



## Tecnologie della lingua

Human Language Technology (HLT)



## Aspetto applicativo dell'elaborazione delle lingue

- Elaborazione delle lingue per affrontare le sfide tecnologiche della società



## Che cosa sono le tecnologie della lingua?

- Tecnologie che riguardano diversi aspetti della ingegneria della lingua: componenti software e strumenti per l'elaborazione delle lingue naturali
- Applicazione delle conoscenze del linguaggio per lo sviluppo di sistemi computazionali in grado di riconoscere, capire, interpretare o produrre una lingua in tutte le sue forme



## Quando è necessario usare la propria lingua?

- Per partecipare alla Società dell'Informazione in modo totalmente naturale;
- La lingua è un mezzo di comunicazione negli affari, nel tempo libero, nel governo e nella società civile in genere;
- La lingua è un mezzo per salvaguardare la propria identità culturale in una società globale caratterizzata dalla diversità culturale e linguistica;
- Essere in grado di fare affari interagendo in diverse lingue, diventa una necessità del commercio globale.





## Le tecnologie della lingua e le sfide della Società dell'Informazione

- La tecnologia della lingua permette agli esseri umani di comunicare con i computer utilizzando la propria lingua e di usarli in maniera più naturale
  - Interazione gradevole (user-friendly)
  - Multilingualità



## Le tecnologie della lingua sono importanti in Europa

- Diversità culturale e linguistica
  - 11 lingue ufficiali che saliranno a più di 20 con l'ingresso nella Comunità dei nuovi paesi candidati
- Aspirazione dei cittadini europei ad integrarsi rispettando profondamente la propria cultura, usi e costumi locali, e la propria lingua.
- La trasparenza della comunicazione (capacità e diritto a comprendere) diventa vitale, sia per i cittadini, che per i consumatori, i governanti e gli uomini d'affari, per partecipare in un ambiente che diventa sempre più denso dal punto di vista informativo e conoscitivo.
- Molti prodotti e servizi per la Società dell'informazione verranno costruiti intorno a componenti che elaborano le lingue.



## Alcuni tipi di applicazione

- Traduzione automatica
- Elaborazione di testi (Text processor)
- Analizzatori e generatori di linguaggio parlato
- Generazione di documenti
- Classificazione di documenti
- Recupero di informazione
- Estrazione di conoscenza da testi
- Apprendimento e insegnamento delle lingue



## Alcuni settori di applicazione

- Produzione di software
  - Sistemi di Information Retrieval
  - Text processors
  - Traduttori personali
  - Analizzatori e sintetizzatori di linguaggio parlato
- Settore bancario
  - Servizi di call centers, commercio elettronico, text mining
- Comunicazione
  - Servizi televisivi
  - Agenzie di stampa e di trasmissione di notizie di vario tipo
  - Servizi Internet
  - Reti cittadine
  - Servizi di prenotazione on-line o via rete telefonica fissa o mobile
- Servizi per l'apprendimento a distanza
  - Sistemi di Tutoring
- Editoria
  - Guide Turistiche Multimediali
  - Produzione di CD
- Enti pubblici
  - Comunicazione con il pubblico, organizzazione di archivi





## Produzione e gestione di risorse linguistiche

- Produzione di basi di conoscenza linguistica per facilitare lo sviluppo delle applicazioni delle tecnologie della lingua
  - Dizionari mono- e multilingue
  - Basi di conoscenza lessicali mono- e multilingue
  - Grandi *corpora* di testi o enunciati scritti o orali con annotazioni linguistiche di vario tipo: fonetiche, morfologiche, sintattiche, semantiche e stilistiche



NLP group



## Servizi di elaborazione linguistica

- Acquisizione automatica di testi
  - Con struttura complessa
    - Multi-font con annesse figure e grafici
  - In vari formati
    - Documenti
    - Libri
    - Cataloghi
- Servizi di traduzione
  - Traduzione assistita di documenti per il cliente
  - Traduzione automatica on-line
  - Consultazione di dizionari on-line, mono- e multilingue, sia generali che specialistici
  - Consultazione di Basi di dati con esempi di traduzioni
  - Consultazione di dizionari di citazioni



NLP group