

## PROTOTIPO DI ANALISI TESTUALE DELLE SENTENZE DELLA CASSAZIONE - ESTRAZIONE AUTOMATICA DI DESCRITTORI

*Manuela Sassi, Alessandra Cinini*

[www.ilc.cnr.it](http://www.ilc.cnr.it)

### PREMESSA

Nel febbraio 2003 si è svolta la prima riunione operativa fra i partner del progetto FIRB "Tecnologie per la giustizia", coordinato dall'Istituto di Ricerca sui Sistemi Giudiziari di Bologna (IRSIG), riunione in cui sono individuate alcune attività di supporto al WP1-Sistema informativo per il monitoraggio del sistema giudiziario. In particolare, l'Istituto di Linguistica Computazionale (ILC) del CNR, ha messo a disposizione alcuni strumenti, già ampiamente sperimentati nella ricerca linguistica, per estrarre informazioni linguistico-concettuali da grandi quantità di dati testuali come, ad esempio, l'insieme delle sentenze della Cassazione.

Per le sentenze della Cassazione è stata realizzata una funzione, basata su un componente del progetto ILC denominato "Linguistic Miner" (LM) [1], per l'individuazione, la classificazione e l'utilizzazione di formule linguistiche relative ai nomi propri che, opportunamente adattato, può permettere la sperimentazione dell'anonimazione dei documenti giuridici, primo passo per la loro pubblicazione su web.

Parallelamente si è studiata l'individuazione di regole per la ricerca di schemi linguistici allo scopo di affinare gli strumenti di classificazione delle sentenze, in modo tale da rendere l'operazione di Information Retrieval più produttiva rispetto a un sistema di classificazione generico o ad ampio spettro.

### IL DATABASE TESTUALE DI PARTENZA

Sono stati recuperati i testi delle sentenze della Corte di Cassazione presenti sul CD distribuito dal Ministero della Giustizia, nel formato txt. Tali testi sono stati poi codificati per l'inserimento nel DBT (Data Base Testuale, Brevetto CNR di Picchi E.) [2], che è un software di gestione di banche dati testuali, frutto dell'esperienza, ormai vicina ai 40 anni, dei ricercatori dell'ILC.

L'indicizzazione dei testi, opportunamente trattati, ha generato 75 archivi per un totale di 92.481.142 parole (corrispondenti a circa 700 mega). Gli archivi delle Sentenze della Cassazione possono essere consultati, integralmente e contemporaneamente, con l'interfaccia DBT-Corpus, che permette di effettuare ricerche globali su tutti i testi. Su questo tema, e cioè la consultazione di grandi depositi testuali e le possibilità di ricerca offerte dal DBT si veda nell'Allegato 1 un esempio di interrogazione e dei risultati ottenuti.

## CARATTERISTICHE DEL DBT

Il software DBT è stato concepito come un sistema di trattamento e interrogazione di testi che tiene in considerazione l'aspetto filologico e facilita lo studio dei particolari fenomeni linguistici. Si caratterizza anche per il rispetto totale delle qualità lessicografiche della risorsa linguistica a cui viene applicato e per la capacità di ottenere, rapidamente e in modo interattivo, funzioni di analisi testuale automatica, quali:

- la ricerca libera di parole nel testo (con possibilità molto estese di mascheramento e di combinazione di termini attraverso gli operatori booleani);
- la ricerca di lemmi, che, attraverso il motore morfologico, recupera tutte le forme del lemma dato e/o dei sinonimi;
- il calcolo delle frequenze in ordine alfabetico o decrescente;
- le concordanze sotto vari formati (direttamente pubblicabili via RTF) e in vari ordinamenti (secondo l'ordine di presenza nei testi, oppure secondo l'ordine alfabetico delle forme che seguono/precedono nella frase la parola oggetto di ricerca);
- gli index locorum, arricchiti da riferimenti puntuali;
- il calcolo dell'indice di leggibilità, basato sulle formule statistiche più accreditate.

Tutto ciò, unito ad altri ulteriori vantaggi (possibilità di analisi morfologica, integrazione di un tagger automatico della lingua con output marcato, ecc.) ha permesso di definire e sviluppare, a partire dallo stesso strumento computazionale, una strategia di trattamento dell'informazione con fini multidisciplinari. Il DBT ha la capacità di gestire grandi quantità di dati con risparmio di tempo e di spazio-disco, e una delle sue funzioni statistiche più produttive è la co-occorrenza: "Per co-occorrenza statistica (Mutual information) si intende il calcolo della probabilità che hanno le parole di un testo di essere associate, in quel testo, con una o più parole prestabilite".

## IL TAGGING AUTOMATICO

I 75 archivi DBT sono stati interamente lemmatizzati tramite la funzione di tagging automatico, che individua il *lemma*, ossia la chiave di ricerca in un comune dizionario cartaceo, le categorie grammaticali di ciascun lemma e quelli relativi alla forma, secondo la funzione svolta nella frase dalla parola data.

In relazione a questa fase dobbiamo osservare che la qualità dei testi ha presentato alcuni problemi nel riconoscimento dei lemmi in quanto sono presenti refusi; tra i più frequenti, anche se parzialmente superabile è la accentazione, sia in forma di lettera accentata sia di lettera seguita da apice (es.: *attivitá'* e *attivitá*). Il tagger riconosce entrambi i casi, ma dato il frequentissimo uso delle virgolette semplici (ottenute tramite l'uso dell'apice prima e dopo la parola o frase) ciò può produrre errore nell'analisi. Altri tipi di refusi sono banali errori di battitura, comunque quando si utilizzano grosse quantità di dati, come in questo caso, la sintesi a livello statistico permette di ridurre al minimo l'influenza dell'errore sul risultato finale.

Il prodotto della fase di marcatura viene archiviato per essere poi utilizzato nelle fasi successive. Un esempio si può vedere nella tabella seguente, dove si mostra una porzione di file "taggato".

2124	LA	LA	R	FS
2125	RICORRENTE	RICORRENTE	SN	NS
2126	LAMENTA	LAMENTARE	VTP	S3IP
2127	CHE	CHE	C	
2128	LA	LA	R	FS
2129	SENTENZA	SENTENZA	SF	FS
2130	IMPUGNATA	IMPUGNATO	A	FS
2131	NON	NON	BS	
2132	CONTERREBBE	CONTENERE	VTRP	S3DP
2133	UNA	UNO	RI	FS
2134	ADEGUATA	ADEGUATO	A	FS
2135	ESPOSIZIONE	ESPOSIZIONE	SF	FS
2136	DELLE	DI	E	FP
2137	RAGIONI	RAGIONE	SF	FP
2138	DEL	DI	E	MS
2139	DECIDERE	DECIDERE	VTIP	F
2140	,	,		
2141	IN	IN	E	
2142	RELAZIONE	RELAZIONE	SF	FS
2143	ALLA	A	E	FS
2144	DEDOTTA	DEDOTTO	A	FS
2145	VIOLAZIONE	VIOLAZIONE	SF	FS
2146	DEL	DI	E	MS
2147	CONTRADDITTORIO	CONTRADDITTORIO	SM	MS
2148	.	.		

Nella prima colonna compare il numero progressivo con cui vengono archiviati i record; nella seconda vi sono le parole così come si trovano nel testo, nella terza appare il lemma, in quarta e quinta colonna vi sono i codici grammaticali (ad esempio: S FP equivale a Sostantivo nella forma Femminile Plurale e VTI S3IP equivale a verbo Transitivo/Intransitivo nella forma della 3 persona singolare dell'Indicativo Presente).

## RICERCA DEI DESCRITTORI

L'operazione di marcatura sopra descritta permette di applicare ai testi le regole per la ricerca di schemi linguistici predefiniti. Questo avviene attraverso un componente del LM denominato "textminer", la cui *console* di gestione permette l'attivazione di tali regole, l'analisi e l'archiviazione dei risultati ottenuti dalla loro applicazione ai testi marcati opportunamente selezionati.

In particolare è stata studiata e applicata la regola che individua tutte le coppie di sostantivi separate dalla preposizione "DI", frutto di una analisi dei testi in oggetto e dell'esperienza nel settore, per cui sono stati ritenuti significativi per la classificazione automatica delle sentenze.

Prendendo come esempio la frase contenuta nella tabella precedentemente inserita, si può dedurre che, data la regola anzidetta, nel risultato appariranno i

seguenti casi: "ESPOSIZIONE (DI) RAGIONE" e "VIOLAZIONE (DI) CONTRADDITTORIO", dato che in entrambi i casi si tratta di sostantivi connessi dalla preposizione "DI". Oltre a questa si è ritenuta significativa l'applicazione di un'altra regola che permette di recuperare anche le coppie di sostantivi e aggettivi collegati.

In tal modo sono stati ottenuti dei risultati, che abbiamo denominato "Bigrammi", contenenti nel primo caso le coppie di sostantivi connessi dalla preposizione "DI", nel secondo caso le coppie sostantivo-aggettivo.

I risultati ottenuti sono stati organizzati in file di testo in formato tabellare, dove le colonne contengono i seguenti dati:

- Campo1 lemma del primo sostantivo (es.: *abbandono*)
- Campo2 forma del primo sostantivo (es.: *abbandoni* o *abbandono*)
- Campo3 lemma del secondo sostantivo o dell'aggettivo (es.: *alloggio* o *consapevole*)
- Campo4 forma del secondo sostantivo o dell'aggettivo (es.: *alloggi* o *consapevoli*)
- Campo5 Identificativo della sentenza da cui è stato estratto il bigramma (1990007336, dove si trova l'anno e il N. di archivio).

Di seguito si mostra un estratto di entrambi:

		SOSTSOST.TXT		
ABBAGLIAMENTO	ABBAGLIAMENTO	SELVAGGINA	SELVAGGINA	0199007336
ABBANDONO	ABBANDONO	ABITAZIONE	ABITAZIONE	0200002217
ABBANDONO	ABBANDONO	ABITAZIONE	ABITAZIONE	0199712829
ABBANDONO	ABBANDONO	ADOTTATO	ADOTTATA	0199107357
ABBANDONO	ABBANDONO	AFFARE	AFFARE	0199804906
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199700999
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199700999
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199700999
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199700999
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199100159
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199100556
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199608830
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199512242
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199403514
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199403514
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199713045
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199713045
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199808530
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199808530
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199808530
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199806586
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199808530
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199001931
ABBANDONO	ABBANDONO	ALPE	ALPE	0199402755
ABBANDONO	ABBANDONO	ALVEO	ALVEO	0199805868
ABBANDONO	ABBANDONO	ALVEO	ALVEO	0199110607
ABBANDONO	ABBANDONO	ALVEO	ALVEO	0199110607
ABBANDONO	ABBANDONO	ALVEO	ALVEO	0199110607
ABBANDONO	ABBANDONO	APPARTAMENTO	APPARTAMENTO	0199011443
ABBANDONO	ABBANDONO	APPARTAMENTO	APPARTAMENTO	0199001931

		SOSTAGG.TXT		
ABBAGLIO	ABBAGLIO	MONUMENTALE	MONUMENTALE	200010921
ABBAINO	ABBAINO	DISCUSO	DISCUSO	199801498
ABBAINO	ABBAINO	POSTO	POSTO	199501456
ABBAINO	ABBAINI	POSTO	POSTI	200008480

ABBAINO	ABBAINO	PREESISTENTE	PREESISTENTE	199501803
ABBAINO	ABBAINO	PROPOSTO	PROPOSTA	199501803
ABBAINO	ABBAINO	RICOSTRUITO	RICOSTRUITO	199501803
ABBAINO	ABBAINO	RICOSTRUITO	RICOSTRUITO	199501803
ABBAINO	ABBAINO	RICOSTRUITO	RICOSTRUITO	199501803
ABBANDONO	ABBANDONO	ACCERTATO	ACCERTATO	199704972
ABBANDONO	ABBANDONO	ACCERTATO	ACCERTATO	199607137
ABBANDONO	ABBANDONO	ACCERTATO	ACCERTATA	199607137
ABBANDONO	ABBANDONO	ACCERTATO	ACCERTATO	199009313
ABBANDONO	ABBANDONO	AFFETTIVO	AFFETTIVO	199800120
ABBANDONO	ABBANDONO	AFFETTIVO	AFFETTIVO	199003346
ABBANDONO	ABBANDONO	ALLEGATO	ALLEGATO	200010682
ABBANDONO	ABBANDONO	ARTICOLATO	ARTICOLATO	199012156
ABBANDONO	ABBANDONO	ASSOLUTO	ASSOLUTO	199408651
ABBANDONO	ABBANDONO	AVVENUTO	AVVENUTO	199110607
ABBANDONO	ABBANDONO	AVVENUTO	AVVENUTO	200014548
ABBANDONO	ABBANDONO	AVVENUTO	AVVENUTO	199504196
ABBANDONO	ABBANDONO	COMPLETO	COMPLETO	199006389
ABBANDONO	ABBANDONO	COMPLETO	COMPLETO	199501802
ABBANDONO	ABBANDONO	COMPLETO	COMPLETO	199604912
ABBANDONO	ABBANDONO	COMPLETO	COMPLETO	199307745
ABBANDONO	ABBANDONO	COMPLETO	COMPLETO	199805556

Questi dati - SOSTSOST.TXT e SOSTAGG.TXT - possono essere trattati con qualunque piattaforma in quanto sono in formato testo delimitato da tabulatori. Abbiamo scelto di inserire anche la forma, perché, mentre il lemma può essere utile per fare la sintesi, è la forma che può essere recuperata dai testi.

Con l'obiettivo di mostrare l'utilità di questi dati, oltre a quella che ovviamente ne può trarre un sistema automatico di classificazione che utilizzi tali bigrammi per associare ogni singola sentenza ad un certo numero di voci.

Si evidenziano di seguito alcune analisi, in particolare si propone l'analisi del termine *abbandono* in entrambi i casi: sia abbinato al sostantivo successivo che all'aggettivo.

Nel primo elenco si sono verificate 530 occorrenze, nella tabella seguente si estrapolano solo quelle con frequenza maggiore di 2.

ABBANDONO DI:

MINORE	67	ALVEO	4
LAVORAZIONE	55	CAUSA	4
LAVORO	49	CRITERIO	4
POSTO	47	IMMOBILE	4
ATTIVITA'	29	ISTANZA	4
DOMANDA	21	MERCE	4
ALLOGGIO	19	POSIZIONE	4
CASA	18	TERRENO	4
DIRITTO	11	CANE	3
DOMICILIO	11	COLTIVAZIONE	3
CANTIERE	8	INCARICO	3
FIGLIO	6	RICHIESTA	3
FONDO	6	SEDE	3
TETTO	6	SISTEMA	3

Lo stesso procedimento è stato adottato nel secondo elenco, dove si sono verificate 163 occorrenze, da cui si estrapolano quelle maggiori di 2.

ABBANDONO:

VOLONTARIO	16	RITENUTO	4
DEFINITIVO	12	SOSTANZIALE	4
MORALE	10	TEMPORANEO	4
MATERIALE	9	AVVENUTO	3
TOTALE	7	IMPLICITO	3
EFFETTIVO	6	INGIUSTIFICATO	3
COMPLETO	5	INTERVENUTO	3
ACCERTATO	4	PRETESO	3
CONSEQUENTE	4	SPONTANEO	3
IMMEDIATO	4	TACITO	3
PRIVO	4		

Una sintesi ulteriore permette di individuare le sentenze da cui sono stati estratti i precedenti bigrammi, impostati sul lemma abbandono.

199001536	ABBANDONO di ATTIVITA'	4		
	ABBANDONO di LAVORO	3		
	ABBANDONO di POSTO	1		
			Totale 199001536	8
199001801	ABBANDONO di GEMELLO	1		
	ABBANDONO di MINORE	3		
			Totale 199001801	4
199001861	ABBANDONO di POSTO	1		
			Totale 199001861	1
199001885	ABBANDONO di ATTIVITA'	1		
			Totale 199001885	1
199001892	ABBANDONO di ATTIVITA'	4		
	ABBANDONO di LAVORO	3		
	ABBANDONO di POSTO	1		
			Totale 199001892	8
199001931	ABBANDONO di ALLOGGIO	1		
	ABBANDONO di APPARTAMENTO	1		
			Totale 199001931	2

## PROSSIMI OBIETTIVI E CONCLUSIONI

I testi così trattati permettono l'applicazione di altre regole, studiate ad hoc, per ottenere risultati e informazioni da utilizzare per l'analisi di altri fenomeni o ad altri fini. Ad esempio individuare gli elementi o le porzioni di testo che forniscano i parametri utili allo studio della giurimetria.

Comunque si vuole sottolineare come tale prototipo possa contribuire a favorire lo sviluppo della Società dell'Informazione nel campo della giustizia, e che applicando tali strumenti innovativi a un settore tanto cruciale come questo si possa veramente contribuire al tanto citato E-Government, ma soprattutto alla E-Democracy, fornendo strumenti più sofisticati alla ricerca della Informazione e alla comprensione dei fenomeni giuridici.

[1] Picchi E., Ceccotti M.L., Cucurullo S., Sassi M., Sassolini E., "Linguistic Miner: an Italian Linguistic Knowledge System" in LREC 2004 IV International Conference, Proceedings, Lisbona, 26-28 maggio 2004, pp.1811-14.

[2] Picchi, E. DBT 3 - Data Base Testuale: Guida all'uso, versione 3.1. Lexis Ricerche s.r.l. su licenza del C.N.R., Roma, 1997.

## ALLEGATO 1

Per rimanere nel campo semantico del termine *abbandono*, si è proceduto ad esaminare nel dettaglio il caso di "abbandono di minore" per selezionare le sentenze che lo menzionano. Nel caso dei Bigrammi si è recuperato solamente quei casi in cui si usa il sostantivo *abbandono*, ma non il verbo *abbandonare*. Nell'esempio seguente invece si è studiata una query che riuscisse a recuperare più informazione possibile su tale tema. Per questo motivo si è proceduto ad interrogare con il DBT il corpus completo delle sentenze, utilizzando il motore morfologico e selezionando quindi tutte le forme sia del sostantivo che del verbo.

Si è poi impostata una famiglia di ricerca complessa che collega le forme suddette recuperate nel corpus. In pratica si definiscono gruppi di parole che poi vengono riutilizzati nella definizione di un ulteriore gruppo. Nella fam.1 si trovano tutte le forme di *figlio* e derivati, nella fam.2 quelle di *adottato*, nella fam.3 quelle di *minore* e nella fam.4 tutte le voci verbali di *abbandonare* e le forme di *abbandono*. I primi tre gruppi di parole vengono associati in OR (|) e tutti insieme, tra parentesi, si associano con l'operatore AND (&) alla fam.4.

Si ottiene quindi la fam.5, comprensiva di tutte le precedenti, opportunamente collegate sia in OR che in AND. Si procede poi alla selezione dei parametri: *massima distanza* delimita il raggio di azione della ricerca combinata, ossia le parole da cercare combinate non possono distare più di 4 elementi; *dimensione contesti* definisce la grandezza del contesto, che qui risulta essere 5 parole prima e 5 dopo quelle trovate; con *sequenza* si richiede il rispetto della sequenza data nell'espressione della ricerca, quindi verranno recuperati i casi in cui abbandono (verbo o sostantivo) precedano gli altri termini collegati, per un totale di 38 forme diverse.

Per completezza di informazione si riportano tutti i 168 contesti in cui si sono verificate le condizioni richieste. Di questi 166 sono pertinenti, escludendo i 2 (il n.16 e il n.68) per l'ambiguità della parola *adottato*. I contesti sono raggruppati per sentenza.

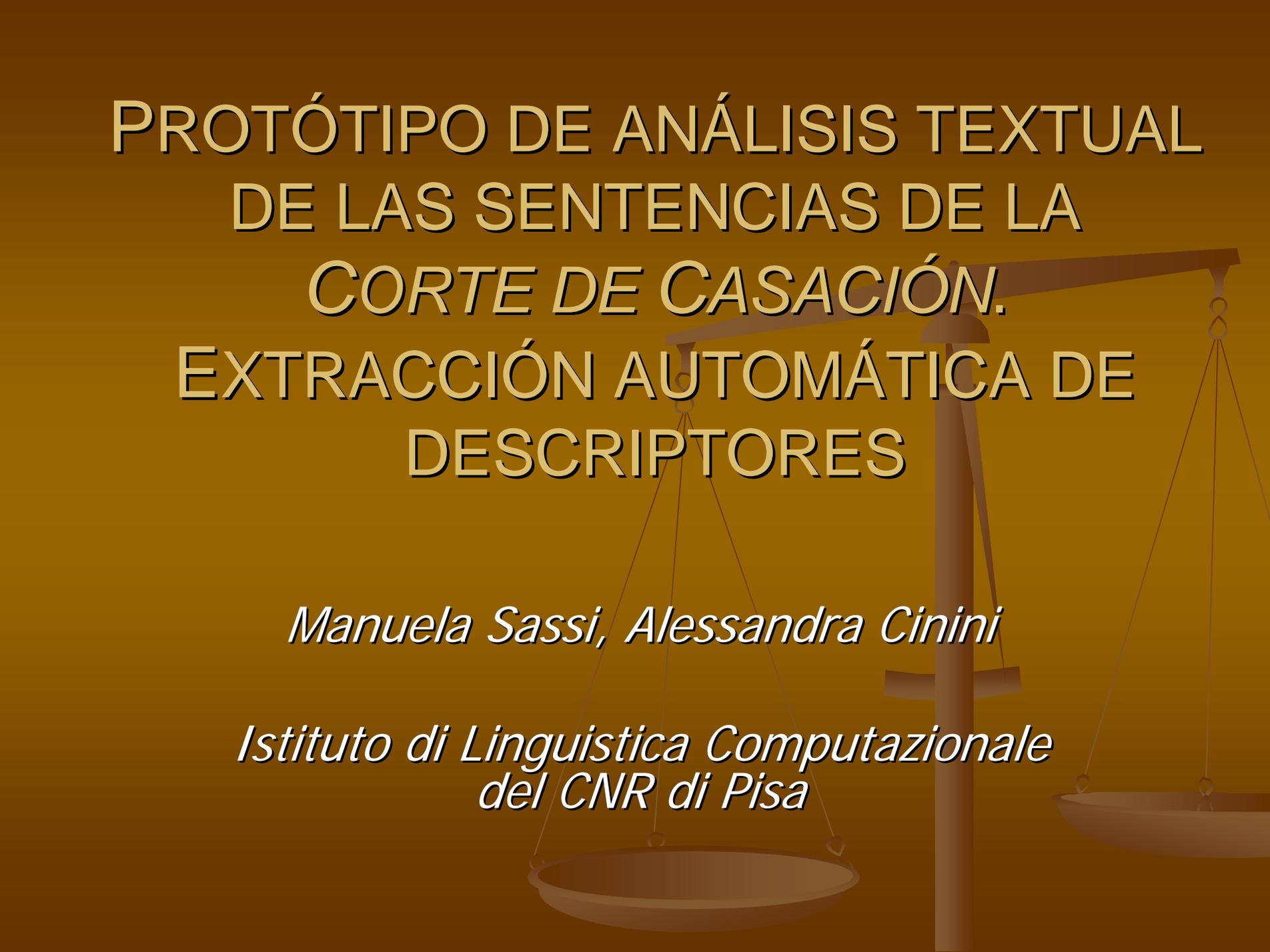
199000232	134) la menzione dello stato di *abbandono* dei *minori*, di cui anzi erano	1
	135) acquisire prove sullo stato di *abbandono* dei due *minori* nel loro Paese. Anche	1
	136) sola menzione dello stato di *abbandono* dei *minori*, dei quali erano indicati	1
		3
199000903	131) La giovanissima madre non avrebbe *abbandonato* il *figlio*: non appena ristabilitasi	1
	132) stessa coppia abbia o no *abbandonato* il *figlio* di loro figlia. Il	1
	137) atti idonei essendo stati l'*abbandono* del proprio *figlio* immediatamente dopo il parto	1
		3
199001801	133) e per contestare dia vere *abbandonato* moglie e *figli*, ai quali dichiarava di	1
	138) sosteneva che lo stato di *abbandono* delle *minori* **AM** era divenuto irreversibile:	1
	139) determinato uno stato di oggettivo *abbandono* delle *minori*, soprattutto morale.	1
	140) dimostrare che lo stato di *abbandono* dei *minori*, da valutarsi all'epoca	1
		4
199001838	141) nell'accertamento dello stato di *abbandono* dei *minori*, non una prova determinante	1
		1
199002118	142) variamente incidere sulla situazione di *abbandono* del *minore*, nel senso di escluderla	1
	143) minore. Lo stato di *abbandono* del *minore* non potrebbe escludersi in relazione	1
	144) d'appello escludeva che l'*abbandono* del *minore* da parte della madre avesse	1
		3
199002397	145) alla sussistenza della situazione di *abbandono* del *minore* e delle circostanze che in	1
		1

199003307	146) volonta' del genitore di non *abbandonare* il *figlio*, quando nella realta' storica	1
	148) incontestabile pregressa situazione di *abbandono* del *minore* non avrebbe potuto	1
	149) sull'esistenza dello stato di *abbandono* del *minore*, perche' tra	1
	150) regolatrice, lo stato di *abbandono* del *minore* all'epoca della dichiarazione di	1
	151) dimostrare come lo stato di *abbandono* del *minore*, verificatosi all'epoca della	1
		5
199003365	152) merito, la situazione di *abbandono* dei *figli*. Con sentenza del 6	1
	153) ad evidenziare lo stato di *abbandono* dei *minori*. In particolare lo specifico	1
	154) conseguenze derivate dall'esperienza di *abbandono* cui i *minori* erano stati sottoposti	1
		3
199003369	147) 3) la *** aveva *abbandonato* suo *figlio* quando non aveva ancora un	1
	155) nella sentenza impugnata circa l'*abbandono* del *figlio* da parte della madre,	1
	156) sol per questo, *abbandono* dei *figli* come, invece,	1
	157) circa il preteso stato di *abbandono* del *figlio* da parte della Inriieri agli	1
		4
199006389	159) in ordine allo stato di *abbandono* della *minore*; ha trascurato di collegare	1
		1
199006494	158) che li avevano portati ad *abbandonare* i *figli*, dei caratteri della loro	1
		1
199009312	160) materiale concretante lo stato di *abbandono* del *minore*, si procede ai sensi	1
		1
199009313	161) la irreversibilita' della situazione di *abbandono* dei *minori* Rodolfo e Mauro	1
		1
199009315	162) tutto dall'eventuale stato di *abbandono* del *minore*, laddove il regime giuridico	1
	163) straniero avesse come presupposto l'*abbandono* del *minore*. Sotto il primo profilo	1
	164) all'accertamento dello stato di *abbandono* del *minore*) il motivo e' senza	1
		3
199011726	165) ritenuto il sostanziale stato di *abbandono* del *minore* ad un distacco graduale	1
	166) precedente sotto il profilo dell'*abbandono* del *minore*, che in presenza	1
		2
199103353	10) di una sostanziale situazione di *abbandono* del *minore*. Tutti i suesposti motivi	1
		1
199104101	11) Grappa denunciava lo stato di *abbandono* dei *minori* Graziella e Federico, nati	1
	12) o no lo stato di *abbandono* dei *minori*, ha ricercato agli elementi	1
	13) elementi confermativi dello stato di *abbandono* dei *minori*, laddove e' evidente che	1
		3
199104528	14) richiamando, per il caso di *abbandono* che concerna il *minore* straniero, non il solo	1
		1
199105881	15) sia una precisa volonta' di *abbandonare* il *figlio* ma e' sufficiente che i	1
		1
199107357	17) dell'accertamento dello stato di *abbandono* dell'*adottata*, a norma dell'art	1
		1
199107358	18) non sussisteva lo stato d'*abbandono* del *minore*, attese le premure e	1
	19) della concreta situazione di *abbandono* del *minore* costituisce un mezzo efficace di	1
		2
199108541	16) espresso da questa Corte (*abbandonando* quello *adottato* in passato), e che	1
		1
199109912	20) abbia verificato lo stato di *abbandono* dei *minori*, laddove tale giudice ha	1
	21) una valutazione di merito sull'*abbandono* del *minore* e sull'opportunita' del suo	1
	22) specie, la situazione di *abbandono* dei *minori* adottandi; e deduce cio'	1
	23) esplicita valutazione dello stato di *abbandono* del *minore*. Il tutto, unitamente	1
	24) dell'autorita' giudiziaria straniera sull'*abbandono* di quel *minore*, e sull'opportunita' di	1
	25) con riferimento allo stato di *abbandono* del *minore* straniero il solo riferimento al	1
	26) il riconoscimento della situazione di *abbandono* di un *minore* in Italia, si puo'	1
	27) sussistenza dello stato di *abbandono* del *minore*, cosi' come delineato dalla	1
		8
199113110	28) circa le condizioni di *abbandono* del *minore*, svolgeva, fra l'	1
		1
199204211	29) perche' la situazione di *abbandono* della *minore* non era stata accertata con	1
	30) da una formale dichiarazione di *abbandono* del *minore*, essendo invece sufficiente che	1
	31) null'altro e' che un *abbandono* del *minore* e, quindi, proprio	1
		3
199204395	32) ritenuto esistente lo stato di *abbandono* delle *minori* e, per essere connessi	1
	33) genitore non deriva automaticamente l'*abbandono* del *minore*, sia perche' questa e'	1
	34) essere considerato in stato d'*abbandono* il *minore* che, pur messo in	1
	35) una oggettiva situazione di *abbandono* delle *minori* nonostante l'attaccamento dei genitori	1
		4

199209273	36) di tempo dimostravano il concreto *abbandono* dei *minori* confermato dal fallimentare ritorno	1
	37) elementi indicativi della situazione di *abbandono* dei *figli* dovuta a causa non transitoria	1
		2
199211604	38) limitata ai soli casi di *abbandono* del *minore* da parte di uno dei	1
	39) della norma ai casi di *abbandono* del *minore* da parte di uno dei	1
		2
199212241	40) all'esistenza dello stato di *abbandono* dei *figli* minori, avendo basato il	1
		1
199301738	41) che quindi lo stato di *abbandono* del *minore* fosse imputabile ad una causa	1
	52) che quindi lo stato di *abbandono* del *minore* fosse imputabile ad una causa	1
		2
199302971	42) la identificazione dello stato di *abbandono* della *minore*, ed altresì in ordine	1
	43) lumeggiare la precedente situazione di *abbandono* del *minore*. Pertanto, l'avere	1
		2
199304050	44) nuova indagine sullo stato di *abbandono* del *figlio* da parte della madre.	1
		1
199304151	45) giudizio circa lo stato di *abbandono* dei *minori* su affermazioni meramente tautologiche,	1
		1
199304544	46) cioè, dello stato di *abbandono* del *minore*). E non v'è	1
		1
199307776	47) la sussistenza dello stato di *abbandono* della *minore* non dovuto a causa di	1
	48) dell'accertamento dello stato di *abbandono* della *minore*, mentre avrebbe dovuto disporre	1
		2
199308360	49) rilevando che lo stato di *abbandono* del *minore* non era imputabile alle condizioni	1
		1
199310011	50) di accertare lo stato di *abbandono* del *minore* e, di conseguenza,	1
		1
199312701	51) genitoriale: così il volontario *abbandono* del *minore*, posto in essere,	1
		1
199403044	53) non esiste la situazione di *abbandono* del *minore* Giancarlo *** che è	1
	54) la prova dello stato di *abbandono* del *minore*, che è frutto di	1
	55) infondato. La situazione di *abbandono* di un *minore*, e cioè la assenza	1
		3
199403588	56) decreto dichiarativo dello stato di *abbandono* del *minore* non compete al genitore che	1
		1
199403810	57) considerando comprovato lo stato di *abbandono* della *minore*, essendo esclusa la	1
	58) che vi era stato oggettivo *abbandono* della *minore* da parte di Serenella ***	1
		2
199404643	59) sussistente lo "stato di *abbandono*" della *minore* Giovanna *** violando,	1
	60) dello "stato d'*abbandono*" del *minore* e preordinato, quindi,	1
	61) meno dello "stato di *abbandono*" del *minore*, presupposto per la dichiarazione	1
		3
199409278	62) necessario che lo stato di *abbandono* del *minore* risulti in modo esplicito dal	1
	63) omessa menzione dello stato di *abbandono* dei *minori* dei quali erano indicati i	1
		2
199500078	64) e che lo stato di *abbandono* della *minore* si ricavava dalla prodotta documentazione	1
		1
199511054	65) genitori non si interessano o *abbandonano* i *figli*, ma anche quando con	1
	66) L'oggettivo stato di *abbandono* della *minore*, non determinato da cause	1
		2
199513100	67) affermato che la situazione di *abbandono* della *minore* derivava dalla totale anaffettività	1
		1
199600066	68) non esclude la possibilità di *abbandonare* conclusioni precedentemente *adottate*,	1
		1
199602236	69) avrebbe dichiarato lo stato di *abbandono* della *minore*. Il padre, infatti	1
		1
199607139	71) gli accertamento sulla situazione di *abbandono* della *minore* non comportava nullità del	1
		1
199607141	72) quanto alla situazione di *abbandono* del *minore* prima del suo ricovero in	1
	73) di accertamento della situazione di *abbandono* del *minore*, era emersa la loro	1
		2
199607618	70) avendo il padre naturale *abbandonato* la *figlia*, la situazione della minore	1
	74) utilizzato, facendo riferimento all'*abbandono* della *figlia* da parte del padre,	1
		2
199609576	75) sia una precisa volontà di *abbandonare* il *figlio*, ma è sufficiente che	1
	77) per insussistenza dello stato di *abbandono* nei confronti delle *minori*, né da parte dei	1
		2

199610656	76) nella specie lo stato di *abbandono* della *minore*. 4. Avverso quest'	1
		1
199702404	78) palesando l'irrecuperabile stato di *abbandono* del *figlio*. La Corte escluse,	1
	79) palesando l'irrecuperabile stato di *abbandono* del *figlio* <sup>m</sup> . Conclusivamente, la Corte	1
		2
199704054	81) proprio convincimento in relazione all'*abbandono* della *minore* che per molti anni e'	1
		1
199704619	80) sia una precisa volonta' di *abbandonare* il *figlio*, ma e' sufficiente che	1
	82) mera presuntivita' dello stato di *abbandono* delle due *minori*, dopo aver premesso che	1
	83) avrebbe presunto lo stato di *abbandono* delle due *minori* senza averlo potuto accertare	1
		3
199705130	84) e che lo stato di *abbandono* della *minore* si ricavava dalla prodotta documentazione	1
		1
199706853	85) che egli non ha mai *abbandonato* il *figlio* in quanto il periodo di	1
		1
199706899	86) in relazione alle ipotesi di *abbandono* del *minore* dalle quali emerge l'interesse	1
	87) secondo cui, l'*abbandono* del *minore* da parte dei genitori,	1
		2
199706927	88) ricorrenti, dello stato di *abbandono* dei *minori*, ma invece stabilisce una	1
		1
199707128	89) dell'insussistenza dello stato di *abbandono* del *minore* che della dichiarazione di adottabilità	1
		1
199710935	90) manifestino una precisa volontà di *abbandonare* il *figlio* ma anche quando con il	1
		1
199800120	91) escluso lo stato di *abbandono* della *minore* e lo avesse ritenuto sussistente	1
	92) lo stato di preteso *abbandono* della *minore* sia rapportabile all'una od	1
		2
199802672	93) inconfutabilmente dimostrato lo stato di *abbandono* della *minore* : in particolare	1
	94) ad escludere la situazione di *abbandono* del *minore*, e' infondato. Costituisce	1
		2
199803326	95) dato adito a sospetti di *abbandono* del proprio *figlio*. Almeno nei confronti di	1
		1
199806035	96) Torino di aver affermato l'*abbandono* del *minore* sulla scorta di considerazioni astratte	1
		1
199809500	97) volonta' del genitore di non *abbandonare* il *figlio*, ove nella realta' manchi	1
	98) alcun dubbio una situazione di *abbandono* della *minore*; che anche	1
	99) riscontro di una situazione di *abbandono* della *minore*, la forza maggiore,	1
	100) possa ovviare allo stato di *abbandono* della *figlia* minore con l'aiuto dei	1
	101) non per lo stato di *abbandono* del *minore*, con la conseguenza che	1
	102) un accertamento dello stato di *abbandono* del *minore*: tale valutazione e',	1
		6
199811241	103) precisa ed esplicita volontà di *abbandonare* il *figlio* da parte dei genitori,	1
	104) avendo ravvisato lo stato di *abbandono* della *minore* nella conferma dell'accertamento,	1
	105) a conferma dello stato di *abbandono* della *minore*, che l'affetto nutrito	1
		3
199811422	106) stato dichiarato lo stato di *abbandono* della *minore*, oltre la decadenza dei	1
		1
199907504	108) la presenza dello stato di *abbandono* della *minore* diffusamente e sotto vari aspetti	1
	109) essenzialmente sull'accertato stato di *abbandono* della *minore*, come evidenziato	1
		2
199909643	110) premesso che lo stato di *abbandono* del *minore* era stato affermato dal giudice	1
		1
199910809	107) precisa volontà dei genitori di *abbandonare* il *figlio*, essendo sufficiente l'oggettiva	1
		1
199912360	111) l'esistenza della situazione di *abbandono* in pregiudizio del *minore* stesso. Bastera', al	1
		1
199912449	112) l'esistenza della situazione di *abbandono* del *minore* rispetto alla madre e alla	1
		1
199913262	113) di un attuale stato di *abbandono* dei *minori*. La censura, che	1
	114) la permanenza dello stato di *abbandono* dei *minori*. Con il terzo motivo	1
	115) data carico dello stato di *abbandono* dei *minori*, instaurando con essi rapporti	1
		3
199913311	116) e' la situazione di *abbandono* in cui il *minore* versi, per essere egli	1
		1
199913419	117) impugnata ribadito lo stato di *abbandono* della *minore* disconoscendo il suo interesse a	1
		1
199914675	118) che la pretesa condizione di *abbandono* della *minore*: a) si "	1

200001095	119) richiedendo una situazione di *abbandono* del *minore* non transeunte per la dichiarazione	1
	120) verifica dell'effettivo stato di *abbandono* del *minore*), con gli ampi poteri	1
		2
200001100	121) incidere concretamente sullo stato di *abbandono* della *minore*. Avverso tale sentenza,	1
		1
200001612	167) in materia di stato d'*abbandono* dei *minori* (v. in particolare	1
	168) esistenza di uno stato di *abbandono* del *minore* per la grave ed irreversibile	1
		2
200005417	122) la verifica dello stato di *abbandono* del *minore* David *** nato il	1
		1
200005580	123) verificare lo stato di effettivo *abbandono* del *minore*, mancando al riguardo qualsiasi	1
	124) minore. Lo stato di *abbandono* del *minore* sarebbe, pertanto, insussistente	1
	125) a sussistere lo stato di *abbandono* del *minore*, configurandolo dal punto di	1
		3
200009802	126) che gli appellanti mai avevano *abbandonato* il *minore*, rimasto sempre legato a	1
	127) tale situazione non dimostrasse l'*abbandono* del *minore*, essendo compatibile con la	1
	128) di legittimita', motivando sull'*abbandono* del *minore* per mancanza di assistenza morale	1
	129) l'accertamento della situazione di *abbandono* permanente del *minore* ed evidenziandosi	1
		4
200012491	130) il risalente stato di *abbandono* del *minore*; l'inaffidabilità della madre	1
		1
200101714	7) di una madre che aveva *abbandonato* in solitudine la *figlia* minore nel paese d'origine	1
	8) ricorrente, la quale avrebbe *abbandonato* la *figlia* minore sola in Marocco non	1
	9) potendosi consentire che il genitore *abbandoni* il *figlio* minore nel proprio paese;	1
		3
200102010	5) che ricorreva una situazione di *abbandono* del *minore*, configurabile anche nell'ipotesi	1
	6) la declaratoria dello stato di *abbandono* del *minore*. Sostiene la *** che	1
		2
200104574	4) mancando lo stato di *abbandono* della *minore*, per avere gli oppositori	1
		1
200105547	1) ad accertare lo stato di *abbandono* della *minore* e la conseguente possibilita' di	1
	2) escluso lo stato di *abbandono* della *minore* sul presupposto che la madre	1
	3) ha escluso lo stato di *abbandono* della *minore* stessa da parte della madre	1
		3



PROTÓTIPO DE ANÁLISIS TEXTUAL  
DE LAS SENTENCIAS DE LA  
*CORTE DE CASACIÓN.*  
EXTRACCIÓN AUTOMÁTICA DE  
DESCRIPTORES

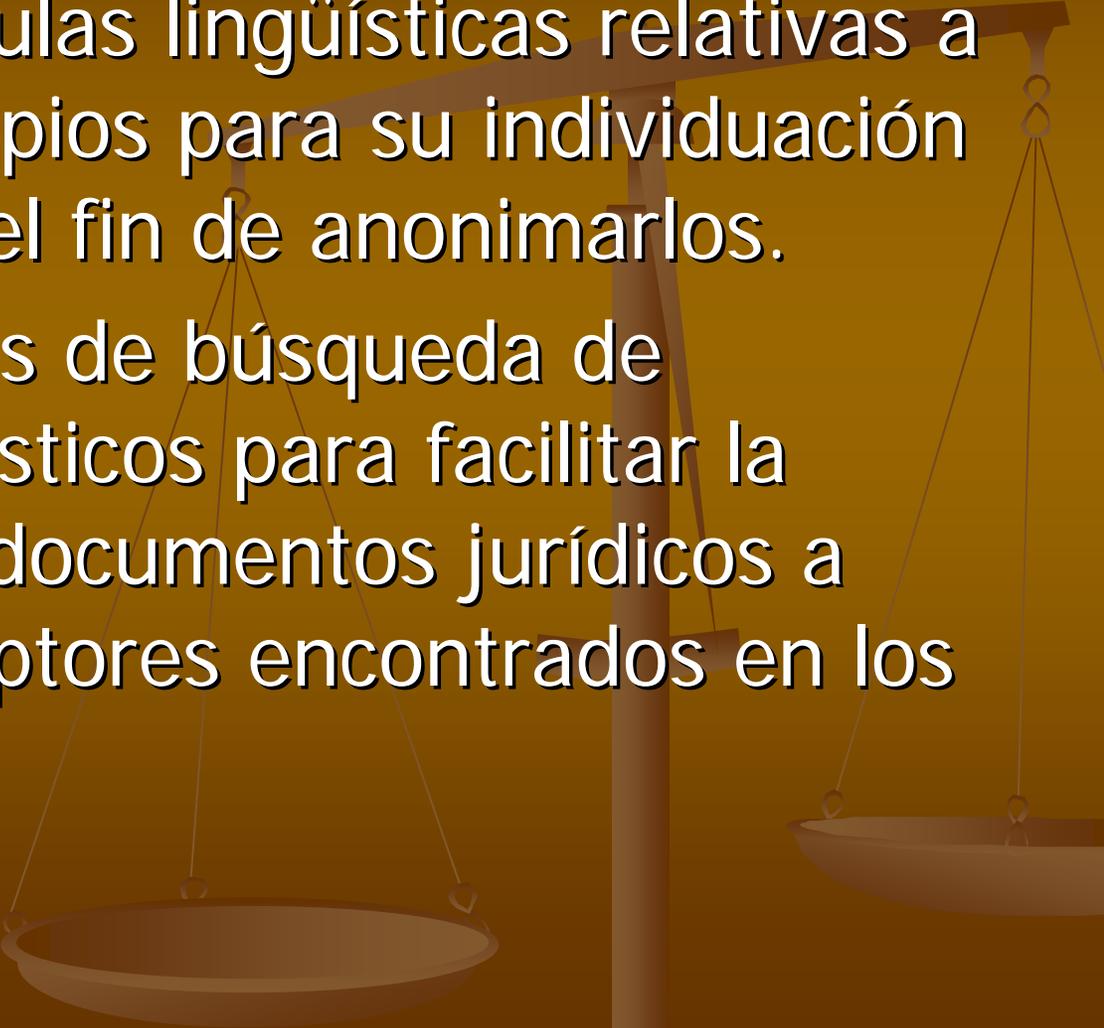
*Manuela Sassi, Alessandra Cinini*

*Istituto di Linguistica Computazionale  
del CNR di Pisa*

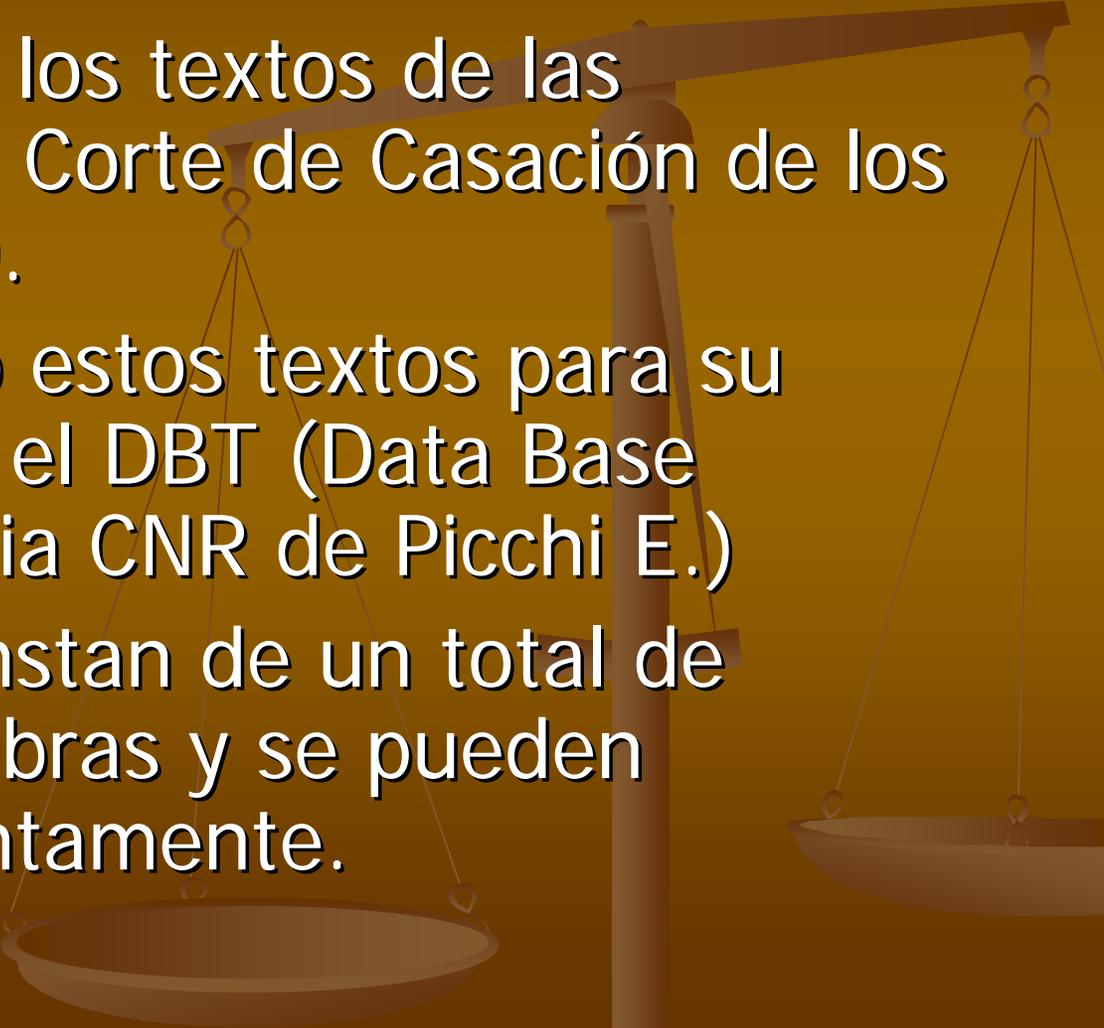
# Introducción

- En el ámbito del proyecto "Tecnologías para la justicia", coordinado por el *Istituto di Ricerca sui Sistemi Giudiziari* del CNR di Bologna (IRSIG), el *Istituto di Linguistica Computazionale* del CNR (ILC), ha brindado instrumentos, ya experimentados en la investigación lingüística, para la extracción de informaciones lingüístico-conceptuales desde grandes cantidades de datos textuales.

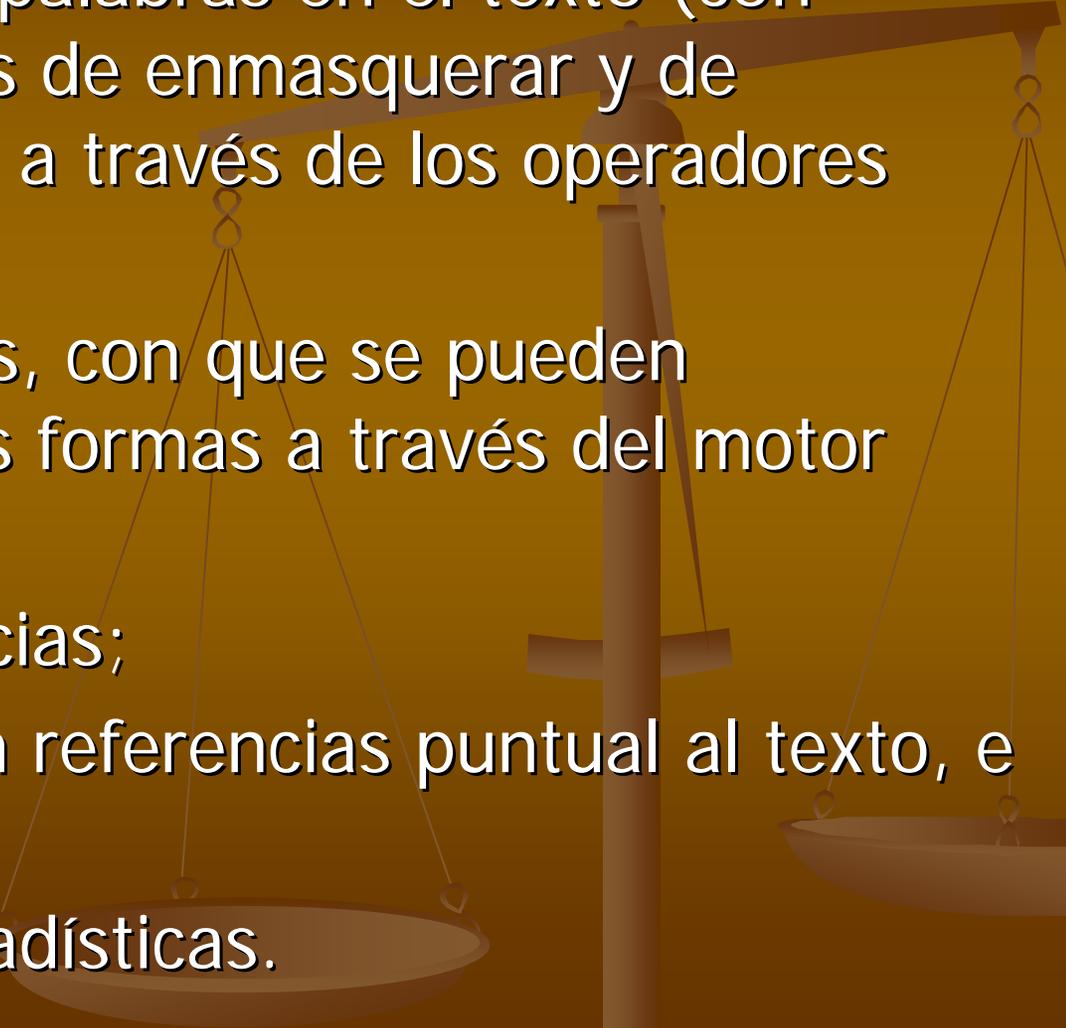
# OBJETIVOS

- Estudio de fórmulas lingüísticas relativas a los nombres propios para su individuación en el texto con el fin de anonimarlos.
  - Estudio de reglas de búsqueda de esquemas lingüísticos para facilitar la clasificación de documentos jurídicos a través de descriptores encontrados en los mismos textos.
- 

# LA BASE INICIAL DE DATOS TEXTUALES

- Se han utilizado los textos de las sentencias de la Corte de Casación de los años 1990-2000.
  - Se han marcado estos textos para su elaboración con el DBT (Data Base Testuale, Licencia CNR de Picchi E.)
  - Los archivos constan de un total de 92.481.142 palabras y se pueden consultar conjuntamente.
- 

# CARACTERÍSTICAS DEL DBT

- Búsqueda libre de palabras en el texto (con varias posibilidades de enmascarar y de combinar términos a través de los operadores booleanos);
  - búsqueda de lemas, con que se pueden encontrar todas las formas a través del motor morfológico;
  - cálculo de frecuencias;
  - concordancias, con referencias puntual al texto, e index locorum;
  - co-ocurrencias estadísticas.
- 

# ANÁLISIS AUTOMÁTICO

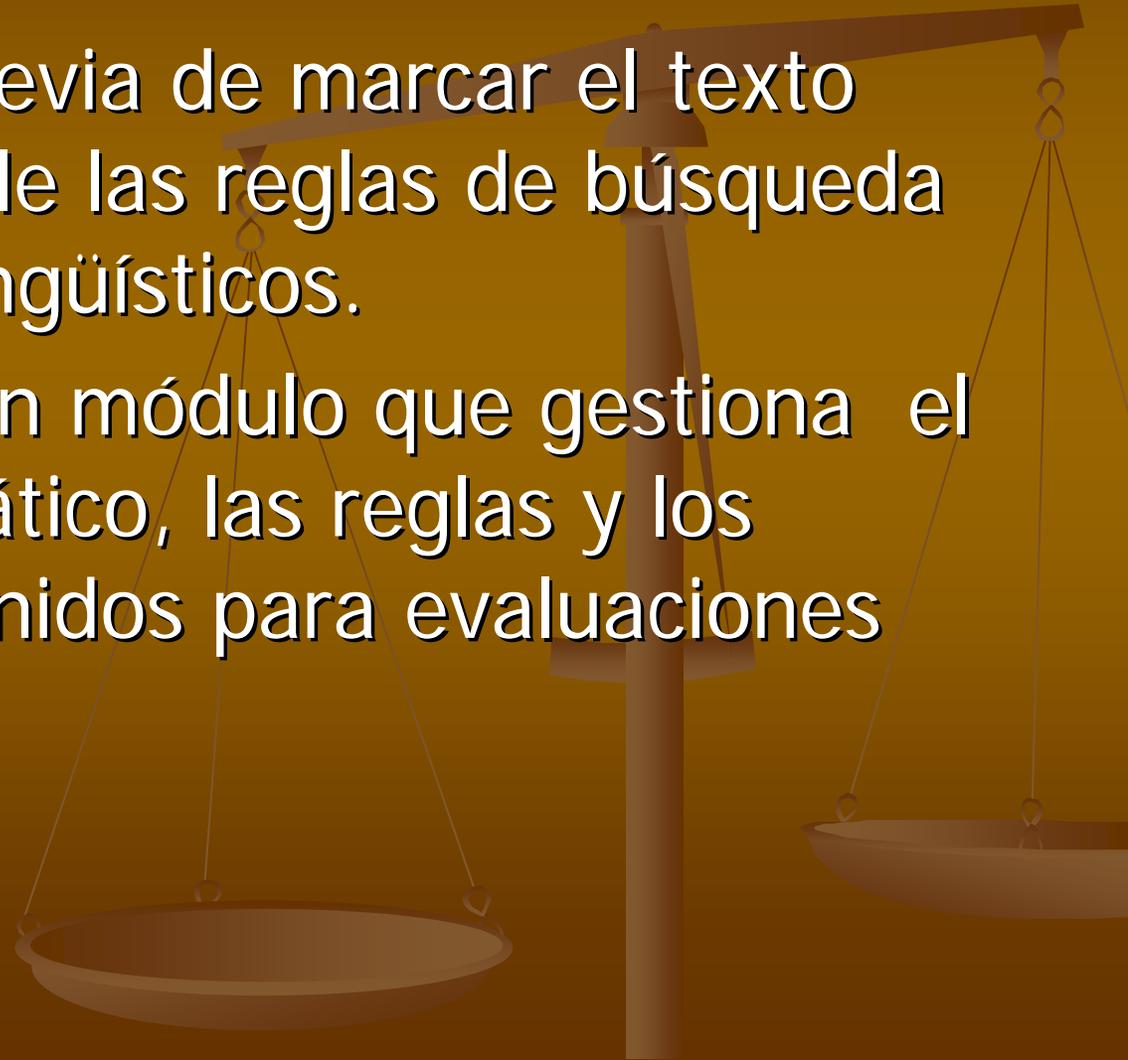
- Los archivos DBT han sido totalmente lematizados con la función de análisis automático, que asigna a cada palabra el *lema*, o sea la llave de entrada en el diccionario, las categorías gramaticales de cada lema y de sus formas, según su papel en el contexto dado.
- El producto de esta fase de análisis se archiva para ser procesado posteriormente, aplicándole las reglas de búsqueda de esquemas lingüísticos.



2124	LA	LA	R	FS
2125	RICORRENTE	RICORRENTE	SN	NS
2126	LAMENTA	LAMENTARE	VTP	S3IP
2127	CHE	CHE	C	
2128	LA	LA	R	FS
2129	SENTENZA	SENTENZA	SF	FS
2130	IMPUGNATA	IMPUGNATO	A	FS
2131	NON	NON	BS	
2132	CONTERREBBE	CONTENERE	VTRP	S3DP
2133	UNA	UNO	RI	FS
2134	ADEGUATA	ADEGUATO	A	FS
2135	ESPOSIZIONE	ESPOSIZIONE	SF	FS
2136	DELLE	DI	E	FP
2137	RAGIONI	RAGIONE	SF	FP
2138	DEL	DI	E	MS
2139	DECIDERE	DECIDERE	VTIP	F

# Búsqueda de descriptores (1)

- La operación previa de marcar el texto permite aplicarle las reglas de búsqueda de esquemas lingüísticos.
- Para esto hay un módulo que gestiona el análisis automático, las reglas y los resultados obtenidos para evaluaciones estadística.



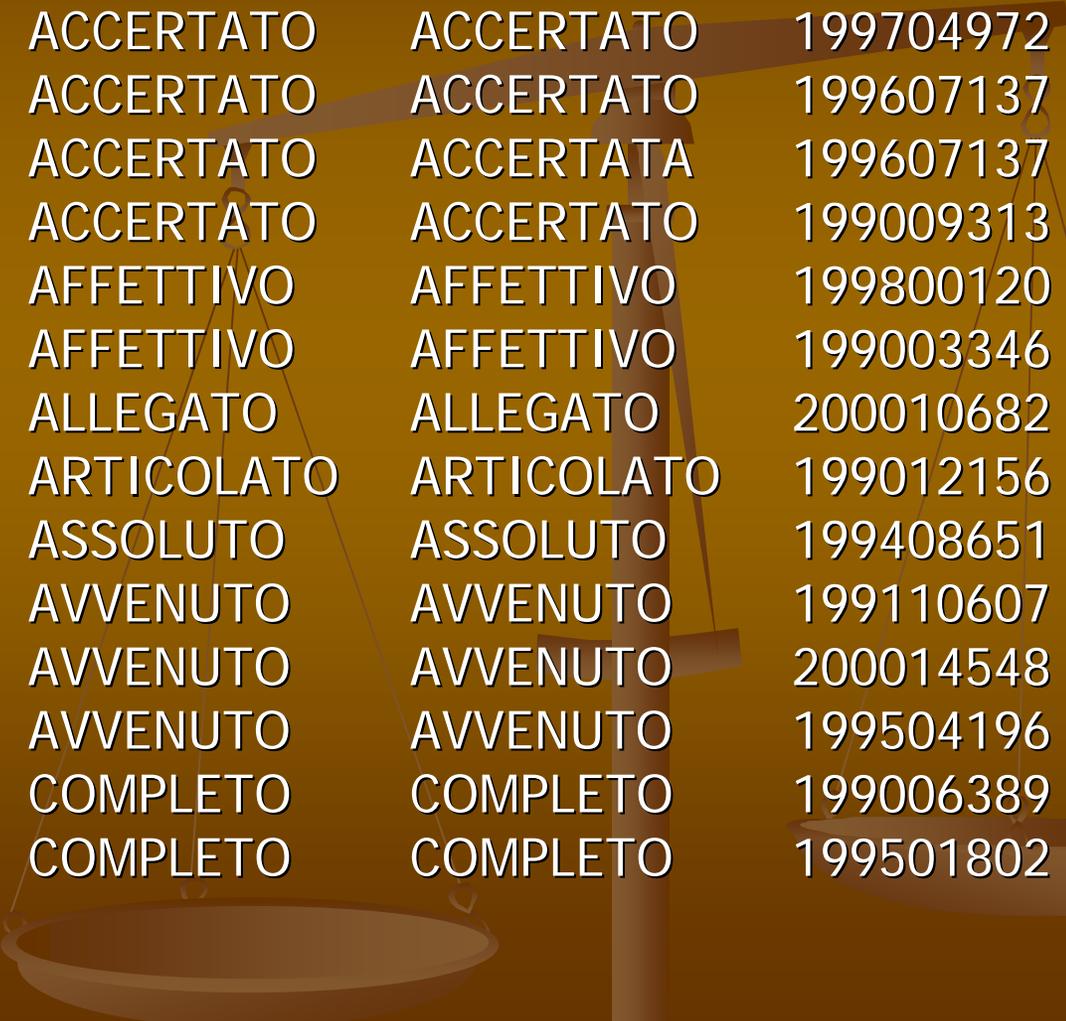
# Búsqueda de descriptores (2)

- In particular se ha implementado una regla para encontrar los sintagmas formados por dos sustantivos conectados por la preposición "DE". Por ejemplo en la frase de la tabla anterior resulta el caso de "ESPOSIZIONE (DI) RAGIONE"
- Otra regla permite de recuperar los sintagmas formados por sustantivo y adjetivo. Ej.: SENTENZA IMPUGNATA.

# PRIMEROS RESULTADOS (1)

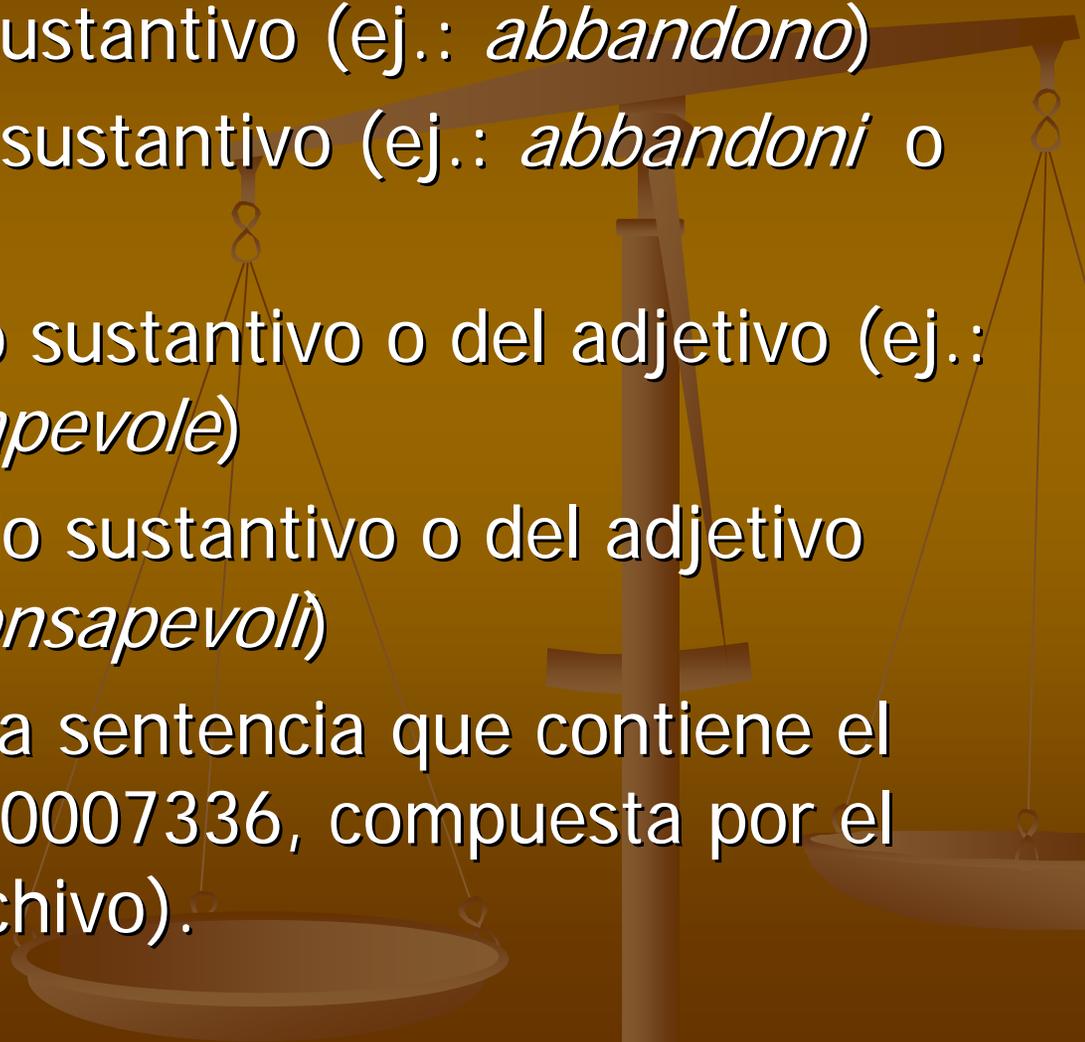
ABBANDONO	ABBANDONO	ABITAZIONE	ABITAZIONE	0200002217
ABBANDONO	ABBANDONO	ABITAZIONE	ABITAZIONE	0199712829
ABBANDONO	ABBANDONO	ADOTTATO	ADOTTATA	0199107357
ABBANDONO	ABBANDONO	AFFARE	AFFARE	0199804906
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199700999
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199700999
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199700999
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199700999
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199100159
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199100556
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199608830
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199512242
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199403514
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199403514
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199713045
ABBANDONO	ABBANDONO	ALLOGGIO	ALLOGGIO	0199713045

# PRIMEROS RESULTADOS (2)

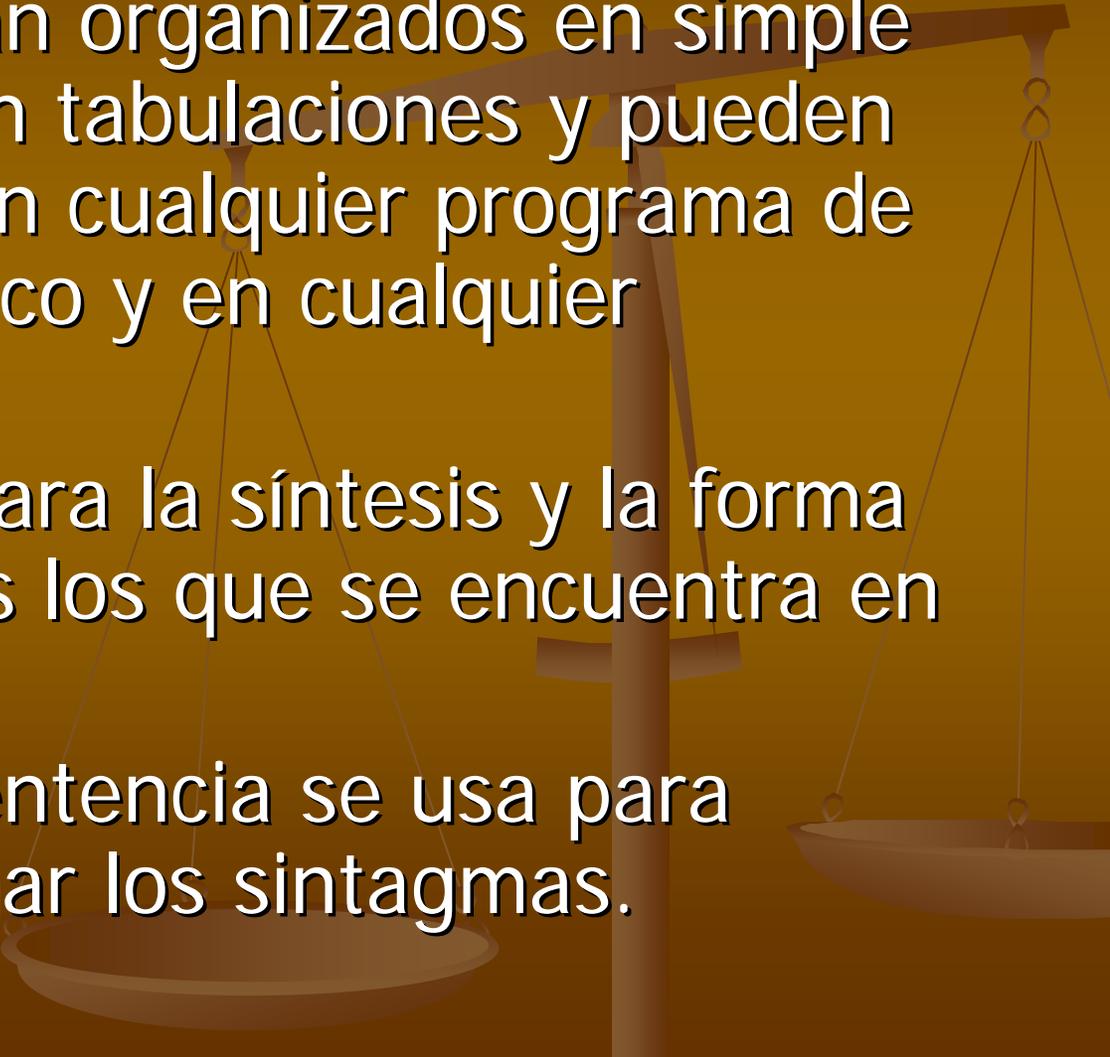


ABBANDONO	ABBANDONO	ACCERTATO	ACCERTATO	199704972
ABBANDONO	ABBANDONO	ACCERTATO	ACCERTATO	199607137
ABBANDONO	ABBANDONO	ACCERTATO	ACCERTATA	199607137
ABBANDONO	ABBANDONO	ACCERTATO	ACCERTATO	199009313
ABBANDONO	ABBANDONO	AFFETTIVO	AFFETTIVO	199800120
ABBANDONO	ABBANDONO	AFFETTIVO	AFFETTIVO	199003346
ABBANDONO	ABBANDONO	ALLEGATO	ALLEGATO	200010682
ABBANDONO	ABBANDONO	ARTICOLATO	ARTICOLATO	199012156
ABBANDONO	ABBANDONO	ASSOLUTO	ASSOLUTO	199408651
ABBANDONO	ABBANDONO	AVVENUTO	AVVENUTO	199110607
ABBANDONO	ABBANDONO	AVVENUTO	AVVENUTO	200014548
ABBANDONO	ABBANDONO	AVVENUTO	AVVENUTO	199504196
ABBANDONO	ABBANDONO	COMPLETO	COMPLETO	199006389
ABBANDONO	ABBANDONO	COMPLETO	COMPLETO	199501802

# DESCRIPCIÓN DE LOS DATOS

- 1 lema del primer sustantivo (ej.: *abbandono*)
  - 2 forma del primer sustantivo (ej.: *abbandoni* o *abbandono*)
  - 3 lema del segundo sustantivo o del adjetivo (ej.: *alloggio* o *consapevole*)
  - 4 forma del segundo sustantivo o del adjetivo (es.: *alloggi* o *consapevoli*)
  - 5 Identificativo de la sentencia que contiene el bigrama (ej.: 1990007336, compuesta por el año y el N. de archivo).
- 

# FUNCIÓN DE LOS DATOS

- Estos datos están organizados en simple fichero-texto con tabulaciones y pueden ser utilizados con cualquier programa de análisis estadístico y en cualquier plataforma.
  - El lema es útil para la síntesis y la forma es útil porque es los que se encuentra en el texto.
  - El número de sentencia se usa para agrupar u ordenar los sintagmas.
- 

## *ABBANDONO DI:*

POSTO	47
ATTIVITA'	29
DOMANDA	21
ALLOGGIO	19
CASA	18
DIRITTO	11
DOMICILIO	11
CANTIERE	8
FIGLIO	6
FONDO	6
TETTO	6
ALVEO	4
CAUSA	4
IMMOBILE	4

## *ABBANDONO:*

VOLONTARIO	16
DEFINITIVO	12
MORALE	10
MATERIALE	9
TOTALE	7
EFFETTIVO	6
COMPLETO	5
ACCERTATO	4
CONSEQUENTE	4
IMMEDIATO	4
PRIVO	4
RITENUTO	4
SOSTANZIALE	4
TEMPORANEO	4





199001536	ABBANDONO di ATTIVITA'	4
	ABBANDONO di LAVORO	3
	ABBANDONO di POSTO	1
199001801	ABBANDONO di GEMELLO	1
	ABBANDONO di MINORE	3
199001861	ABBANDONO di POSTO	1
199001885	ABBANDONO di ATTIVITA'	1
199001892	ABBANDONO di ATTIVITA'	4
	ABBANDONO di LAVORO	3
	ABBANDONO di POSTO	1
199001931	ABBANDONO di ALLOGGIO	1
	ABBANDONO di APPARTAMENTO	1

# CONCLUSIONES

- Los archivos ya marcados permiten la aplicación de otras reglas en relación a otros objetivos más complejos, como por ejemplo son los temas de la jurimétrica para evaluar las sentencias desde el punto de vista *reato-pena*.
  - Se quiere subrayar como este tipo de trabajo, no obstante su especificidad, pueda contribuir al desarrollo de la TICs para todos.
- 