

ALCIDE

Analysis of Language and Content In a Digital Environment

Giovanni Moretti, Sara Tonelli, Stefano Menini, Rachele Sprugnoli Digital Humanities Joint Research Project – http://dh.fbk.eu Fondazione Bruno Kessler, Trento



WHAT

Web-based platform for Historical Content Analysis of contemporary history textual sources



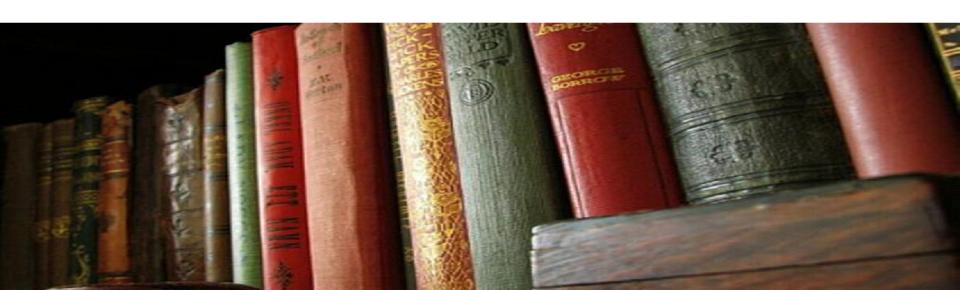
HOW

Combination of Human Language
Technologies and interactive visualizations
in a digital environment tailored to the
needs of historians

Topic segmentation and recognition Parsing Sentence breaking Part-of-speech tagging Morphological segmentation Part-of-speech tagging Named entity recognition Named entity recognition Automatic summarization Temporal Information Processing Relationship extraction Sentiment analysis Language identification

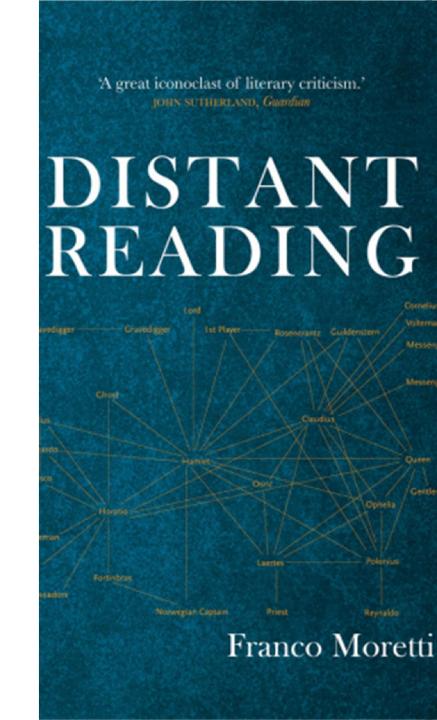
WHY

DH is the field in which Humanities and Human Language Technologies can interact and support each other



- Distant + Closed readings
- Time bounded information
- Flexibility
- Ease of use

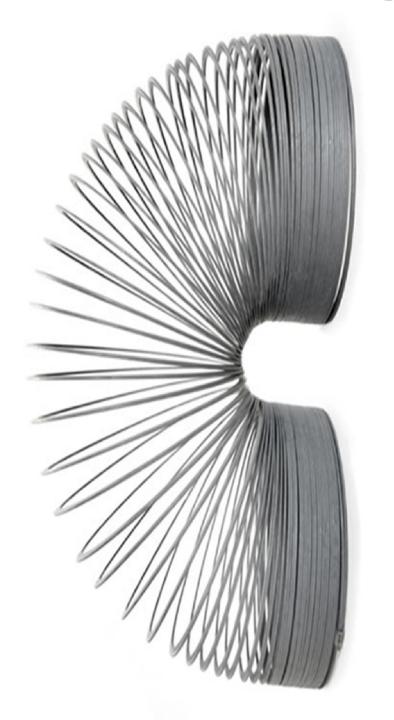
Multilinguality



- Distant + Closed readings
- Time bounded information
- Flexibility
- Ease of use
- Multilinguality



- Distant + Closed readings
- Time bounded information
- Flexibility
- Ease of use
- Multilinguality



- Distant + Closed readings
- Time bounded information
- Flexibility
- Ease of use
- Multilinguality



- Distant + Closed readings
- Time bounded information
- Flexibility
- Easy of use
- Multilinguality



Case Studies

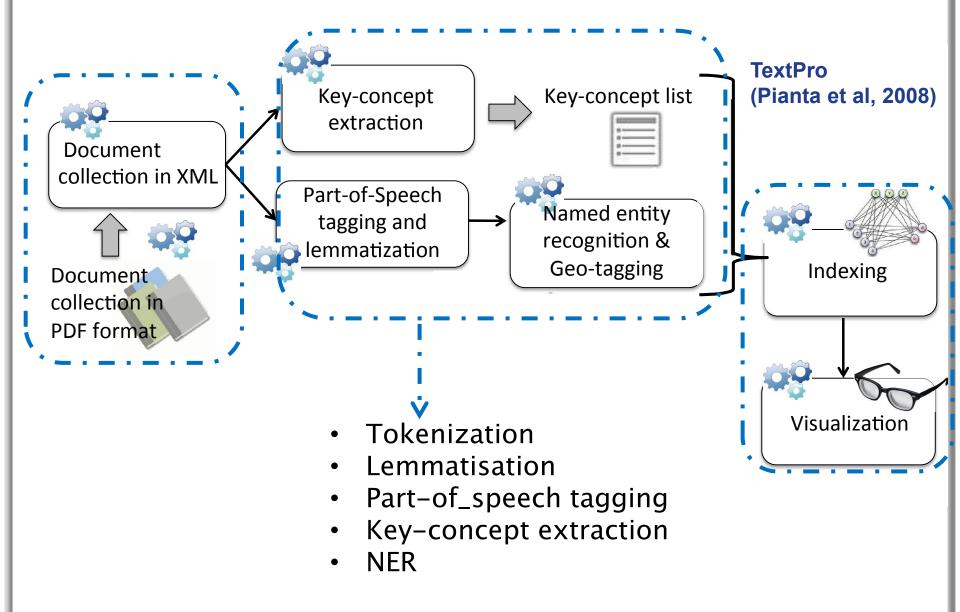
- Complete collection of Alcide De Gasperi's writings (1901 – 1954)
- Around 3,000 documents,
 3,000,000 words





- 1960 USA Presidential campaign speeches
- More than 800,000 words for both Kennedy and Nixon

Text Analysis Workflow



Key-concepts and Named Entities

- Key-concept extractor
 - preference to abstract nouns
 - preference to short key-concepts

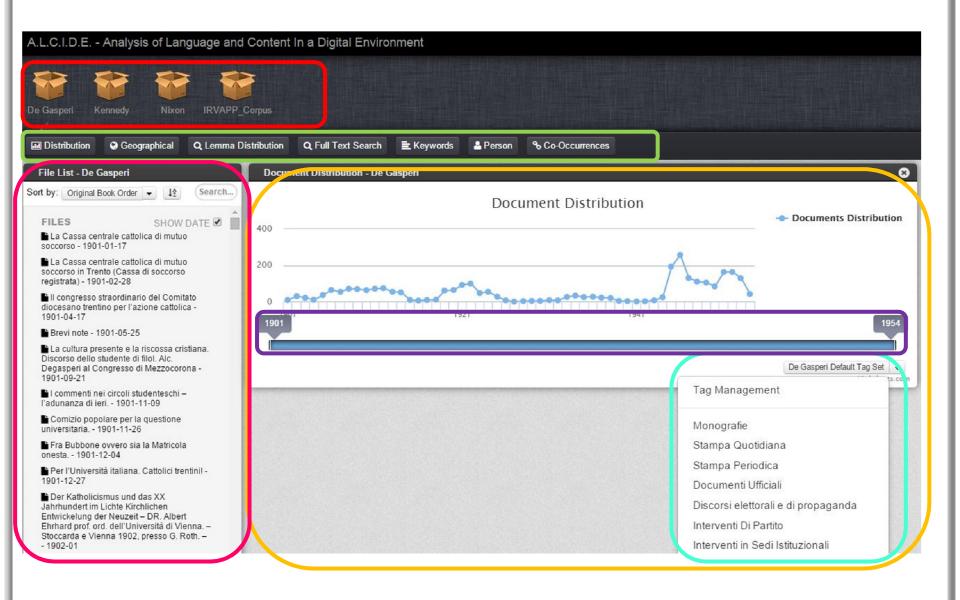
NER module F1 score before filtering

EVALITA 2007 → Sample of De Gasperi's writings

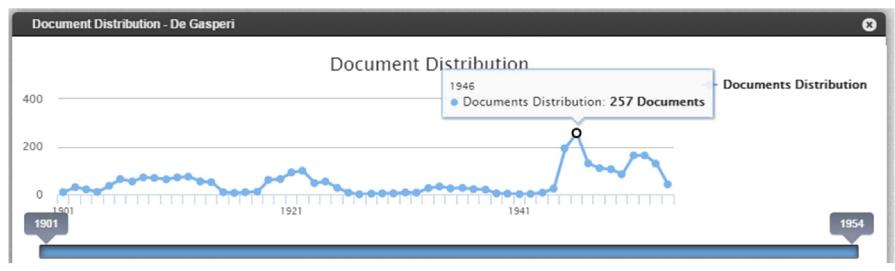
PER 92.12 PER 75.75

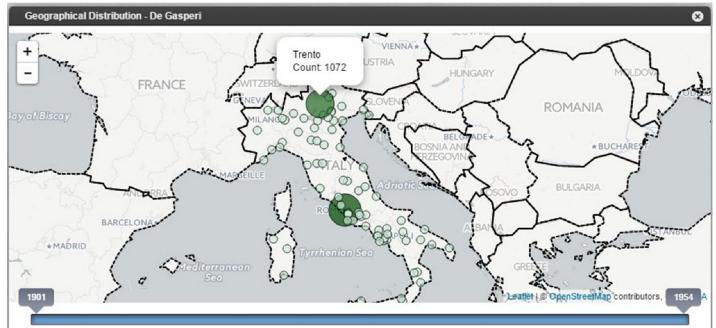
GPE 85.54 GPE 86.23

Graphical Interface

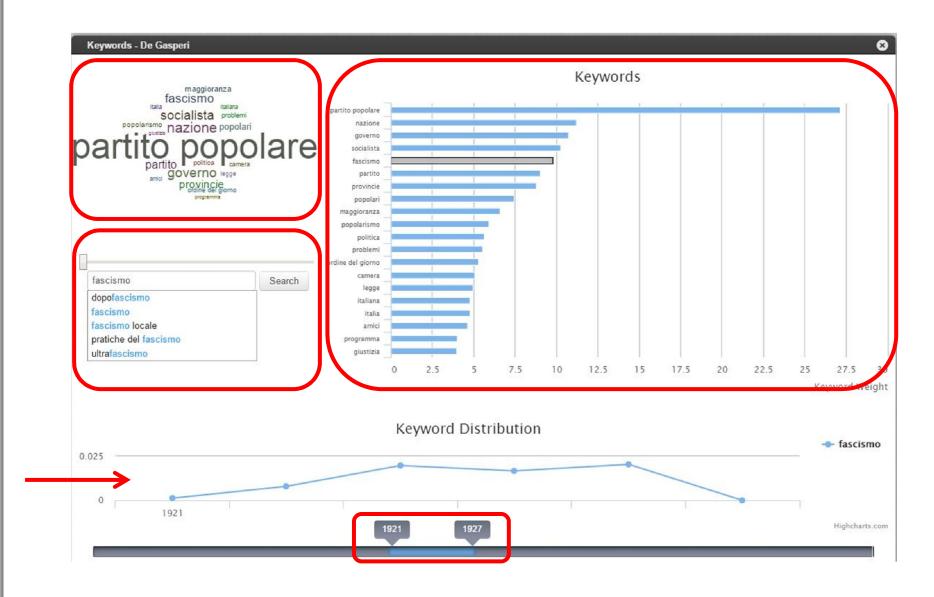


Metadata Visualization

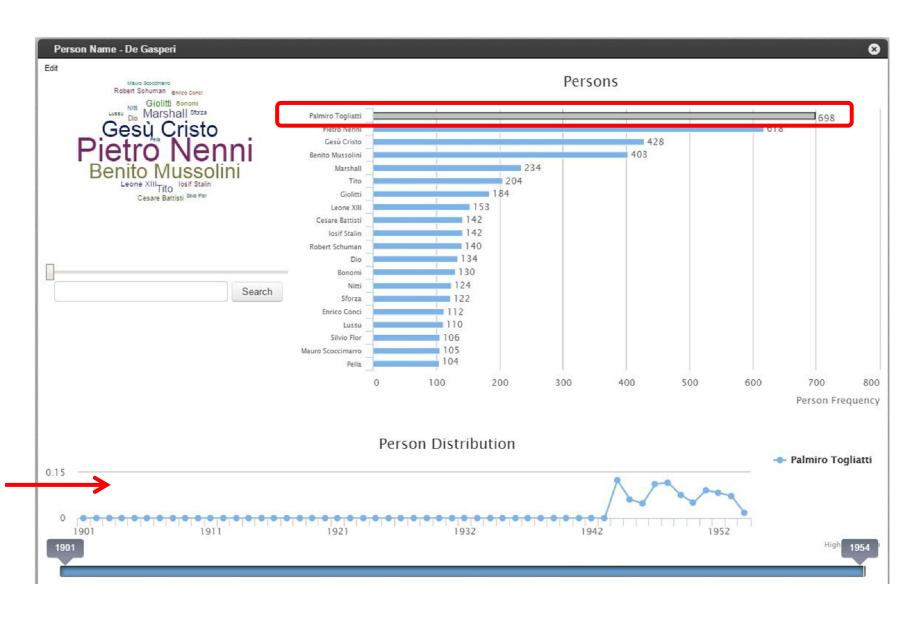




Key-Concept Extraction



Named Entity Recognition



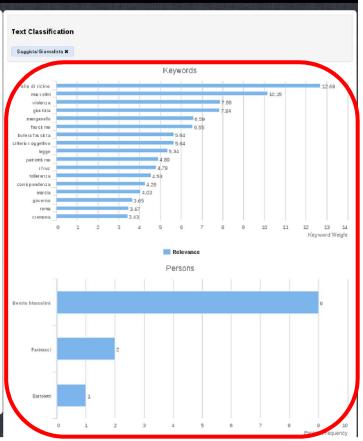
Close Reading

A.L.C.I.D.E. - Analysis of Language and Content In a Digital Environment

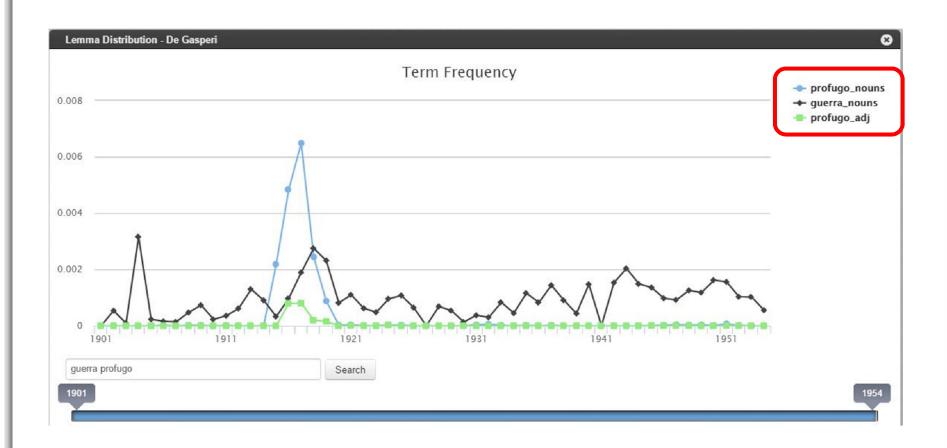
La marcia di...Sfruz

Trento,1923-03-30

Quando Mussolini o i suoi generali richiamano alla disciplina e condannano con gradita precisione le violenze degli eroi della sesta giornata, nella stampa fascista e nelle riunioni di partito si risponde con un atto di sottomissione: qui e là però gli avvenimenti ed il linguaggio che riprende poco dopo ci dicono che la sottomissione è formale, ma non nello spirito. Tuttavia che ciò accada nelle provincie per le quali la marcia di Roma fu o doveva essere la fi ne di tutto un lungo periodo di conflitti sanguinosi, è ancora spiegabile, perché l'uso ripetuto e prolungato della violenza produce una perturbazione delle coscienze così profonda che gli stessi esaltatori d'un metodo di eccezione da loro ritenuto necessario ma transitorio, né dianzi prevedevano né oggi sanno dominare. Ma tale spiegazione e il compatimento che ne deriva non si confanno a regioni che del fascismo hanno ignorata la epoca cruenta. Nella Venezia Tridentina, se si eccettuano le due imprese di Bolzano e di Trento, le quali ebbero stile e sviluppo militare e carattere antistatale, ma non furono rivolte contro l'incolumità delle persone, o di un gruppo di persone, nessun episodio grave si ebbe a deplorare, né per difesa né per offesa. Le risse domenicali non insanguinarono né i centri né i villaggi della provincia, cosicché la tradizione d'ordine, di quieta tolleranza, di rispetto alla legge, non subì interruzioni notevoli. Lo spirito di violenza non è quindi nel Trentino né endemico né epidemico. Se tuttavia qui e là fa capolino anche fra noi, vuol dire che vi viene importato di proposito e forzando il naturale sviluppo delle cose. Vestigia terrent e, per conto nostro, ci preoccupiamo che anche le faville più leggere non suscitino qualche incendio. Onde ci chiediamo se sia lecito che un giornale locale che fa professione di fascismo pubblichi minacce e intimidazioni che (in una corrispondenza da Tione) polemizzando con un o.d.g. di una assemblea di capifamiglia stampi bravate consimili: «Illustri capi di famiglia, fate il piacere di tacere anche e soprattutto perché oltre al Partito fascista che vostro malgrado oggi governa energicamente e sagacemente potrebbe svegliarsi quel santo manganello che vi farebbe ingoiare, ecc.». Come? Oltre il Governo, del quale vantate l'energia e la sagacia, invocate ancora il manganello? Oltre la forza della legge minacciate la violenza del vostro bastone? Sentite questa, che viene a proposito. Nella «Cremona Nuova» (che è pur l'organo di quell'agnellino di Farinacci) si è svolto un dibattito a proposito delle recenti sentenze dei tribunali contro i somministratori di olio di ricino. Nella discussione intervenne anche l'avv. Bartoletti che, messo in rilievo la severità dell'art. 154 del Cp., secondo il quale, infatti, chiunque usa violenza o minaccia per costringere qualcuno a fare, tollerare od omettere qualche cosa è punito con la reclusione fi no ad un anno e con la multa fi no a lire mille, ecc.; così conclude (e le sue parole scritte sul Cremona Nuova, dell'on. Farinacci hanno un valore anche maggiore) commentando le osservazioni di un fascista che trovava meno punite le manganellate dell'olio di ricino: «lo, tuttavia, ho osservato a quel giovine che non si possono misurare le consequenze dei colpi di manganello e quindi si possono produrre anche lesioni che sono punite con un minimo di cinque anni di reclusione, o la morte, e che, ad ogni modo, non ha più ragione d'essere né manganello né olio di ricino, oggi consule Mussolini. Ed è giusto per questo, a mio credere, che i magistrati hanno messo da banda la benignità, con la quale interpretavano la legge per l'addietro». Noi non condividiamo naturalmente il criterio soggettivo del collaboratore della Cremona Nuova, il quale, se Mussolini, non fosse al Governo, troverebbe forse inapplicabile per simili casi l'art. 154; ma il criterio soggettivo o quello oggettivo, dato che al Governo d'è proprio Mussolini, conduce alla stessa conclusione. Invece i signori di Tione - e, pare, gli altri che raccolgono la voce loro - pensano che consule Mussolini si debba tuttavia far ricorso al manganello il quale dovrebbe fornire una sanzione suppletoria a quella delle leggi, pur applicate con energia e con sagacia! E magari fra chi statuisce così osseguienti principii si distingue gualche impiegato governativo! La verità è che anche nelle valli più alpestri una strana emulazione sembra rivelarsi. Se Mussolini ha fatto la marcia su Roma, perché non vorremo noi intraprendere la marcia di ... Sfruz? Sfruz è un paesello di 500 abitanti circa con 108 ettari di agro comunale. Ora che cosa può essere accaduto su codesti 108 ettari di china montana e che razza di delinguenti alberga codesto pacifi co nido, se lo stesso giornale, ultimo numero, stampa da colà una corrispondenza che termina così: «Lottiamo per una causa giusta ed umana, e se nella legge non troveremo, una pronta giustizia – i fascisti italiani la faranno da loro. Roma fulcro di civiltà e progresso fu conquistata dal fascismo come lo furono tutti i comuni d'Italia. Sfruz è un granellino di sabbia in mezzo al mare che potrebbe essere travolta dalla bufera fascista che epurerà l'atmosfera satura di miasmi, d'ingiustizia e malvagità». Non sappiamo proprio di che si tratti, ne c'interessa di saperlo, poiché pensiamo che gli organi del governo Mussolini – carabinieri, sindaco, sottoprefetto, pretore – siano suffi cienti a salvaguardare la causa della giustizia e del patriottismo, quando fossero in pericolo. Ma quei giovanotti di lassù pensano diversamente e proclamano che se la legge non fa giustizia essi faranno giustizia fuori della legge, minacciando di travolgere nella bufera fascista quel miserabile «granellino di sabbia» che risponde al dolce nome di Sfruz, Come Roma..., tanto che dire: come Mussolini... Ehi, signori! Siete in arretrato. Mussolini ha detto e ripetuto che siamo già entrati nel secondo tempo, nel tempo cioè in cui, valendosi delle forze sociali e politiche dello Stato, si assimila, si rinvigorisce, si ricostruisce. Non si scatenano bufere, ma si lavora nell'ordine e nella disciplina, non si fa dello «stupido antieroico illegalismo» (parole sue), ma si dà forza e obbedienza alle leggi. Senonché non intendiamo ricordare questa paternale proprio per voi e solo a voi. Voi siete un caso fra tanti che abbiamo colto fra gli ultimi, ad illustrazione degli altri. Il monito vale per tutti. Il Trentino non è ancora guarito dalle piaghe della guerra e dei martiri che, durante la guerra sovratutto, gli fece soffrire lo straniero; deve oltre a ciò fare opera di adattamento alle nuove condizioni economiche - l'adattamento anche se per il meglio comporta sempre uno sforzo -, deve assuefarsi ad un nuovo tipo di fi nanza e di amministrazione; è tutto un lavoro questo che esige spirito di solidarietà e di fattivo patriottismo fi n nell'ultimo paesello di montagna, e a questo lavoro sono chiamati quei pochi che, superando l'attuale crisi economica nonché emigrare, possono dedicare qualche attività alla pubblica cosa. Ebbene questa è opera urgente di italianità ch'esige concordia, equità, tolleranza e presuppone un ambiente di pace. A tale opera noi diamo il nostro concorso e nessun altro ammonimento abbiamo ripetuto ai nostri, se non questo. Perché e a che pro devono risuonare ancora parole di odio e di minaccia? Oggi non si marcia, si lavora e si ricostruisce, per l'Italia, e per questa terra fi nalmente sua.



Advanced Search Functions



Future Work

- Upload new corpora
- Improve the precision of current functionalities
- Add new functionalities:
 - extraction of temporal expressions and events
 - sentiment analysis
 - semantic similarities



THANK YOU!

Email: sprugnoli@fbk.eu

Web Site: http://dh.fbk.eu

Twitter: @DH_FBK



