

AILALI' "AUDIO INDEXING" PER LA LETTURA E L'ASCOLTO DELLA LETTERATURA ITALIANA

Piero Cosi*, Cinzia Avesani*
Maurizio Falghera**, Cristiana Giacometti**, Elena Giacometti**,
Gianni Lazzari***, Fabio Pianesi***, Roberto Gretter***, Fabio Tesser***

* ISTC-SPFD CNR

Istituto di Scienze e Tecnologie della Cognizione - Sezione di Padova "Fonetica e Dialettologia", Consiglio Nazionale delle Ricerche

www: <http://nts.csrf.pd.cnr.it/>

e-mail: {cosi,avesani}@csrf.pd.cnr.it

** "IL NARRATORE" Associazione Culturale

www: <http://www.ilnarratore.org>

e-mail: {mauri,cris,elena}@ilnarratore.com

*** ITC-IRST

Istituto Trentino di Cultura – Centro per la Ricerca Scientifica e Tecnologica

www: <http://www.itc.it/IRST/index.htm>

e-mail: {lazzari,pianesi,gretter,tesser}@itc.it

SOMMARIO

Vengono illustrate brevemente le linee guida del progetto AILALI' ("Audio Indexing" per la Lettura e l'Ascolto della Letteratura Italiana), di cui vengono inoltre descritti i principali obiettivi, le principali attività di sviluppo, le possibili applicazioni e le future possibilità di sfruttamento.

DESCRIZIONE SINTETICA

La disponibilità in rete di informazioni multimediali attraverso "biblioteche digitali" realizzate da importanti attori del settore (giornali, radio, televisioni) ha innescato la richiesta di applicazioni avanzate, sia da parte dei produttori di informazione che del pubblico, per un'efficace ricerca ed utilizzazione dell'enorme mole di dati ormai disponibili. E' appunto da queste motivazioni che sorge la necessità di sviluppare nuove tecnologie in grado di migliorare la gestione e l'accesso all'informazione.

La proliferazione di informazioni multimediali nel *World Wide Web* ha fortemente accelerato l'introduzione di motori di ricerca per immagini, filmati e materiali audio. Circa il 60% del materiale audiovisivo disponibile sulla rete contiene parlato. Di circa 12.000 *files* considerati analizzando la comunità virtuale di *Yahoo*, da una nota società di software, ben il 59% contiene parlato, il 40% è rappresentato da musica e l'1% da altre informazioni che non rientrano nelle prime due categorie. Il 90% del parlato è in inglese con circa tutto il resto nelle principali lingue europee.

Da questi dati si evince facilmente che un sistema in grado di indicizzare velocemente ed automaticamente tutto questo materiale risulterebbe di essenziale importanza per un'efficace utilizzazione delle informazioni multimediali disponibili. Proprio in quest'ottica sia l'Istituto Trentino di Cultura, in particolare, il Centro per la Ricerca Scientifica e Tecnologica (ITC-IRST) di Pantè di Povo (Trento) che l'Istituto di Scienze e Tecnologie della Cognizione, in particolare, la Sezione di Padova "Fonetica e Dialettologia" (ISTC-SPFD) (ex Istituto di Fonetica e Dialettologia) di Padova del Consiglio Nazionale delle Ricerche sono ormai da anni fortemente impegnati in questo settore.

Anche l'Associazione Culturale "il Narratore" è da tempo direttamente impegnata a promuovere e diffondere nel mondo l'arte della lettura e dell'ascolto della Letteratura Italiana. A questo scopo utilizza tutti gli strumenti telematici e digitali (compresa la rete Internet) attualmente a disposizione, attraverso la sperimentazione di strumenti nuovi e originali (soprattutto per il miglioramento dell'interattività *on-line*) e il

perfezionamento di canali di comunicazione più tradizionali (audiocassette, CD-audio, CDMP3, radio, performance teatrali, sessioni di lettura e ascolto nelle scuole, presenza in festival, ecc.).

Le attività di questo progetto saranno principalmente incentrate sullo sviluppo di tecnologie mirate alla realizzazione, elaborazione ed indicizzazione di archivi digitali relativi alla Letteratura Italiana focalizzati alla lettura ed all'ascolto. Per la realizzazione di questi archivi saranno, infatti, utilizzati i più innovativi strumenti tecnologici e culturali di accesso alle fonti della cultura letteraria anche in funzione di una loro possibile fruizione da parte di utenti disabili.

Nell'ambito di questo progetto, infatti, saranno realizzati archivi audio di letture e storie registrate in voce da narratori, attori e scrittori professionisti: tali archivi audio (+ testi scritti) digitalizzati saranno messi gratuitamente a disposizione di tutti gli utenti Internet nel formato audio compresso MP3 e nel formato testo PDF. Tutti gli archivi saranno automaticamente indicizzati sul segnale audio in modo da facilitare una loro rapida consultazione. L'utente sarà in grado di formulare le sue "query" in modalità testuale e di ricevere le corrispondenti informazioni sotto forma di collegamenti alle sezioni audio che soddisfino le query. Tutti gli archivi potranno essere facilmente "ascoltati" e consultati in rete sia direttamente, qualora la registrazione audio dell'opera letteraria in questione sia disponibile, sia mediante un sistema di sintesi da testo scritto in grado di leggere il testo corrispondente.

Seguendo una struttura attualmente in fase di definizione nel sito dell'associazione il Narratore (<http://www.ilnarratore.org>) gli archivi saranno suddivisi in tre collezioni:

- Audio Antologia della Letteratura Italiana
- Visioni del Mondo
- Voci della Terra

Due importanti novità di questo progetto riguarderanno senza dubbio l'utilizzazione delle più attuali metodologie per il riconoscimento automatico del parlato al fine di realizzare un'efficace indicizzazione del segnale audio e, parallelamente, l'introduzione delle più recenti innovazioni informatiche nel campo della sintesi vocale da testo scritto (TTS) per favorire la lettura di testi della Lingua Italiana non ancora registrati.

La principale tematica riguarderà l'audio segmentazione o indicizzazione del materiale audio relativo alla lettura di vari brani dei più famosi autori della Letteratura Italiana per un loro efficace ascolto e per una ricerca guidata, senza trascurare però altre attività di interesse quali l'audio segmentazione e classificazione di segnali audio in base alla sorgente sonora (parlato, musica, tipo di musica), l'eventuale identificazione dei narratori e la segmentazione e classificazione del contenuto.

Vi è da sottolineare che in questo progetto non sussiste il problema del riconoscimento "da zero" del materiale audio da indicizzare in quanto, considerando la Letteratura Italiana, il materiale a disposizione è fedelmente letto da alcuni attori e speaker professionisti. Le metodologie di riconoscimento automatico saranno, infatti, applicate non per il riconoscimento del parlato da indicizzare, ma per l'allineamento temporale con il testo corrispondente.

OBIETTIVI E ATTIVITÀ

I principali obiettivi culturali del progetto possono essere di seguito elencati:

- Promuovere la lettura e l'ascolto della migliore letteratura italiana.
- Recuperare e diffondere l'arte della narrativa orale italiana (dai contastorie alle letture ad alta voce) attraverso la collaborazione con cantastorie, *story tellers*, attori, narratori e scrittori.
- Mettere a disposizione di tutti gli utenti internet nuovi strumenti didattici ed educativi, di aggiornamento e studio, per l'apprendimento della lingua italiana e dei suoi dialetti (interattività online, database audio, ipertestualità, *voice-chat*, ecc.).
- Realizzazione su scala internazionale di collaborazioni letterarie, scambio critico d'idee, promozione di una cultura dell'amicizia e della conoscenza fra i popoli mediata dall'arte dello scrivere e del narrare, in pratica la creazione di una comunità web d'ascolto.
- Realizzare contenuti con strumenti tecnologici e culturali sempre più innovativi di accesso alle fonti della cultura letteraria.
- Realizzare contenuti con strumenti tecnologici e culturali sempre più innovativi di accesso alle fonti della cultura letteraria per gli utenti disabili (privi di vista e altre disabilità).

Per quanto riguarda le specifiche attività di sviluppo di AILALI, queste si possono elencare in:

- Sviluppo di un sito web aggiornato in grado di offrire una sempre migliore navigabilità, consultazione e utilizzazione dei prodotti culturali offerti, con particolare attenzione alla sua trasformazione in un sito multilinguistico e predisposto per l'ottimizzazione d'accesso ai disabili della vista e altri disabili per la ricerca di testi e brani audio.
- Realizzazione e continuo ampliamento dell'archivio dell'Audio Antologia della Letteratura Italiana fino a comprendere tutti gli autori più significativi sia classici sia contemporanei, e creare una ricca sottocollezione comprendente gli autori contemporanei esordienti nel mercato editoriale.
- Realizzazione di un archivio specificamente dedicato alla lettura di fiabe e storie per ragazzi in italiano ed eventualmente in altre lingue europee.
- Indicizzazione audio automatica degli archivi sonori.
- Introduzione di procedure automatiche per la lettura da testo scritto di archivi testuali non ancora disponibili in formato audio.
- Miglioramento dell'interattività con gli utenti attraverso strumenti web come forum, mailing list, chat audio, ecc.
- Creazione di una comunità virtuale - perlomeno su scala italiana ed europea - intorno al progetto sull'arte della lettura e dell'ascolto (comprendente singoli individui, gruppi di lettura e d'ascolto, gruppi di scrittori e contastorie, *mailing list*, gruppi di discussione, siti web, ecc.).
- Cooperazione con archivi audio on line nazionali e internazionali.
- Realizzazione di programmi radiofonici per la rete Internet.
- Formazione nelle scuole, eventi culturali sul territorio, convegni, seminari e ricerche centrati sull'arte della lettura ad alta voce e della narrazione in generale.

ATTIVITÀ FUTURE

Relativamente alle principali attività future ricordiamo:

- Ampliamento del sito con altri archivi di Audio Antologia della Letteratura in altre lingue europee oltre all'italiano
- Ampliamento dell'archivio Visioni del Mondo con letture tratte dalla tradizione filosofica, religiosa, scientifica e artistica europea
- Ampliamento dell'archivio Voci della Terra (in particolare con l'inserimento di registrazioni audio della tradizione poetica e del racconto orale-*story telling* di varie culture e linguaggi europei ed extra-europei)

COORDINAMENTO DELLE ATTIVITÀ

Fra le intenzioni dei proponenti, oltre alla realizzazione del prodotto finale (il sito stesso, con le informazioni prodotte nell'ambito di questo progetto), vi è anche quella di favorire la messa a punto di strategie comuni per la memorizzazione e distribuzione del materiale audio preparato e quindi di preparare la strada ad una standardizzazione di tutte le metodologie coinvolte.

Le diverse e complementari competenze richieste ai partecipanti, afferenti a settori e discipline quali fonetica e fonologia, elaborazione vocale, elaborazione, riconoscimento e sintesi della voce, codifica audio, trattamento di informazioni testuali, informatica, sviluppo web, ricerca e redazione letteraria, recitazione, ecc. , conferiscono alla fase di coordinamento un ruolo chiave e fondamentale.

Sarà compito del Coordinatore, Piero Cosi, dell'ISTC-SPFD del CNR, stimolare, verificare, integrare ed omogeneizzare le attività svolte dalle Unità Operative nelle varie fasi del Progetto. Il numero limitato delle Unità partecipanti, la loro localizzazione geografica, la solida e pluriennale collaborazione delle Unità in progetti e comitati nazionali ed internazionali faciliterà certamente il compito di coordinamento.

E' impegno del Coordinatore convocare un incontro tecnico su base trimestrale a rotazione fra le Sedi, per la definizione, programmazione e verifica dell'avanzamento delle attività. Saranno sollecitati incontri più frequenti soprattutto durante la prima fase del progetto in cui verranno messe a punto e coordinate le metodologie comuni da utilizzare poi nella seconda e realizzativa fase.

APPLICAZIONI E POSSIBILITÀ DI SFRUTTAMENTO

Nella comunità scientifica internazionale vi è un notevole interesse verso le nuove tecnologie multimediali. In particolare, è unanimemente riconosciuto che una loro introduzione nelle normali applicazioni di riconoscimento automatico della voce e di sintesi da testo scritto potrebbe portare finalmente ad un'efficace robusta e sempre più *human-like* interazione uomo-macchina.

Mediante l'utilizzazione delle informazioni raccolte, elaborate e distribuite in questo progetto si potranno elaborare nuovi sistemi per l'indicizzazione automatica di materiale vocale, in presenza o meno della corrispondente informazione testuale e si potranno effettuare ricerche specifiche direttamente sul materiale vocale raccolto. Come anche illustrato in Figura 1, alcune delle principali applicazioni fortemente interessate al prodotto finale di questo progetto possono essere:

- Applicazioni per l'Insegnamento quali, ad esempio, l'apprendimento e la diffusione della lingua italiana mediante la lettura e l'ascolto della migliore letteratura italiana.
- Diffusione con nuovi strumenti tecnologici e culturali della cultura letteraria per gli utenti disabili (privi di vista e altre disabilità).
- Applicazioni nei settori dell'analisi e della codifica del segnale vocale.
- Applicazioni nei settori della sintesi automatica della voce.
- Utilizzazione delle recitazioni e delle letture eseguite da narratori professionisti per migliorare la qualità e la naturalezza della sintesi da testo scritto. In altre parole, analisi dell'interpretazione "semantica" – che interessa tutte le modalità paralinguistiche (tono, timbro, pause, accentazioni, ecc.) - che coinvolge naturalmente la grammatica e la sintassi dei testi.
- Applicazioni nei settori del riconoscimento automatico della voce.
- Sviluppo di interfaccia multimediali uomo-macchina.

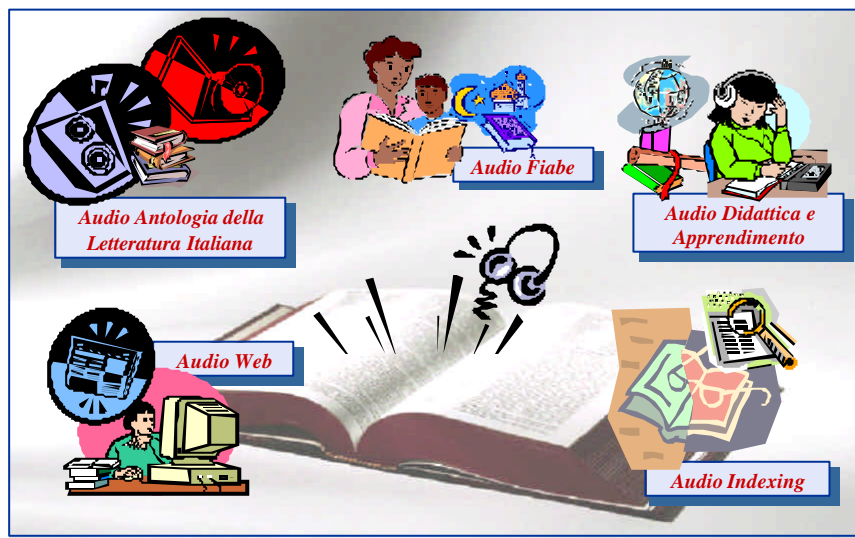


Figura 1. Alcune delle principali applicazioni interessate al prodotto finale di AILALI'

TEMPI, COSTI E RISORSE

AILALI' dovrebbe poter essere completato in 3 anni, nei quali si procederà alla raccolta delle informazioni, alla loro elaborazione ed organizzazione ed infine alla loro distribuzione. Sarà ovviamente necessario acquisire alcune risorse quali:

- Personale tecnico per la realizzazione e il trattamento digitale degli archivi audio.
- Personale tecnico per la realizzazione e gestione del sito internet.
- Diritti d'autore e/o licenze d'uso delle opere realizzate in audio e dei testi on line.
- Strumentazione hardware e software per l'acquisizione e l'elaborazione del segnale audio
- Strumentazione hardware e software per la memorizzazione degli archivi digitali e la loro distribuzione via Internet

I costi previsti per la realizzazione di AILALI' sono di 1.350.000,00 Euro (2.613.964.500 L) che verranno equamente suddivisi in base alle specifiche attività di sviluppo esercitate dalle singole unità operative.

RIFERIMENTI BIBLIOGRAFICI

- [1] S. E. Johnson, P. Jourlin, G. L. Moore, K. Sparck Jones, and P. C. Woodland, "The cambridge university spoken document retrieval system," in Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing, 1999.
- [2] J. Garfalo, E. Vorhees, C. Auzanne, V. Stanford, and B. Lund, "Spoken document retrieval track overview and results," in Proceedings of the 7th Text Retrieval Conference (TREC-7), 1998.
- [3] P. J. Moreno, C. Joerg, JM Van Thong, and O. Glickman, "A recursive algorithm for the forced alignment of very long audio segments," in International Conference on Spoken Language Processing, 1998.
- [4] M. Burrows, "Method for indexing information of a database. u.s. patent 5,745,899," 1998.
- [5] G. Salton and M. J. McGill, " in Introduction to Modern Information Retrieval. McGraw-Hill, 1983.
- [6] B. Eberman, B. Fidler, R. A. Iannucci, C. Joerg, L. Kontothanassis, D. E. Koval-cin, P. J. Moreno, M. J. Swain, and JM Van Thong, "Indexing multimedia for the internet," in In Visual Information and Information Systems. D. P. Huijsmans and Arnold W.M. Smeulders (Eds.) Springer-Verlag, 1999.
- [7] "Linguistic data consortium (LDC)", <http://www ldc.upenn.edu>.
- [8] G. Robson, Inside Captioning, CyberDawg Publishing, Castro Valley, California, 1997.
- [9] C. Silverstein, M. Henzinger, H. Marais, and M. Moricz, "Analysis of a very large altavista query log," in SRC Technical Note 1998-014, 1998.
- [10] G. Salton, Automatic Text Processing: The transformation, Analysis, and Re-trieval of Information by Computer, Addison-Wesley, Reading, Massachusetts, 1989.
- [11] M. Witbrock, "Search engines and beyond," in SearchEngine Meeting, Boston, 1999.
- [12] B. Logan, P. Moreno, JM Van Thong, and E. Whittaker, "An experimental study of an audio indexing system for the web," in International Conference on Spoken Language Processing (ICSLP), 2000.
- [13] A. Singhal and F. Pereira, "Document expansion for speech retrieval," in ACM SIGIR'99, 1999.
- [14] R.Valentino Merletti. Leggere ad alta voce, Mondadori, MI, 1996.
- [15] R.Valentino Merletti. Raccontar storie, Mondadori, MI, 1998.
- [16] D. Pennac, Come un romanzo, Feltrinelli, MI, 2000.