

Il software per l'analisi qualitativa TRANSANA

Fabio Malfatti, *Centro Ricerche EtnoAntropologiche, Siena*

02/06/2007



(Dettagli al termine del documento)

La storia

La storia di Transana inizia nel 1995 negli USA, più precisamente all'Università del Wisconsin. Cris Fressnacht, ai tempi, era uno studente di sociologia che stava scrivendo una tesi basata sull'analisi di un gran numero di interviste registrate. Come accade spesso, invece di andare al cuore del problema, ossia scrivere la tesi, iniziò a girarci intorno scrivendo un programma che gli consentisse di analizzare le interviste. Ad un certo momento si rese conto che non sarebbe riuscito a terminare la tesi e finire il programma. Con saggia decisione terminò la tesi e donò il codice del programma all'università a condizione che venisse completato e distribuito gratuitamente.

Nel 2000 il Wisconsin Center for Education Research¹, incaricò David Woods di terminare il lavoro di sviluppo del programma. David rielaborò parti significative e aggiunse una serie di funzioni in modo da renderlo più flessibile ed adattabile. Finalmente nell'ottobre del 2001 venne rilasciata la prima distribuzione, la 1.0

La versione originale del programma Transana era scritta utilizzando come linguaggio di programmazione il Delphi², di proprietà della Microsoft, cosa che avrebbe limitato il funzionamento solamente al sistema operativo Windows e non rendeva Transana completamente conforme alle specifiche per la distribuzione open-source, contenendo almeno un componente commerciale. Nel 2003 David Woods ha iniziato a riscrivere Transana in Python, un linguaggio di programmazione open-source e multiplatforma. La nuova versione 2.0 è stata terminata agli inizi del 2005. I codici sorgente di Transana sono disponibili sul sito del progetto per chi voglia effettuare modifiche o semplicemente studiarne la struttura. Nel gruppo di sviluppo, oltre a David Woods, occorre ricordare gli importanti apporti di Jonathan Beavers, Nate Case, Mark Kim, Rajas Sambhare e David Mandelin. (Woods, 2006)³

Negli anni, attorno al progetto Transana, anche grazie ai forum di discussione presenti sul sito www.transana.org, si è aggregata una nutrita comunità di ricercatori provenienti da differenti settori disciplinari, accomunati dall'interesse per lo strumento e dalla filosofia del progetto. Comunità che ha avuto parte attiva nella realizzazione della versione 2.0 e che continua ad essere uno dei motori della evoluzione del programma.

A causa dei crescenti costi di sviluppo e delle scarse donazioni da parte degli utilizzatori (a fine 2006 erano state scaricate 40.000 copie di Transana, ma solo pochi avevano versato il libero contributo che veniva richiesto), i download dalla versione 2.20, che include nuovi e importanti aggiornamenti, è a pagamento. Il costo resta comunque molto limitato: 50 US\$ (circa 37 €) per la versione singolo utente e 500 US\$ (circa 370 €) per la versione multiutente, (in questo caso la quota è intesa unica per tutte le copie necessarie al progetto di ricerca). La distribuzione, viene mantenuta sotto licenza GPU-GPL, per cui continueranno ad essere disponibili i codici sorgente. E' importante però sottolineare che, considerato il basso costo, è opportuno incoraggiare ogni utilizzatore della versione singolo utente a contribuire, in modo da renderne possibile il continuo aggiornamento.

Oltre al pagamento diretto, lo sviluppo del software può essere appoggiato richiedendo al gruppo di lavoro che sviluppa Transana modifiche, personalizzazioni o applicazioni specifiche. La

¹ Wisconsin Center for Education Research <http://www.wcer.wisc.edu>

² **Delphi** è sia un linguaggio di programmazione che un ambiente di sviluppo. È stato creato dalla Borland (che per un certo periodo ha cambiato il suo nome in Inprise). Il linguaggio Delphi, precedentemente conosciuto come Object Pascal (una versione di Pascal orientata agli oggetti), è stato sviluppato inizialmente per Microsoft Windows, di recente sono state pubblicate versioni per GNU/Linux (Kylix) e per il framework .NET di Microsoft (<http://it.wikipedia.org/wiki/Delphi>)

³ Woods, D. 2006, 'Transana history', comunicazione personale

realizzazione di queste modifiche è subordinata ad una lista di priorità o al finanziamento diretto della modifica.

Caratteristiche principali

Il software è stato progettato sulle necessità dei ricercatori che vogliono analizzare dati video o audio in forma digitale, per consentire di realizzare analisi e gestire i dati in modo molto sofisticato.

Il punto focale del programma sta nella continua e costante aderenza, in tutte le fasi di lavoro, tra testo, trascrizioni, categorie utilizzate e sintagmi audio video delle registrazioni originali. Sfruttando le potenzialità dei computer, il programma offre la possibilità di utilizzare un solo strumento in tutte le fasi della ricerca, offrendo strumenti per la trascrizione, l'annotazione, l'assegnazione di parole chiave e strumenti per l'interrogazione del data base.

Transana è stato realizzato in due versioni: multiutente e singolo utente. Le funzioni sono essenzialmente le stesse, ma il data base della versione multiutente risiede su di un server, permettendo ad un gruppo anche numeroso di ricercatori di lavorare contemporaneamente via internet o rete locale ad un unico o più progetti e di comunicare tra loro.

Lo spazio disco necessario per lavorare è limitato praticamente alla dimensione dei file audio e video dell'archivio, che non vengono modificati né alterati, mentre i dati (riferimenti ai sintagmi,

descrizioni, parole chiave ecc.) vengono archiviati in un data base di tipo SQL che può essere esportato in formato XML.

Transana può lavorare con file codificati nei formati comunemente disponibili, compresi windows media video e audio (wmv e wma), mpeg1, mpeg2, mpeg 4, quick time, mp3 ed altri. E' possibile utilizzare anche file in formato DV ma date le dimensioni dell'archivio (12 Gb per ogni ora di registrazione), è sconsigliato. Transana è disponibile sia per piattaforme windows che Mac OSX (è allo studio una versione Linux).

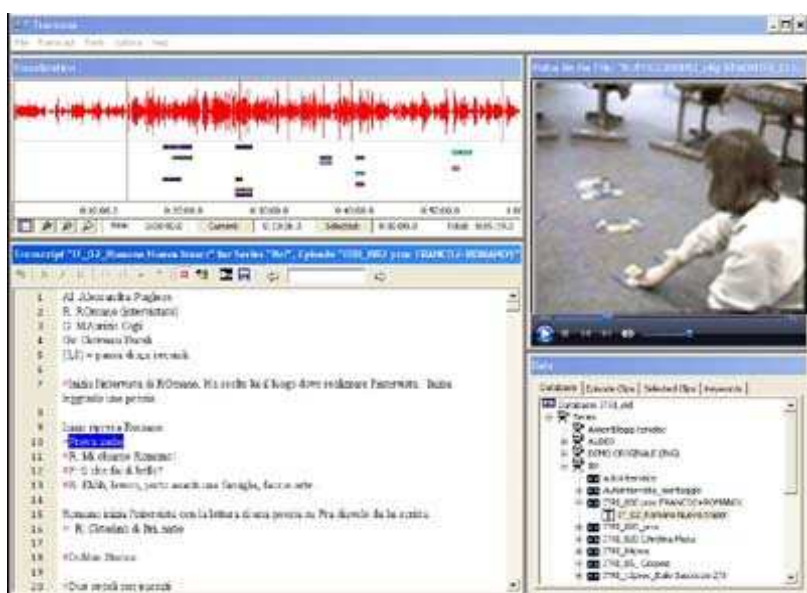


Figura 1

Gli strumenti per l'analisi, comprendono la creazione di sintagmi chiamati *clip*, alle quali assegnare parole chiave. Le clip sono contenute in *collezioni* che possono essere riprodotte in sequenza, simulando un assemblaggio. In una collezione possono essere inserite clip provenienti da qualunque file audio/video (*episodio*) inserito nel data base. E' importante sottolineare che la creazione di clip 'autonome' non comporta la necessità di ulteriore spazio su disco rispetto a quello occupato dall'archivio, infatti il programma archivia nel suo data base solamente i riferimenti al video originale. Le parole chiave assegnate possono essere esplorate attraverso vari strumenti grafici e testuali.

In tutte le fasi dell'analisi viene mantenuta una stretta 'aderenza' tra testo, trascrizioni, categorie utilizzate e sintagmi audio video. E soprattutto di condividere le analisi con altri ricercatori.

In particolare Transana consente di:

- Identificare e recuperare facilmente i sintagmi significativi di audio e video.
- Gestire grandi archivi audio/video contenenti centinaia (e potenzialmente migliaia) di ore di registrazione.
- Organizzare sintagmi video (video clip), provenienti dallo stesso o da differenti files audio/video, in categorie significanti, come strategia per sviluppare ed espandere la comprensione teorica dei contenuti.

- Applicare parole chiave alle clip
- Visualizzare e stampare rapporti sull'uso delle parole chiave in modalità grafica e testuale.
- Realizzare complesse ricerche di dati e verifica di ipotesi teoriche su grandi archivi di registrazioni audio/video
- Condividere i dati dell'analisi con colleghi di un gruppo di lavoro

Metodologie di analisi applicabili

Transana, utilizza come base di partenza per l'analisi, un documento di testo. E' quindi direttamente applicabile a tutte le analisi basate sulla trascrizione o sulla annotazione, come l'Analisi Conversazionale (Conversation Analysis), ma anche su analisi di tipo etnografico, di cinesica, prossemica ecc. Il collegamento tra testo e punto della registrazione inserito manualmente con una precisione di 1/1000 di secondo.

Il testo di riferimento non deve essere necessariamente una trascrizione, ma può consistere in una annotazione, schedatura o altro, per cui si estende a tutte le tipologie di analisi che utilizzano la scrittura come riferimento.

Gli strumenti per gestire la riproduzione dei video/audio permettono di effettuare le normali operazioni di controllo della riproduzione (avanzamento veloce, pausa, riproduzione, riavvolgimento, ripresa della riproduzione 2 secondi indietro o programmabile direttamente dall'utente), con una combinazione di tasti o con un pedale USB programmabile. E' possibile spostarsi all'interno della registrazione utilizzando il cursore del riproduttore, la forma d'onda, il grafico delle parole chiave ed il testo e riprodurre la registrazione a velocità accelerata o rallentata.

Nel programma sono stati integrati i principali codici della notazione Jefferson, che possono essere introdotti con una combinazione dei tasti o mediante l'utilizzo di apposite icone.

La creazione e la manipolazione di sintagmi (clip) è l'equivalente elettronico della tradizionale operazione di 'taglio' dei documenti di testo (trascrizioni o altro) per suddividerli in sintagmi analiticamente significativi, che venivano trasferiti su schede successivamente organizzate per tematiche. Transana consente al ricercatore di ricollocare rapidamente le 'schede' nel loro contesto originale del file audio/video, ma soprattutto di non perdere mai, in ogni istante del processo di analisi, il collegamento tra testo e sintagma audio/video riproducibile a piacere.

E' possibile applicare molteplici parole chiave alle clip create e realizzare la ricerca utilizzando operatori logici (and, or, not). In questo modo vengono facilitate le operazioni di ricerca e di verifica delle ipotesi, permettendo di muoversi rapidamente all'interno della singola registrazione o trasversalmente con la visualizzazione grafica delle parole chiave sotto forma di colori sulla scala temporale.

Transana copre un vuoto nell'ambito dei software analitici. Molti software per l'analisi qualitativa, oltre ad essere costosi, sono progettati per lavorare con sintagmi estratti dalle registrazioni audio/video e non con grandi archivi di video o audio grezzi. Ad un unico archivio audio/video possono fare riferimento molteplici processi di analisi realizzato contemporaneamente o in tempi diversi.

Caratteristiche principali:

- Sistema avanzato per la trascrizione e annotazione dei documenti audiovisivi
- Visualizzazione contemporanea di forma d'onda, testo, contenuto video
- Collegamento definibile dall'utente tra testo e sintagma audio/video di riferimento
- Navigazione nella registrazione visualizzata utilizzando il testo, la forma d'onda o il media player integrato
- Possibilità di creare molteplici gruppi di parole chiave
- Sistema di ricerca per parole chiave sull'intero archivio
- Creazione di vari data base indipendenti
- Assegnazione di parole chiave a tutti gli oggetti (episodi, trascrizioni, clip)
- Possibilità di creare *collezioni*, sequenze di sintagmi selezionati (clip) senza dover duplicare sintagmi video/audio dell'archivio originale

- Le collezioni possono essere create sia manualmente che attraverso ricerca per parole chiave e operatori logici (AND, OR, NOT)
- Può essere abbinato ad un sistema SRB (Storage Resource Broker) che consente di gestire i trasferimenti dei file audiovisivi via internet in modo efficiente e criptato di per garantire la privacy di dati sensibili
- Nella versione multiutente, vari ricercatori possono operare (collegati via internet o mediante una rete locale) sullo stesso data base di dati⁴, comunicare tra loro e visualizzare il lavoro degli altri in tempo reale.
- Consente di esportare le trascrizioni e annotazioni in formato rtf e di dati delle clip in formato txt.

Le motivazioni della scelta

- Funzioni disponibili
- Costo del software. Gratuito sino al 3 aprile 2007. Dalla versione 2.20 Transana non è più gratuito: il costo è di 50 USD per la versione singolo utente (circa 37 Euro) e 500 USD (circa 370 Euro) per la versione multiutente. In questo caso la licenza è valida per tutti i ricercatori del progetto.
- Software open source, può quindi essere modificato per adattarlo a esigenze particolari di una ricerca (il codice è scritto in Python, come data base utilizza MySQL nella versione multiutente e un embedded MySQL nella versione singolo utente)
- Facilità nell'uso (relativamente ai software di analisi qualitativa).
- Flessibilità nell'applicazione a differenti metodologie di analisi.
- E' stato studiato come software multiplatforma, con possibilità di utilizzo sia su sistemi operativi Windows che Macintosh OSX (è allo studio una versione Linux).
- E' stato sviluppato da un team universitario ed il progetto ha aggregato una discreta comunità internazionale di ricercatori, nelle più diverse discipline. Al momento sono stati realizzati nel mondo 40.000 download del programma.
- E' stato progettato per lavorare su grandi archivi audio/video (potenzialmente limitati solamente dallo spazio disco) con accesso multiutente via rete (versione multiutente)
- Esistono traduzioni del software in Italiano

Punti di debolezza

- Mancanza di uno strumento di gestione del data base (integrazione, importazione, esportazione, selettiva tra data base)
- Assistenza disponibile solo tramite forum in Inglese
- Manuale solo in Inglese (è in corso una traduzione all'italiano)
- Gestione del data base relativamente rigida

Che cosa non è transana:

- Non è un programma per la consultazione di archivi da parte di utenti generici
- Non è un programma di montaggio video/audio
- Al momento consente di effettuare analisi solamente su file di tipo audio e video (è in fase di studio della possibilità di integrare la gestione di immagini)
- Non integra strumenti grafici per la costruzione di mappe concettuali o reti complesse di relazioni tra oggetti

Informazioni e download (in lingua Inglese) : www.transana.org

Fabio Malfatti info@fabiomalfatti.it / fabiomalf@iol.it



Questo articolo viene rilasciato sotto licenza Creative Commons / Science Commons con la formula: Non commerciale - 2.5, per dettagli: <http://creativecommons.org/licenses/by-nc/2.5/it/legalcode>

L'opera può essere di riprodotta, distribuita, esposta e comunicata al pubblico alle seguenti condizioni: deve essere chiaramente attribuita la paternità dell'opera citando l'autore, è gradita informazione via e-mail. Quest'opera non può essere usata per fini commerciali. L'opera può essere tradotta. Ogni volta che l'opera viene usata o distribuita, deve essere fatto secondo i termini di questa licenza, che va comunicata con chiarezza. In ogni caso, possono essere concordati col titolare dei diritti d'autore utilizzi non consentiti da questa licenza. Niente in questa autorizzazione altera o limita i diritti dell'autore.

⁴ Gi archivi audio/video devono essere disponibili localmente (su rete locale o su computer locale) a causa dei limiti strutturali delle comunicazioni internet.