



**SERVIZIO DI EPIDEMIOLOGIA**  
Via Sabaudia, 164 – 10095 GRUGLIASCO (TO)  
☎ 011.40.188.221 fax 011.40.188.201  
E-mail: [sepi@epi.piemonte.it](mailto:sepi@epi.piemonte.it)



All'interno dei metodi di data mining, negli ultimi anni sono state sviluppate tecniche e strumenti per l'estrazione di informazioni e conoscenza da documenti di tipo testuale; la branca dell'informatica che si occupa di elaborare tali strumenti da un punto di vista teorico e utilizzarli nei vari contesti applicativi è generalmente indicata come text mining.

Obiettivo dei metodi di text mining è l'estrazione di informazioni da testi non strutturati, e può comportare varie forme di elaborazione, come l'analisi sintattica del linguaggio naturale, l'individuazione di pattern sulla base di un'analisi statistica dei termini presenti nei testi, l'individuazione dei concetti, l'individuazione di gerarchie di concetti come tassonomie e ontologie. Compiti tipici delle applicazioni di text mining sono la categorizzazione e il clustering di documenti testuali, il riassunto automatico, l'estrazione di named entities (generalmente persone, organizzazioni e luoghi), la sentiment analysis e l'opinion mining.

All'interno del Servizio da tempo è presente interesse rispetto a queste tecniche e, all'interno del piano di lavoro 2016, è prevista una piccola attività relativa a 'Valutazione e utilizzo di nuovi metodi per le attività di data mining (text mining, forecasting, analytics, machine learning, ...)'. Inoltre, all'interno di alcuni progetti sono state sviluppate prime esperienze di utilizzo di strumenti per il text mining.

Allo scopo di aggiornarci rispetto alla situazione e alle evoluzioni dei metodi di text mining e alle attività svolte in questo ambito è previsto un

### INCONTRO TECNICO

## L'analisi automatica dei testi: text mining, information extraction e natural language processing

**Martedì 13 dicembre 2016**

**Dalle ore 10.45 alle 13.00**

**Aula Formazione Servizio di Epidemiologia, ASL TO3  
Via Sabaudia 164 - Grugliasco**

- **Ore 10.45 - 11.00:** Evoluzione della quantità e qualità dell'informazione (*Marco Dalmasso*)
- **11.00 – 12.10:** Stato dell'arte del text mining per l'estrazione di informazioni (*Daniele Radicioni – Ricercatore presso il Dipartimento di Informatica – Docente di Sistemi cognitivi*)
- **Ore 12.10 - 12.25:** Il trattamento di dati testuali provenienti da Pronto soccorso (*Matteo Di Cunzolo - tesista del Corso di laurea in Informatica*)
- **12.25 – 12.40:** Esperienze di text mining all'interno del Servizio di Grugliasco (*Selene Bianco*)

Al termine delle presentazioni, i partecipanti potranno confrontarsi sull'interesse e la necessità di approfondire e sviluppare gli argomenti trattati. L'incontro è rivolto agli operatori del Servizio di Grugliasco e della rete epidemiologica regionale con interesse verso i temi del data mining e dell'uso di metodi non tradizionali di trattamento dati.