lavoro.

4.4 Le modalità di rilevazione

La rilevazione delle informazioni necessarie a AW View e ad AW Database RL avviene tramite :

- Relativamente al profilo sociodemografico: questionari autocompilati somministrati via Internet. In realtà per tutti gli individui inseriti nel panel provenienti dalla Ricerca di Base si conoscono già tutte le informazioni del questionario della Ricerca di Base, in tal caso il ricorso ai questionari autocompilati via Internet è limitato al solo aggiornamento dei dati o alla raccolta di informazioni addizionali.
- Relativamente alla attività di navigazione in Internet: la raccolta delle informazioni è effettuata in modalità completamente automatizzata e trasparente per il navigatore tramite apposito software meter installato sul computer. Il software installato memorizza in tempo reale il log di tutte le attività di navigazione effettuate (URL delle pagine viste, etc) e provvede a scadenze temporali predefinite a trasmettere automaticamente tale log attraverso Internet ad un sistema centrale di raccolta dati, sistema ove avvengono le successive fasi di elaborazione e consolidamento delle informazioni giunte da tutti i panelisti.

Nel caso di computer familiari usati da più membri della famiglia sia all'accensione che dopo ogni periodo di inattività appare una schermata che richiede all'utente del computer la sua identificazione, in modo da poter associare allo specifico individuo della famiglia la sua effettiva navigazione.

4.5 Il periodo di rilevazione

La rilevazione è continuativa. Per ogni individuo appartenente al panel viene rilevata senza interruzione la totalità della navigazione effettuata, dal suo momento di inserimento nel panel e relativa validazione sino al momento della sua cancellazione dal panel stesso.

L'elaborazione e aggregazione delle informazioni raccolte avviene alle scadenze temporali già definite per AW View e AW Database RL.

4.6 Indirizzo Web di pubblicazione della metodologia

www.audiweb.it

5. AW Database RL

5.1 La costruzione di AW Database RL

Il principale risultato al quale si arriva con i dati raccolti dal panel è il mettere a disposizione degli utilizzatori (pianificatori, concessionarie ed editori) una base di dati, AW Database RL, che possa essere elaborata con gli appositi software di Analisi delle Audience e di Valutazione dei Piani, così come oggi già fatto per gli altri mezzi (radio, stampa, tv).

Tale base di dati deve essere a livello individuale (respondent level), cioè contenere, per i singoli individui del campione rappresentativo degli utenti del mezzo, i dati descrittivi dell'individuo in termini sociodemografici e di consumo del mezzo stesso.

I dati sociodemografici e il livello di informazioni a base singoli individui sono necessari per permettere analisi e valutazioni di piani su target definiti liberamente dagli utenti.

La descrizione del consumo del mezzo da parte dei singoli individui sarà rappresentato come descritto in Allegato Note Informative B e si riferirà a ciascuno dei siti/canali per le Property iscritte ad Audiweb e sarà invece descritto con la denominazione "Altro" per tutti i siti/canali delle property non iscritte ad Audiweb.

La costruzione di AW Database RL avviene secondo la procedura delineata nei paragrafi successivi.

Scopo di tale procedura è assegnare a ciascuno degli individui del panel la globalità della sua navigazione in Internet sia da Home che da Work (procedura di Fusione) sia da accessi da Altri Luoghi (procedura di Estensione) sempre tramite computer (trascurando invece per ora accessi tramite device quali PDA e telefoni cellulari).

Viene poi aggiunta in AW Database RL tramite la procedura di Clonazione la navigazione effettuata da altri individui 2-74 anni esclusivamente in Altri Luoghi (Internet caffè, computer condivisi in università, scuole, altre istituzioni, case di amici e parenti etc).

Tali individui utilizzando Internet su computer esclusivamente al di fuori dell'ambiente Home e Work non sono rappresentati in AW Panel e perciò inizialmente non sono neppure rappresentati in AW Database RL. La procedura di Clonazione permette di aggiungere tale tipologia di individui ad AW Database RL che diviene così rappresentativo di tutta la navigazione effettuata in Italia in Internet tramite computer dalla popolazione Italiana 2-74 anni (come già detto resta esclusa la navigazione derivata da terminali mobili quali PDA e telefoni cellulari).

Infine considerato che da AW Census sono ricavabili informazioni censuarie e non campionarie sul totale consumo del mezzo Internet espresso come pagine viste e come tempo totale (depurando AW Census delle componenti non rappresentate da AW Database RL, cioè la navigazione da terminali mobili e la navigazione effettuata al di fuori del territorio italiano) tramite la procedura di Convergenza si verifica che per alcuni dei principali indicatori derivati da AW Census sussista una coerenza con l'analogia stima derivata da AW Database RL.

5.2 Fusione dei dati Home e Work.

La combinazione dei dati Home e Work, rilevata come già esplicitato su due diversi campioni indipendenti, avviene attraverso una procedura di "merging" (fusione) delle due basi di dati.

Considerando ogni individuo del campione Home (del quale perciò già si dispone della navigazione Home rilevata dal meter) che abbia anche dichiarato di navigare in Internet da Work la procedura di “merging” permette di attribuire allo stesso l’attività di navigazione Work rilevata dal meter in un individuo “gemello” rilevato nel campione Work.

L’individuazione del “gemello” sarà effettuata con riferimento sia a variabili di tipo sociodemografico (sesso, età etc) che a variabili relative alle abitudini di navigazione di ciascun individuo nei due ambienti Home e Work.

A valle della procedura di merging a ciascun individuo saranno perciò attribuiti i dati di navigazione completi, effettuati sia da Home che da Work.

Il meccanismo di fusione potrebbe comportare la necessità di integrare il questionario sottoposto ai panelisti (attuali e futuri, sia reclutati in modo probabilistico che online, sia nel panel work che in quello home) con domande specifiche atte a raccogliere informazioni indispensabili per poter procedere ai processi di Fusione/Estensione/Clonazione.

5.3 Verifiche di controllo sulla metodologia di Fusione dei dati Home e Work.

Relativamente ad un sottoinsieme di panelisti Work, Audiweb potrà, previo comune accordo con NNR, prevedere l’installazione del meter anche presso il computer utilizzato dallo stesso individuo presso l’abitazione (Home).

Tale insieme di panelisti, sui quali viene misurato l’effettivo consumo sia in Home che in Work potrà essere utilizzato come benchmark sul quale verificare la coerenza della metodologia applicata nella procedura di Fusione.

5.4 Estensione alla navigazione ALNE, Altri Luoghi (non esclusiva).

Per ogni individuo per il quale è già stata rilevata una navigazione in Home e/o Work si verifica la sua eventuale navigazione anche da Altri Luoghi (sempre tramite computer).

Conoscendo dalla Ricerca di Base la numerosità e il profilo demografico degli individui che navigano da Altri Luoghi oltre che da Home o Work viene costruito un campione casuale di individui di AW Database RL ai quali attribuire una navigazione da Altri Luoghi.

Identificati gli individui ad ognuno di essi vengono aggiunte una serie di sessioni da navigazione, analoghe a quelle effettuate dallo stesso individuo a Home o Work, in modo da rappresentare correttamente anche la frequenza di connessione da Altri Luoghi come rilevata dalla Ricerca di Base.

5.5 La navigazione ALE , Altri Luoghi (esclusiva), rappresentazione per clonazione.

Conoscendo dalla Ricerca di Base la numerosità e il profilo demografico degli individui che navigano esclusivamente da Altri Luoghi viene costruito un campione casuale di individui di AW Database RL atto a rappresentare la numerosità e il profilo demografico di tale popolazione. Ogni individuo di AW Database RL selezionato viene perciò duplicato (clonato) e viene aggiunto ad AW Database RL.

Per ogni individuo clonato ed aggiunto ad AW Database RL a rappresentare la navigazione ALE vengono considerate una serie di sessioni da navigazione dell’individuo dal quale è stato clonato, in modo da rappresentare correttamente anche la frequenza di connessione esclusiva da Altri Luoghi come rilevata dalla Ricerca di Base.

5.6 Estensione di AW Database RL al complemento fra popolazione rappresentata da AW Ricerca di Base e potenziali navigatori rappresentati da AW panel

Aw Panel e la sua trasposizione originale in AW Database RL rappresentano la popolazione potenzialmente navigatrice in Internet da Home o Work come dettagliato nel paragrafo 4.3.1, la clonazione ALE, paragrafo 5.5, ha aggiunto in AW Database RL anche gli individui con navigazione esclusiva da altri luoghi.

AW Ricerca di base è invece un campione rappresentativo della totalità della popolazione italiana 11- 74 anni con l'estensione agli individui di età 2-10 anni e superiore ai 74 anni appartenenti alla famiglia dell'intervistato.

Da AW Ricerca di base è perciò possibile conoscere la numerosità e il profilo demografico della popolazione che non è potenzialmente navigatrice nè da Home nè da Work nè da Altri luoghi.

5.7 Caratterizzazione degli individui di AW Database RL

Per ogni individuo inserito in AW Database RL saranno disponibili oltre alle informazioni sulla navigazione effettuata nel mese sui siti iscritti ad Audiweb (secondo il formato descritto nell'Allegato Note Informative B) anche le seguenti informazioni descrittive :

- Sesso
- Età in anni
- Professione
- Scolarità
- Area Geografica
- Eventuali altre informazioni di profilazione.

5.8 La convergenza fra i principali indicatori derivati da AW Census e le analoghe stime da AW Database RL.

Dalla fonte censuaria utilizzata per AW Census viene depurata, la navigazione da device mobili (PDA/cellulari), la navigazione dal di fuori del territorio italiano ed eventuali altre tipologie di navigazione non rappresentate in AW Database RL (per es. console giochi).

Rispetto ai dati utilizzati in AW Census i valori da usare per la ponderazione AW Database RL dovranno inoltre essere resi conformi alle modalità di conteggio del “meter” installato sui computer dei panelisti (per es. eliminazione delle attività svolte su finestre diverse da quella “in focus”). Tale attività di depurazione dovrà essere svolta applicando nei sistemi di Web Analytics originatori dei dati censuari ulteriori filtri atti a non conteggiare nei dati censuari da usare per la ponderazione gli elementi di pagina vista non rilevati dai “meter” installati in AW Panel.

Considerando le metriche così depurate di Pagine Viste in una situazione ideale dovrebbe esserci una quasi identità di risultati fra i valori ottenuti da AW Census depurato e quelli ottenuti da AW Panel.

I dati di AW Panel derivano ad ogni modo da stime campionarie, affette perciò da un errore statistico causato dalla dimensione ridotta del campione stesso e da un errore residuo di distorsione qualora le procedure di ponderazione già applicate non abbiano corretto in modo completo il profilo socio-demografico e di consumo di internet del panel rispetto all'universo rappresentato.

La disponibilità di alcuni indicatori censuari (perciò non affetti dagli errori sopra illustrati) può essere perciò utilizzata come "ulteriori valori di Universo" da impiegare nella procedura di ponderazione del panel onde ribilanciare eventuali residui disallineamenti rispetto all'Universo rappresentato.

Viene perciò effettuata un'ulteriore ponderazione aggiungendo ai parametri già utilizzati anche alcuni indicatori di fonte censuaria.

La ponderazione su base censuaria verrà effettuata utilizzando la metrica di pagine viste proveniente dall'AW Census in modo da garantire l'omogeneità rispetto alla stessa metrica rilevata dal meter. In particolare verranno utilizzate solo misurazioni che il meter è in grado rilevare e che trovano riscontro anche nella fonte censuaria. A titolo di esempio non verranno considerate le pagine viste prodotte dall'autorefresh.

La nuova ponderazione permette di allineare il campione AW Database RL (Home+Work + Altri Luoghi Probabilistico+Online) rispetto ai seguenti parametri descrittivi dell'Universo degli utilizzatori Internet 2-74 anni:

sexo/età/ area geografica/dimensione aziendale/settore merceologico aziendale/classi di volume di navigazione/classi di tipologie dei siti visitati /pagine viste derivate da AW Census.

La ponderazione avviene utilizzando la metodologia CALMAR di INSEE (Istituto Statistico Nazionale Francese). CALMAR appartiene alla stessa classe di IPF (Iterative Proportional Fitting) già utilizzato nella prima ponderazione applicata ad AW Panel , ampliando però la gamma delle variabili di controllo da quelle puramente categoriche come il sesso da Ricerca di Base a variabili puramente quantitative come il numero di pagine viste.

Le principali metriche da AW Census sono calcolate secondo la seguente procedura:

- Vengono considerati i soli valori di AW Census relativi a siti che utilizzano strumenti di misurazione censuaria certificati da Audiweb e che si siano sottoposti al livello massimo dei controlli stabilito da Audiweb
- Per ognuno di questi siti viene calcolato da AW Census il valore totale di pagine viste, depurandolo delle componenti non rappresentate in AW Database RL (navigazione da devices mobili, PDA/cellulari e al di fuori del territorio italiano) e reso coerente con le modalità di misurazione degli elementi pagine viste utilizzate dal meter (es. eliminazione delle attività svolte su finestre diverse da quella "in focus, come nel caso delle pagine viste in autorefresh").
- Per ogni Categoria del Catalogo AW per la quale siano disponibili da AW Census almeno 5 siti/canali secondo quanto sopra esposto, viene calcolato il valore totale di pagine viste al netto del sito che presenta le peggiori discrepanze assolute (in più o in meno) fra i propri valori da AW Census e da AW Database RL.
- I siti/canali rientranti in categorie per le quali non viene raggiunto il numero minimo di 5 siti sono raggruppati nella Categoria residuale Altro e anche per essi viene calcolato il valore totale di pagine viste al netto del sito che presenta le peggiori discrepanze assolute (in più o in meno) fra i propri valori da AW Census e da AW Database RL.

- I valori così calcolati di pagine viste per tutte le Categorie nelle quali è stato raggiunto il numero minimo di 5 siti più la Categoria residuale Altro sono inseriti nella procedura CALMAR come valori target, in parallelo agli altri target fonte Ricerca di Base/Campioni Probabilistici già esposti. La procedura CALMAR pondererà AW Database RL in modo che i risultati prodotti siano allineati con tutti i valori target.
- Analoga procedura è svolta considerando per l'insieme dei siti sopra esposti i totali pagine viste suddivisi fra sabato, domenica e giorni feriali.
- Analoga procedura è svolta considerando per l'insieme dei siti sopra esposti i totali pagine viste suddivisi per fasce orarie della giornata, che saranno: 0-3, 3-6, 6-9, 9-12, 12-15, 15-18, 18-21, 21-24
- Il panel risultante dalle operazioni di Data Fusion, Estensione/Clonazione ALNE e ALE viene riponderato. Tale riponderazione viene effettuata utilizzando le stesse variabili già impiegate per la ponderazione nell'AW View oltre a quelle relative al confronto tra le pagine viste censuarie e quelle di fonte panel come da paragrafo 5.8.
- Qualora l'indice di efficienza della seconda ponderazione dovesse diminuire in modo eccessivo rispetto all'indice di efficienza della prima ponderazione si procede all'allentamento delle variabili derivate dalle pagine viste fino a quando l'efficienza risulti inferiore in valore assoluto di non più del 15% del valore originario.
- L'allentamento avviene accettando una limitata divergenza fra i valori target imposti e i valori effettivamente ottenuti dalla ponderazione. In parallelo in questo caso saranno attivate le operazioni di verifica e manutenzione della composizione del campione atte ad aumentare gli indicatori di efficienza.

5.9 L'inclusione delle stime da AW Database RL in AV View

I passi illustrati di Fusione/Estensione/Clonazione/Convergenza portano al fatto che i risultati prodotti per le varie metriche e per i vari siti da AW Database RL siano differenti (e stante la logica della procedura esposta, maggiormente rappresentativi) dagli analoghi risultati producibili da AW Panel prima di questa ultima fase. Ne conseguirà la coesistenza di 2 differenti risultati.

5.10 Il periodo di rilevazione

Continuativo su base mensile.

5.11 Indirizzo Web di pubblicazione della metodologia

www.audiweb.it

ALLEGATO Note Informativa A

ESEMPLIFICAZIONE DELLE MODALITA' DI CALCOLO DELLE NUMEROSITA' CAMPIONARIE MINIME

Consideriamo gli "Universi 2+" rappresentati a Dicembre 2007 secondo i dati Nielsen Online.

Risulta un **Universo Home di 30.269.000** (individui con accesso a Internet da casa, di cui solo una parte attivi nel mese) e un **Universo Work di 9.531.000** (individui con accesso a Internet dal lavoro, di cui solo una parte attivi nel mese).

La somma dei due universi risulta perciò pari a 39.800.000 al lordo però delle duplicazioni.

Considerando le seguenti regole riportate in AW_PANEL

*"La numerosità totale del panel sarà di **20.000 individui**, con il vincolo che le proporzioni delle componenti "probabilistica" e "online" saranno rispettivamente il 22,5% (4.500 individui) ed il 77,5% (15.500 individui).*

Considerando i panel HOME_PROB e WORK_PROB sarà mensilmente tenuto sotto controllo il criterio fornito dal cosiddetto "fattore di proiezione medio" (con riferimento alle numerosità di utenti in internet, come risultanti dalla Ricerca di Base, da rappresentare con ogni panelista). Qualora la differenza tra i fattori dei 2 panel fosse superiore al 30%, si attuerà un intervento teso al riequilibrio di detto parametro incrementando adeguatamente la numerosità campionaria del campione che presenta il più alto valore del "fattore di proiezione medio .

Analogha operazione sarà effettuata con riferimento ai campioni HOME_ONLINE e WORK_ONLINE."

Si arriva alle seguenti dimensioni numeriche delle 4 componenti di AW PANEL

CAMPIONI PROBABILISTICI HOME_PROB E WORK_PROB

Dimensione ideale:

Totale = 4.500 individui , peso medio $39.800.000 / 4.500 = 8.844$

Numerosità ideale HOME_PROB = $30.269.000 / 8.844 = 3.422$

Numerosità ideale WORK_PROB = $9.531.000 / 8.844 = 1078$

Dimensione accettabile (ipotizzando una minor dimensione di Work_prob rispetto a quella ideale):

Numerosità accettabile HOME_PROB = $30.269.000 / 8.362 = 3.620$

Numerosità accettabile WORK_PROB = $9.531.000 / 10.831 = 880$

In questa ipotesi il peso di proiezione Work risulta del 30% superiore a quello Home

CAMPIONI CON RECLUTAMENTO ONLINE HOME_PROB E WORK_PROB

Dimensione ideale:

Totale = 15.500 individui, peso medio 39.800.000 / 15.500 = 2.567,7

Numerosità ideale HOME_ONLINE = 30.269.000 / 2.567,7 = **11.788**

Numerosità ideale WORK_ONLINE = 9.531.000 / 2.567,7 = **3.712**

Dimensione accettabile (ipotizzando una minor dimensione di Work_online rispetto a quella ideale):

Numerosità accettabile HOME_ONLINE = 30.269.000 / 2.425 = **12.480**

Numerosità accettabile WORK_ONLINE = 9.531.000 / 3.156 = **3.020**

In questa ipotesi il peso di proiezione Work risulta del 30% superiore a quello Home.

Il campione WORK PROB avrà perciò una dimensione ideale di 1.078 e minima accettabile di 880 casi.

Il campione WORK ONLINE avrà perciò una dimensione ideale di 3.720 e minima accettabile di 3.020 casi.

Il campione HOME PROB avrà perciò una dimensione ideale di 3.422 casi

Il campione HOME ONLINE avrà perciò una dimensione ideale di 11.788 casi

Qualora i campioni WORK si discostino dalla dimensione ideale scendendo sino ai valori della dimensione accettabile , i campioni Home dovranno compensare la numerosita' di casi mancanti al totale di 20.000 casi.

Considerando inoltre la ulteriore regola riportata in AW_PANEL

Saranno inoltre sottoposti a controllo sistematico gli elementi di stratificazione del panel WORK_PROB, (sesso / classi di età / area geografica / settore merceologico e dimensione dell' azienda) e del panel HOME_PROB (sesso / classi di età / area geografica).

Ciascun panel deve rispettare il mix di allocazione campionaria derivante da AW Ricerca di Base rispetto alle categorie delle variabili sopra esposte. Per la numerosità campionaria di ciascuna categoria e' accettata una variabilità massima del 30 % rispetto al valore teorico desunto da AW Ricerca di Base .

Si hanno inoltre le seguenti restrizioni.

Se ad esempio la percentuale di Maschi nell'Universo Home di 30.269.000 fosse del 50%, il campione Home_Prob dovrà avere una percentuale di maschi compresa fra il 65% e il 35% .

Analoghi limiti saranno imposti alle varie categorie delle altre variabili sopra esposte e cioè nel panel WORK_PROB, sesso / classi di età / area geografica / settore merceologico e dimensione dell' azienda e nel panel HOME_PROB sesso / classi di età / area geografica.

Qualora i valori degli Universi Rappresentati differissero in futuro da quelli utilizzati nella presente esemplificazione, anche tutte le numerosità campionarie dovranno essere adeguate alla nuova situazione.

Si tenga presente che tutte le numerosita' riportate si riferiscono al numero di casi "in tab" cioè al numero di panelisti inclusi nell'elaborazione delle tavole statistiche mensili e non considerano invece panelisti esclusi dalla base di calcolo delle statistiche mensili in quanto reclutati nel corso del mese , non piu' partecipanti al panel alla fine del mese o per altri motivi.

ALLEGATO Note

Informativa B “Struttura di AW Database RL”

AW Database RL si compone delle seguenti strutture di dati che saranno rese disponibili agli Utilizzatori con cadenza "mensile" e con riferimento alle navigazioni effettuate nei singoli "mesi" dell'anno. I dati saranno disponibili come file Ascii , prelevabili da un sito PFT.

Per "mese" si intende quanto specificato al punto 5.10 “Il Periodo di rilevazione”.

Demographics Data

- Fields are separated by a « ; »
- Files are in ASCII format and contain a return (CRLF) at the end of each record

- Individual Data File (*Sample_MMYYYY.dat*)

Field	Criteria	Comment
1	Household_id	Household Id
2	Household_member	Individual Id within the household
3	Weight	Individual's weight
4	Criteria 1	Value or answer of criteria 1 of the individual (pl. refer to descriptive file)
...
3+n-1	Criteria n-1	Value or answer of criteria n-1 of the individual (pl. refer to descriptive file)
3+n	Criteria n	Value or answer of criteria n of the individual (pl. refer to descriptive file)

- Descriptive file of criteria and answers (*SampleDesc_MMYYYY.txt*) present within the individual data file (as from the field number 4 of this same file)

Field	Type	Name	Description
1	Numerical	Range	Number of the criteria (from 1 to n)
2	String of characters	Question_Label	Question's label associated with the criteria
3	1 or 2	Type	Type of criteria (1=Qualitative, 2=Quantitative)
4	Empty or Numerical	Value_id	If Type=1, ID of the answer, if not, empty
5	Empty or Numerical	Value_name	If Type=1, label of the answer, if not ,empty

N.B. : This file is thus composed of *L* lines, 1 per criteria of type 2 and as many lines as answers for criterias of type 1.

Example :

1;Sex;1;1;Man
1;Sex;1;2;Woman
2;Age;1;1;15-24 years

2;Age;1;2;25-34 years
2;Age;1;3;35 years and more
3;Age detailed;2;;

Navigation data

- Fields are separated by a « ; »
- Files are in ASCII format and contain a return at the end of each record

Important : Fields *Brand_id*, *Channel_id* and *IA* must be identical from one delivery to another. Any new id then corresponds to a new vehicle (Brand, Rollup, Channel or Internet Software <= example: MSN Windows Live) and an id already used can only be used for the very same vehicle.

Same for *Category_ID* and *Sub_Cat_ID*.

Note : Custom Rollup Up are considered as specific brands

- **Individual audience Data file (*Audience_MMYYYY.dat*)**

Field	Type	Name	Description
1	Numerical	Household_id	ID of Household
2	Numerical	Household_member	ID of individual with the household
3	Date (jj/mm/aaaa)	Day	Date
4	Numerical	Brand_id	ID of Brand
5	Numerical	Channel_id	ID of Channel or 0
6	0 or 1	Tree_level	0 if Brand_id is a Brand 1 si Brand_id is a Rollup
7	Numerical	IA	Type of Internet Application
8.1a	Numerical	Pages , Time Band 1	Number of pages viewed inside Brand /Channel in Time Band 1
8.1b	Numerical	Time , Time Band 1	Times in seconds of navigation inside Brand /Channel in Time Band 1
8.1c	Numerical	Sessions , Time Band 1	Number of sessions inside Brand /Channel in Time Band 1
8.2a	Numerical	Pages , Time Band 2	Number of pages viewed inside Brand /Channel in Time Band 2
8.2b	Numerical	Time , Time Band 2	Times in seconds of navigation inside Brand /Channel in Time Band 2
8.2c	Numerical	Sessions , Time Band 2	Number of sessions inside Brand /Channel in Time Band 2
		Repeated for time bands 3 to 7	
8.8a	Numerical	Pages , Time Band 8	Number of pages viewed inside Brand /Channel in Time Band 8
8.8b	Numerical	Time , Time Band 8	Times in seconds of navigation inside Brand /Channel in Time Band 8
8.8c	Numerical	Sessions , Time Band 8	Number of sessions inside Brand /Channel in Time Band 8
Fields 8 are repeated for 8 time bands , each one of 3 hours (0-2,3-5,6-8,9-11,12-14,15-17,18-20,21-23)			

- **BRANDS File – File *BrandNames_MMYYYY.txt***

Field	Type	Name	Description
1	Numerical	Brand_ID	unique numerical brand ID
2	String of characters	Brand_Name	brand name
3	Numerical	Category_ID	brand category ID
4	Numerical	Sub_Cat_ID	brand sub-category ID

- **CHANNELS file– File *ChannelNames_MMYYYY.txt***

Field	Type	Name	Description
1	Numerical	Channel_ID	unique numeric channel ID
2	String of characters	Channel_Name	channel name
3	Numerical	Category_ID	channel category ID
4	Numerical	Sub_Cat_ID	channel sub-category ID

- **APPLICATIONS file – File *ApplicationNames_MMYYYY.txt***

Field	Type	Name	Description
1	Numerical	IA	unique numeric Internet Software ID
2	String of characters	Application_Name	Name of the Internet Software

- **Hierarchy file CATEGORIES and SUB-CATEGORY – File *CategoryTree_MMYYYY.txt***

Field	Type	Name	Description
1	Numerical	Category_ID	category ID
2	Numerical	Sub_Cat_ID	sub-category ID

- **CATEGORIES file– File *CategoryNames_MMYYYY.txt***

Field	Type	Name	Description
1	Numerical	Category_ID	unique numeric category ID
2	String of characters	Category_Name	site/domain category description.

- **SUB-CATEGORIES file – File *SubCategoryNames_MMYYYY.txt***

Field	Type	Name	Description
1	Numerical	Sub_Category_ID	unique numeric sub-category ID
2	String of characters	Sub_Category_Name	site/domain sub-category description

ALLEGATO Note Informative C “Modalità di utilizzo dei dati censuari”

Il sistema Audiweb prevede l'impiego dei dati da fonte AW Census per due obiettivi:

- Pubblicazione dei dati giornalieri relativi ai siti/canali dei publisher iscritti ad Audiweb che utilizzano la fonte censuaria..
- Utilizzo dei dati censuari dei publisher iscritti per la ponderazione del panel di individui utilizzato per la produzione di AW Database RL.

1. Pubblicazione dati censuari giornalieri

Nel report AW Report i dati censuari sono pubblicati, in modalità automatica con frequenza giornaliera, alle ore 10 del mattino seguente il giorno di rilevazione.

Le informazioni riportate sono relative ai publisher iscritti al report censuario, e in particolare alle seguenti metriche:

- Browser unici
- Pagine viste
- Pagine viste (di cui del totale) che sono prodotte da un autorefresh inferiore a 5 minuti (300”)
- Visite
- Tempo medio per pagina (minuti e secondi)
- Durata visita (minuti e secondi)

E' cura del publisher evidenziare le pagine prodotte da un autorefresh inferiore a 5 minuti (300”) introducendo nel codice della stessa pagina una informazione del tipo di quella sotto riportata:

```
<meta http-equiv="Refresh" content="150; URL=?refresh" />
```

La struttura della pubblicazione prevede, oltre alle metriche suddette, le seguenti informazioni:

- Aggregatore / Custom Property
- Parent
- Brand
- Channel
- Sistema di misurazione browser-based contenente il fornitore censuario utilizzato
- Note

2. Modalità automatica di pubblicazione dei dati giornalieri

E' stato predisposto un flusso di dati provenienti dai provider censuari che permette al sistema di ricezione di rilevare gli stessi e di inserirli nell'interfaccia di pubblicazione.

I provider certificati sono visibili sul sito www.audiweb.it nell'apposita sezione.

E' importante, ai fini di una corretta lettura dei dati prodotti, garantire da parte del publisher la corrispondenza delle informazioni che concorrono a formare le varie aggregazioni di sito e canale.

In particolare la rilevazione censuaria delle pagine di un sito/canale deve essere relativa allo stesso insieme di pagine contenute nel sistema di catalogazione di AW View. Per questo il publisher deve dare tutte le indicazioni necessarie al proprio provider censuario in modo da far costruire, qualora non già disponibili, tutti i report corrispondenti.

3. Dati censuari per la ponderazione del panel

Il panel di individui utilizzato per produrre l'AW Database RL viene ponderato sulla base dei pesi di ciascuno dei profili degli individui che hanno accesso ad internet. Tali pesi vengono dedotti dalla AW Ricerca di base. A valle della Data Fusion Home & Work e della estensione e clonazione (ALNE & ALE) degli individui del panel si ottiene un primo output di informazioni di navigazione in un formato del tipo giorno/fascia oraria. Le fasce orarie previste sono 8 e cioè:

1. 00.00 – 03.00
2. 03.00 – 06.00
3. 06.00 – 09.00
4. 09.00 – 12.00
5. 12.00 – 15.00
6. 15.00 – 18.00
7. 18.00 – 21.00
8. 21.00 – 00.00

A valle di tale output si procede ad effettuare una comparazione, in termini di pagine viste, tra le informazioni del primo database ottenuto (nel formato suddetto) e le corrispondenti informazioni disponibili dalle fonti censuarie.

Affinchè il confronto sia omogeneo è indispensabile che i report censuari siano allineati alla fonte panel (stesse pagine nei siti e canali) e che le pagine viste da fonte censuaria siano nettizzate rispetto ad ogni autorefresh, alle pagine provenienti da accessi stranieri (ip stranieri) e alle pagine provenienti da utenti mobile. Anche in questo caso il publisher deve evidenziare le pagine in autorefresh (tutte) inserendo il codice sopra descritto e indicandole al provider censuario.

Il provider censuario deve creare dei report nettizzati che siano disponibili nel formato giorno/fascia oraria.

La scelta delle pagine viste come metrica di confronto tra panel e censuario nasce dal fatto che le altre metriche non permettono un confronto omogeneo per effetto delle pagine non in focus. Infatti il panel non rileva il tempo speso quando la pagina di un sito è off focus (ad esempio quando si sta lavorando su applicativi locali sul PC o su altre pagine web mentre è ancora attiva in background la pagina in oggetto di rilevazione).

Successivamente si procederà ad utilizzare le aggregazioni disponibili, sulla base di categorie (vedi esempio successivo) di appartenenza.

Il confronto sarà quindi effettuato prendendo il totale delle pagine viste di ciascuna frazione temporale, cioè tipo giorno/fascia oraria provenienti dalle 2 fonti (panel e censuaria).

Esempio: il raggruppamento Automotive > Automotive Information è costituito da:

Virgilio > Auto & Motorbox

Tiscali > Automotive

La Repubblica > La Repubblica Motori

Corriere della Sera > Automobili.com

Le pagine viste di queste 4 componenti vengono confrontate sulle 2 fonti e relativamente alla tipologia giorno feriali, cioè tutti i giorni feriali del mese in osservazione e sulla fascia oraria.

Esempio: nel mese di febbraio 2008, per i giorni feriali e per la fascia oraria 09.00 – 12.00 e per l'aggregazione automotive verranno sommate le pagine dei giorni 1, dal 4 all'8, dall'11 al 15, dal 18 al 22 e dal 25 al 29. Tale somma si riferisce ai contributi complessivi dei siti/canali sopra evidenziati.

Lato panel viene effettuata la stessa operazione (ovviamente a valle dell'espansione da panel già disponibile) estraendo, per gli stessi siti/canali le pagine viste sommate riferite alla tipologia giorno feriale/fascia oraria (09.00 – 12.00).

Dal confronto dei 2 valori emergeranno differenze che saranno utilizzate come elementi di input ad un sistema statistico di IPF (con algoritmo Calmar) il quale produrrà i fattori di ulteriore riponderazione sui panelisti con conseguenti nuovi output in termini di audience sui singoli siti/canali e sui profili.

Esempio di aggregazioni su categorie

Category	Sub-category
Automotive	Automotive Information
Education & Careers	Career Development
Entertainment	Broadcast Media
Entertainment	Multi-category Entertainment
Entertainment	Sports
Entertainment	Videos/Movies
Multi-category Commerce	Classifieds/Auctions
News & Information	Current Events & Global News
News & Information	Directories/Local Guides
News & Information	Multi-category News & Information
Search Engines/Portals & Communities	General Interest Portals & Communities
Search Engines/Portals & Communities	Member Communities
Search Engines/Portals & Communities	Search
Telecom/Internet Services	E-mail
Travel	Multi-category Travel