

## Il Linguaggio Naturale: una Ricchezza da Sfruttare

*Il trattamento del “linguaggio naturale” continua a essere una tematica di ricerca e di sviluppo di grandissima importanza per la pervasività che caratterizza i suoi risultati nei confronti di numerosi settori di applicazione. Dopo una prima lunghissima fase di ricerca di modelli computazionali, si è aperta una seconda fase, di raccolta di supporti all'ingegnerizzazione dei prodotti di “linguaggio naturale” e alla loro integrazione in sistemi e servizi più ampi. Si tratta delle cosiddette “risorse linguistiche”, il cui panorama ha oggi acquisito una sua stabilità. Il mondo del trattamento del “linguaggio naturale” sta diventando sempre più complesso sia sul piano della ricerca sia del mercato. Questo perché, ormai, è presente l'intera filiera, che va dalla ricerca e progettazione di modelli più avanzati (multimodalità, “semantic web”, ecc.), al consolidamento delle tecniche di raccolta di risorse linguistiche, all'elaborazione di metodologie di apprendimento automatico, fino allo sviluppo di processi specifici mirati alla produzione di sistemi “intelligenti” robusti a supporto di applicazioni reali. Entriamo in questo mondo complesso con Remo Raffaelli, Direttore Commerciale di Synthema, società da sempre impegnata a scoprire il valore aggiunto del “linguaggio naturale”.*

*Cosa si intende, esattamente, per “linguaggio naturale”?*

È un tipo di comunicazione che non prevede alcun tipo di struttura o formalismo. Può essere, per esempio, una comunicazione telefonica o epistolare. È un formato libero ma con un altissimo contenuto informativo.

*In cosa consiste la sua elaborazione?*

L'elaborazione del “linguaggio naturale”, detta anche NLP (dall'inglese *Natural Language Processing*), è il processo che mira a rendere informaticamente utilizzabile l'informazione espressa in linguaggio umano, scritto o parlato, mediante l'impiego di tecnologie che si rifanno alla linguistica, alla statistica e all'intelligenza artificiale.

*È un processo difficile?*

È reso particolarmente difficile e complesso a causa delle caratteristiche intrinseche di ambiguità del linguaggio umano. Per questo motivo, il processo di elaborazione viene suddiviso in fasi diverse, tuttavia simili a quelle che si possono incontrare nel processo di elaborazione di un linguaggio di programmazione.

*Vale a dire?*

Una prima fase di *analisi lessicale*, in cui vi è la *scomposizione* di una espressione linguistica in *token* (in questo caso le parole), una seconda fase di *analisi sintattica*, in cui si procede all'arrangiamento dei *token* in una struttura sintattica (ad albero: *parse tree*), e una fase conclusiva di *analisi semantica* dove si arriva all'assegnazione di un significato (semantica) alla struttura sintattica e, di conseguenza, alla espressione linguistica.

*La ricerca ha contribuito a confermare l'importanza del “linguaggio naturale”?*

Sì. Il linguaggio è uno degli aspetti fondamentali del comportamento umano e costituisce una componente essenziale della nostra vita quotidiana. Sia nelle sue manifestazioni scritte sia in

quelle parlate, rappresenta il più comune mezzo di comunicazione e di interazione tra le persone. Sul suo conseguimento da parte di sistemi artificiali si sono concentrati sforzi notevoli non appena la tecnologia dei calcolatori ha fornito strumenti all'altezza del compito.

*Che importanza può avere il "linguaggio naturale" per il mondo delle aziende?*

Quando un'azienda utilizza nello scambio, nel rapporto con il mondo esterno, con il mercato, l'informazione in "linguaggio naturale", questa informazione ha sicuramente uno scopo primario, che è quello della comunicazione. Questa informazione, però, normalmente resta inutilizzata e non diventa un ricchezza informativa per l'azienda in quanto è molto difficile da utilizzare. Per l'essere umano è normale scambiarsi informazioni in "linguaggio naturale".

Queste informazioni possono diventare un asset importante anche per l'azienda, ma purtroppo non è facile gestirle.

*Come può essere sfruttata al meglio una simile ricchezza?*

Un aiuto importante viene fornito dal *text mining* e in particolare dal *knowledge mining*. Un potentissimo strumento che consente di estrarre conoscenza da materiale non strutturato. Siccome la *conoscenza* è uno degli asset più importanti che un'azienda ha o cerca di avere, disporre degli strumenti che permettono di ottenere questa conoscenza in modo automatico o semiautomatico è un passo molto importante in questo processo.

*Come si legano conoscenza informativa e sistemi di supporto decisionale?*

In generale, avere a disposizione strumenti che consentono di estrarre e classificare l'informazione, sicuramente facilita il compito degli operatori nei processi decisionali. Perché fornisce loro uno strumento di controllo assolutamente oggettivo, che mette in evidenza l'informazione nascosta nei vari documenti a disposizione dell'azienda. Informazione che non rappresenta solo l'insieme dei concetti racchiusi nei documenti, ma esprime anche come questi concetti sono relazionati l'uno all'altro. Questo può far scattare dei segnali importanti di attenzione per chi sta utilizzando tale informazione. E può essere di grande supporto nella formazione delle scelte decisionali.

*Intervista a cura di Titti Acone*