



Un portale semantico per la promozione e l'insegnamento dell'italiano nel mondo

A. Del Bimbo, A. Ferracani, N. Hosseini, M. Bertini

Media Integration and Communication Center

Università degli Studi di Firenze

Firenze, Italy

Il Portale

Il portale VIVIT è stato sviluppato dal 2010, finanziato dai fondi FIRB, dal MICC (Media Integration and Communication Center) e dall'Accademia della Crusca, con contributi delle unità di ricerca dell'Università di Modena e Reggio Emilia e della Università di Padova.

L'obiettivo principale è stata la creazione di una **piattaforma web** per la condivisione di media e documenti, che mettesse a disposizione inoltre **risorse interattive** per l'auto valutazione del livello di competenza nella lingua italiana

Gli utenti finali dell'applicazione sono italiani di seconda e terza generazione che vivono all'estero, per i quali i contenuti del portale costituiscano approfondimenti sulla cultura italiana ed esercizi didattici



II MICC

Il Media Integration and Communication Center (MICC) è un **centro di eccellenza** istituito dal MIUR presso l'Università degli Studi di Firenze nel 2001.

Il MICC è un centro di ricerca **interdisciplinare** avanzata nel campo del processamento, interpretazione e trasmissione di immagini e video, del **multimedia**, delle telecomunicazioni e degli studi che riguardano l'informatica legale e le **tecnologie di internet**.

Il MICC è diretto dal Professor **Alberto Del Bimbo**, ordinario alla Facoltà di Ingegneria dell'Università degli Studi di Firenze



8 chiavi per l'italiano

8 chiavi per l'italiano



1. Come è nato l'italiano?

Dal latino ai volgari. L'italiano, come le altre lingue romanze e tutti i dialetti, deriva dal latino volgare. Con la caduta dell'Impero romano, quindi a partire dal V sec. d.C., il...



1. Come è nato l'italiano?



2. Che differenza c'è tra lingua e dialetto?



3. Com'è l'italiano di oggi?



4. Si scrive come si parla?

1 - 4 su 8 totali

CondiVivit

Vuoi raccontare la tua storia? Usa questa scheda per parlarci del tuo rapporto con la lingua e la cultura italiana, delle origini della tua

Immagini



Ritratto di Galileo
(nella galleria "Letteratura")



Papa Clemente IX (Giulio Rospigliosi)
(nella galleria "Letteratura")



I templi di Era a Paestum
(nella galleria "Campania")

Vedi tutto

Audio/Video

Da fotografia animata a decima musa: il cinema - Parte 2

Da "Il mezzo del messaggio" - RAI Radio 3

Laboratorio di ceramica a Cividale del Friuli
Dal programma "Serenio variabile" - Rai Due (2012)

I promessi sposi - Capitolo X
Lettura di Silvia Cecchini

Vedi tutto

Lingua

- > Storia linguistica d'Italia
- > Strutture della lingua italiana
- > Varietà dell'italiano
- > Dialetti e altri idiomi d'Italia
- > Lingua ed emigrazione italiana nel mondo

Didattica dell'italiano

- > Percorsi linguistico-didattici

Letteratura, teatro e arti

Il sistema



DrupalTM

Data la natura **collaborativa** del progetto, con contributi pubblicati da diverse unità di ricerca, si è implementata **una piattaforma CMS** (Content Management System) di base per dare agli editori la possibilità di lavorare direttamente sui documenti del portale senza alcuna supervisione da parte degli sviluppatori.

VIVIT si basa su **Drupal**, software PHP opensource per il quale sono stati implementati moduli e una base di dati MySQL personalizzati.



Gestione dell'informazione

I contenuti sono organizzativi in **due macro-categorie principali**:

- contenuti che possono essere **letti e commentati**, che includono le risorse che riguardano la storia linguistica e la cultura italiana;
- contenuti che richiedono **partecipazione attiva e interazione** da parte degli utenti (quiz, esercizi di comprensione anche multimediali).

Tassonomie sono state utilizzate per creare relazioni fra le risorse del corpus di documenti: tali tassonomie sono state redatte sia **manualmente** dagli editori, sia attraverso **contenuti generati dagli utenti** (tags) o utilizzando **tecniche semiautomatiche**.

Le **relazioni** fra risorse dunque possono essere definite dagli editori dei contenuti e/o dai professionisti del settore usando appositi campi disponibili durante il processo di sottomissione e pubblicazione dei documenti, **utilizzando tali tassonomie**.



Dante Alighieri

- > "De monarchia"
- > L'Inferno (1911)
- > Dante Alighieri, "Convivio", Venezia, 1521

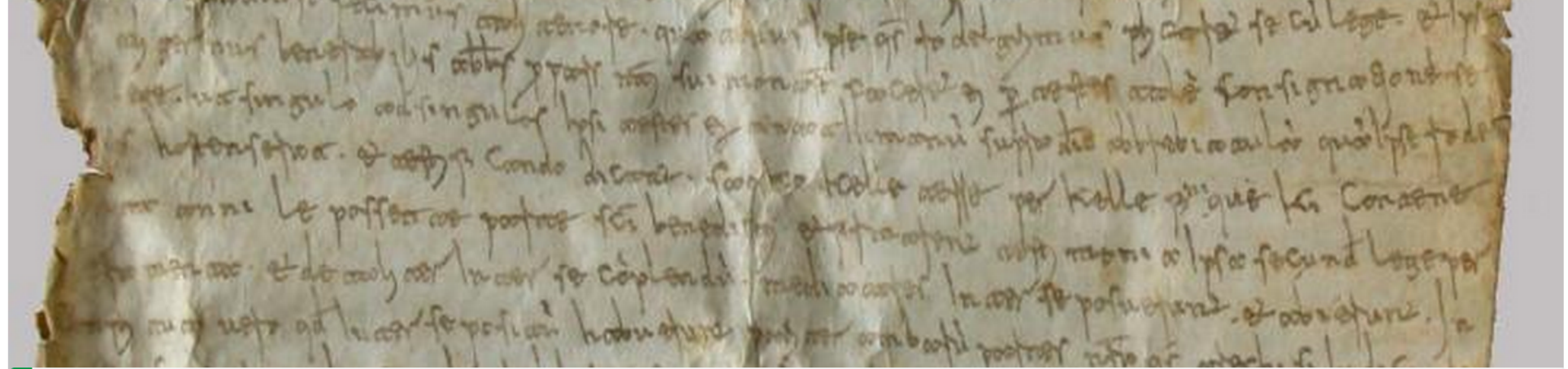
- > Aggiungi questa pagina
- > Personalizza

8 chiavi per l'italiano

Società e costume

- > Comunicazioni di massa
- > Cucina
- > Migrazioni
- > Moda e design

Archivi digitali



1. Come è nato l'italiano?

Dal latino ai volgari. L'italiano, come le altre lingue romanze e tutti i dialetti, deriva dal latino volgare. Con la caduta dell'Impero romano, quindi a partire dal V sec. d.C., il latino classico restò solo lingua scritta, mentre, sulla base delle

L'italiano resterà però almeno fino al secondo dopoguerra quasi esclusivamente lingua scritta e i dialetti svolgeranno la funzione di lingue di comunicazione.

Le pagine correlate più lette



Francesco Petrarca
(in Autori e opere)



Giovanni Boccaccio
(in Autori e opere)



Dante Alighieri
(in Autori e opere)



Dalla frammentazione del latino alla formazione delle lingue neolatine
(in Storia linguistica d'Italia)

Moduli per esercizi

Sono state studiate **soluzioni interattive avanzate** per la creazione e l'integrazione di materiali didattici all'interno del portale.

I moduli interattivi consentono la navigazione attraverso serie di **quiz/domande** che vanno dai **cloze test**, agli esercizi a risposta singola o multipla, da vari tipi di campi di input ad esercizi di **associazione immagine/parola**.

Feedback visuale è sempre fornito al discente sulla correttezza o meno delle risposte, inoltre si fornisce la possibilità di **completare e controllare automaticamente i vari quiz**.

Riflessione sulla lingua: imperfetto/passato prossimo 1

Leggi questo brano, tratto (con alcune variazioni) dall'intervista ad Elton.

Io **sono nato** in Brasile, ma sono di origine italiana, sono di origine trentina, veneta e calabrese. I miei due bisnonni paterni **erano** di Trento, Fornace e Civezzano. Da parte di mia madre, mia nonna **era** figlia di veronesi, di San Bonifacio, e mio nonno **era** calabrese, di Santa Sofia d'Epiro. Quando **ero** bambino e **andavo** a casa di mio nonno, dalla parte del mio papà, non **capivo** niente di quello che **diceva**, perché il suo **era** un dialetto, e quindi **parlavo** con lui sempre in portoghese. Allora **dicevo** sempre, fin da piccolo, che quando sarei andato a scuola, avrei imparato anche la lingua italiana perché **volevo** capire quello che mio nonno **diceva**. Quando **ho cominciato** a studiare l'italiano, **sono andato** a casa del nonno per provare a parlare con lui. Ma lì mi **sono reso conto** che lui non **capiva** quello che **dicevo** io, perché io **parlavo** italiano, imparato a scuola e dalle grammatiche, e lui **parlava** dialetto, quindi **erano** due cose diverse l'una dall'altra.

Osserva il testo con attenzione e clicca sulle informazioni corrette; per vedere la spiegazione vai alla pagina seguente.

Nel brano letto, Elton usa l'imperfetto per

« 1 di 2 »

- ☐ descrivere persone, luoghi o situazioni nel passato.
- ☐ indicare un'azione finita.
- ☐ indicare un'azione puntuale.
- ☐ indicare un'abitudine nel passato.

Controlla

Imperfetto/passato prossimo: esercizio 1

Completa i testi tratti da un forum usando l'imperfetto o il passato prossimo.

1. Scegli tra il passato prossimo e l'imperfetto. A volte sono possibili entrambi.



SOS: ricerca nonna italiana!
Cerco notizie di mia nonna materna, Joana Fiorentin, emigrata in Brasile quando 20 anni. Lì due anni in un campo di caffè, ma poi mio nonno che un piccolo negozio di alimentari e a lavorare per lui. un po' alla volta e nel 1920 . nello stato di Santa Caterina in una piccola fazenda e in cinque anni quattro figli, due maschi e due femmine. Io da una delle figlie nel 1990. La nonna però So che da Castelfranco veneto e che due sorelle.
Grazie a chi mi aiuta a trovare sue informazioni! Isabela Silva Pereira

Controlla Soluzione Ripristina

Livelli di accesso



Il portale VIVIT offre diversi livelli di accesso:

- **utenti non autenticati**: possono navigare il sito e ricercare contenuti;
- **utenti autenticati**: possono navigare, cercare, fare il **bookmark** dei contenuti e condividerli con altri utenti, inviare e ricevere **messaggi** privati, pubblicare **storie**, postare commenti;
- **collaboratori**: possono in più **pubblicare** e **categorizzare** i propri contenuti;
- **editori**: possono **inserire** ed **editare** ogni tipo di contenuto, incluse le risorse dei collaboratori e creare specifiche **relazioni** di dominio;
- **amministratori**: tutti i permessi dei livelli precedenti e funzioni di amministrazione del sito.

L'obiettivo finale non è solo la creazione e fruizione di contenuti professionali, ma la costituzione di una **comunità di utenti** attivamente coinvolta nella vita del portale e nella condivisione di documenti anche di esperienza e vita personale.

Bookmark dei contenuti



Andrea Ferracani

[Profilo utente](#) [Messaggi](#) [Racconta la tua storia](#) [Esci](#)



vivit

Vivi italiano

Il portale dell'italiano nel mondo

[CHE COS'È VIVIT?](#) [IL PROGETTO](#) [L'ACCADEMIA DELLA CRUSCA](#) [CREDITI](#)

Cerca nel sito:

Cerca

I miei preferiti

- ▼ Dante Alighieri
 - > "De monarchia"
 - > L'Inferno (1911)
 - > Dante Alighieri, "Convivio", Venezia, 1521
- > [Aggiungi questa pagina](#)
- > [Personalizza](#)

8 chiavi per
l'italiano

I miei preferiti

Elenca elementi

Aggiungi voce

Preferiti preimpostati

Voce del menu

Attivato

Espanso

Pubblico

Operazioni

⊕ Dante Alighieri



[modifica](#)

[elimina](#)

⊕ "De monarchia"



[modifica](#)

[elimina](#)

⊕ L'Inferno (1911)



[modifica](#)

[elimina](#)

⊕ Dante Alighieri, "Convivio", Venezia, 1521



[modifica](#)

[elimina](#)

[Salva configurazione](#)



Bookmark dei contenuti

Lingua

- > Storia linguistica d'Italia
- > Strutture della lingua italiana
- > Varietà dell'italiano
- > Dialetti e altri idiomi d'Italia
- > Lingua ed emigrazione italiana nel mondo

Didattica dell'italiano

- > Percorsi linguistico-didattici

Letteratura, teatro e arti

- > Letteratura e teatro
- > Arti

Società e costume

- > Comunicazioni di massa
- > Cucina
- > Migrazioni
- > Moda e design

Cronologia

Iscritto da

4 anni 2 giorni

Vivit



Andrea Ferracani  usa Vivit

Preferiti

- ▼ Dante Alighieri
 - > "De monarchia"
 - > L'Inferno (1911)
 - > Dante Alighieri, "Convivio", Venezia, 1521

I tuoi amici di Facebook su VIVIT



Daniele Pezzatini



Niccolò Becchi

Pubblicazione storia



Andrea Ferracani

[Profilo utente](#) [Messaggi](#) [Racconta la tua storia](#) [Esci](#)



Vivi italiano
Il portale dell'italiano nel mondo

[CHE COS'È VIVIT?](#) [IL PROGETTO](#) [L'ACCADEMIA DELLA CRUSCA](#) [CREDITI](#)

Cerca nel sito:

Cerca

I miei preferiti

- ▼ Dante Alighieri
 - > "De monarchia"
 - > L'Inferno (1911)
 - > Dante Alighieri, "Convivio", Venezia, 1521
- > [Aggiungi questa pagina](#)
- > [Personalizza](#)

8 chiavi per
l'italiano

Crea Contributo utente

immagine:

Choose File

No file chosen

Carica

Dimensione massima del file: 2 MB

Estensioni permesse: *png gif jpg jpeg*

Inserire una immagine (dimensioni massime 300x300 pixel: le immagini maggiori saranno ridimensionate automaticamente)

Titolo: *

Testo:

☒ Mostra il sommario nella vista completa

Funzionalità sociali

Gli utenti hanno la possibilità non solo di fare browsing dei contenuti, ma anche di condividerli sui principali social networks: **Facebook**, **Twitter**, **Google**.

Inoltre la creazione dell'account sul portale può essere effettuata anche via **Facebook Login** (implementata con il Facebook PHP SDK).

Il modulo di connessione Facebook consente all'utente la visualizzazione della foto del profilo su VIVIT e di mostrargli gli **amici di Facebook** già presenti su VIVIT.

La funzionalità migliora il **coinvolgimento** dell'utente su VIVIT e contribuisce all'allargamento della comunità.

Vivit



Andrea Ferracani 🇺🇸 usa Vivit

Preferiti

- ▼ Dante Alighieri
 - > "De monarchia"
 - > L'Inferno (1911)
 - > Dante Alighieri, "Convivio", Venezia, 1521

I tuoi amici di Facebook su VIVIT



Daniele Pezzatini



Niccolò Becchi



Stefania Iannizzotto



Vera Gheno



Analisi dei testi e semantica

Ad integrazione del sistema di taggatura manuale e del modulo di gestione delle tassonomie, VIVIT mette a disposizione un **sistema automatico** di estrazione di **topics** e **named entities** (nomi, entità, date e luoghi), che gira come servlet sul server del MICC.

Il servizio, chiamato “**Homer**”, permette l’input diretto di testi da parte degli utenti o via API e può generare risposte con liste di topics, entities, e termini analizzando documenti o url.

La servlet espone diversi **parametri configurabili**: gli utenti possono settare filtri sul tipo di entità da rilevare, numero di risultati da restituire, formato di output (JSON o XML).



Analisi del testo - funzionalità principali

- **rilevazione della lingua**: il servizio determina la lingua del documento da analizzare: è un passaggio di preprocessing essenziali per le analisi successive. In particolare è la base per procedere con l'eliminazione delle stop-word;
- **eliminazione delle stop-word**: l'obiettivo è eliminare tutte le parole prive di significato semantico quali **congiunzioni**, **articoli** etc. Il passaggio è basato su liste di stop-words e necessita la conoscenza della lingua del documento da processare.

Può essere considerata una procedura di **filtraggio**, benefica per altre tecniche di analisi.

In caso di incertezza nella rilevazione della lingua la rilevazione delle stop-word è usata al fine della **disambiguazione** della lingua;



Analisi del testo - funzionalità principali

- **estrazione di tag-cloud:** servizio utile per avere una rappresentazione veloce di un documento. Esegue una analisi di **frequenza** delle parole del documento e sceglie le N parole più comuni. Per essere efficace è necessario prima togliere le stop-word, altrimenti è molto probabile che esse siano le parole più frequenti (secondo l'empirica legge di Zipf).
- **rilevazione dei topic:** servizio che seleziona le parole di un documento via analisi statistica. Usando la tecnica **LDA** è possibile selezionare le parole che non necessariamente sono le più frequenti, come nell'estrazione tag-cloud, ma più significative.
- **named entity detection:** servizio che identifica nel testo classi di risorse quali nomi di persona, date, luoghi. Si tratta di un servizio utile per la rilevazione di **concetti semantici** in documenti siano tali parole frequenti o no (il secondo caso è più probabile).



Analisi del testo - Homer

Analyse text

Text to be analysed: Scritto in latino e suddiviso in tre libri, è l'unico trattato in prosa che Dante abbia terminato. La scrive quando gli è stata confermata la sua condanna a morte e l'impresa di Arrigo VII è ormai fallita. Dante riprende e sviluppa – una per ogni libro – le tre idee politiche in cui crede fermamente, già presenti anche nel Convivio, e che saranno approfondite in molti		Analysis: Entities detection ☐ HTML input format ☐ HTML input article extraction ☒ JSON output format	Process
Text URL: 			
Topic detection parameters:			
Number of topics (-1 for autodetection): 3			
Number of keywords (-1 for autodetection): 7			
Entity detection parameters:			
☒ All entities	☒ Person entities	☐ Place entities	
☐ Date entities	☐ Time entities	☐ Currency entities	
☐ Percentage entities	☐ Organization entities		

<http://shrek.micc.unifi.it:8080/homer/>

Analisi del testo - Homer

8 chiavi per
l'italiano

Lingua

- > Storia linguistica d'Italia
- > Strutture della lingua italiana
- > Varietà dell'italiano
- > Dialetti e altri idiomi d'Italia
- > Lingua ed emigrazione italiana nel mondo

Didattica dell'italiano

- > Percorsi linguistico-didattici

Letteratura, teatro e arti

- > Letteratura e teatro
- > Arti

Società e costume

Dante Alighieri

IMMAGINI	AUDIO/VIDEO	OPERE	APPROFONDIMENTI	CONTATTO COL TESTO	STRUMENTI DI RICERCA
					
Andrea del Castagno, "Ritratto di Dante", 1450		Così nel mio parlar voglio esser aspro	Dante Alighieri, "Convivio", Venezia, 1521		
					1 di 4 >>

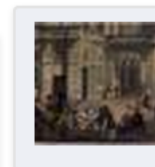


Materiali collegati

Elenco dei materiali collegati a questa scheda

Scheda collegata:

- 1. L'età di Dante e il Trecento
- 1.1 La poesia in Toscana: Dante e lo Stilnovo



Tutte le vicende d'Italia, linguistiche, culturali e politiche, sono rimaste segnate dall'opera di Dante. Con il suo nome, innanzi tutto, il nome d'Italia si è fatto conoscere nel mondo.

Già nella raccolta di poesie giovanili della *Vita Nova*, Dante esprime chiaramente la necessità di una lingua letteraria che andasse al di là dei **singoli volgari italiani**: la "lingua del sì"; questa idea verrà elaborata nel *Convivio* e quindi definita nel trattato *De vulgari eloquentia*, all'interno di una visione che abbraccia l'intera Europa.

La *Commedia* rappresenta l'attuazione concreta di questo pensiero e diventa anche occasione per lanciare l'idea dell'Italia come entità politica da ricostruire.

I canti del poema cominciano a diffondersi ovunque via via che escono dalla penna del suo autore e «tutta l'Italia sembra dar l'assalto, con un esercito di copisti, alla *Commedia*; in realtà ne è assalita e sconvolta» (così nelle parole del filologo e linguista Ignazio Baldelli, 1922–2008).

Con Dante la lingua italiana inizia ad affermarsi rapidamente al di sopra delle varietà di tradizioni linguistiche e culturali esistenti: già nel 1332, secondo il trattatista padovano **Antonio** da Tempo, il toscano letterario ha il primato assoluto tra i volgari d'Italia. L'opera di Dante, sostenuta dalla fortuna del *Canzoniere* di **Petrarca** e del *Decameron* di **Boccaccio**, farà sì che la lingua italiana si collochi a pieno titolo accanto alle altre grandi lingue nazionali, nonostante l'assenza di un'unità politica.

Analisi del testo - funzionalità principali

```
<rss version="2.0">
  <channel>
    <title>Homer RSS feed</title>
    <link>http://shrek.micc.unifi.it:8080/homer/Homer</link>
    <numresults>7</numresults>
    <item>
      <person>
        <keyword>Vita Nova</keyword>
        <tf>0.004385965</tf>
        <occurrences>1</occurrences>
        <confidence>1.0</confidence>
      </person>
    </item>
    <item>
      <person>
        <keyword>Ignazio Baldelli</keyword>
        <tf>0.004385965</tf>
        <occurrences>1</occurrences>
        <confidence>1.0</confidence>
      </person>
    </item>
    <item>
      <date>
        <keyword>1922-2008</keyword>
        <tf>0.004385965</tf>
        <occurrences>1</occurrences>
        <confidence>1.0</confidence>
      </date>
    </item>
  </channel>
</rss>
```

```
<item>
  <person>
    <keyword>Dante</keyword>
    <tf>0.004385965</tf>
    <occurrences>1</occurrences>
    <confidence>1.0</confidence>
  </person>
</item>
<item>
  <date>
    <keyword>1332</keyword>
    <tf>0.004385965</tf>
    <occurrences>1</occurrences>
    <confidence>1.0</confidence>
  </date>
</item>
<item>
  <person>
    <keyword>Antonio Da Tempo</keyword>
    <tf>0.004385965</tf>
    <occurrences>1</occurrences>
    <confidence>1.0</confidence>
  </person>
</item>
</channel>
</rss>
```



Risorse aggiuntive

Il progetto VIVIT dà inoltre accesso a risorse esterne sviluppate nel contesto dei progetti FIRB, sviluppate negli anni al MICC che sono connesse al campo di applicazione della **linguistica computazionale**.

lit lessico dell'italiano televisivo Nuova ricerca Modifica

Videoteca

ID	Title	Status	XML-TEI
1	cd1 - belardinelli/bico/20060714	annotated	http://deckard.micc.unifi.it/crusca/lit/
2	cd1 - belardinelli/bico/20060715	annotated	http://deckard.micc.unifi.it/cruscaxml/
3	cd1 - belardinelli/bico/20060716	annotated	http://deckard.micc.unifi.it/cruscaxml/
4	cd1 - belardinelli/bico/20060717	annotated	http://deckard.micc.unifi.it/crusca/lit/

Trasmissione

Trasmissioni nel video:
Il commissario Rex

Dati Trasmissione:

Rete: Rai Uno
Titolo: Il commissario Rex
Data: 14/07/2006
Tassonomia: Programmi TV -> Fiction -> Serie
Parlato: 00:09:36:209

Metadati

Persona: pm1
Sesso: Maschio
Tipo di conversazione: dialogo
Tipo di parlato: esecutivo
Tipo di Voce: In campo
Parlante: Interno

Intervallo Battuta: 00:00:02:665 - 00:00:05:174 00:00:00:000

la batteria // no adesso no //
che c'è Rex ? // che succede ? //
scappa Rex ! via ! corri !
Fritz occupati del cane !
h h ! questa la prendo io // e tu saresti un super piediplatti eh ?
il cane è scappato //
quelli vogliono ammazzarvi // scappate ! / appena potete ! //
ho detto che vi uccideranno mi capite o no ? //
ai diavolo !
forza scendete ! / andate là dentro !
adesso puoi scendere //
dove andiamo ? / il capo mi ha riservato una calorosa accoglienza ? eh ? / di solito non parlo mai con i galoppini // non capiscono mai alla prima //

MICC ha sviluppato la piattaforma **LIT** (“Lessico dell'Italiano Televisivo”) e **LIR** (“Lessico dell'Italiano Radiofonico”), **motori di ricerca** e **gestionali** di archivi, che offrono la possibilità di pubblicare, annotare e ricercare flussi audiovisivi concessi da TV nazionali e network radio e che consentono di ricercare estratti di parlato marcati con lo standard **TEI/XML**.