

Cognitive Computing



SAPIENZA
UNIVERSITÀ DI ROMA



Master in Data Intelligence e Strategie Decisionali
Dipartimento di Scienze Statistiche

Paolo Dell' Olmo

Outline

Cognitive Computing
Text Mining
Social Media Listening
Application

Cognitive Computing

Definition

Cognitive computing describes technology platforms that, broadly speaking, are based on the scientific disciplines of artificial intelligence and signal processing. These platforms encompass machine learning, reasoning, **natural language processing**, speech recognition and vision (object recognition), human–computer interaction, dialog and narrative generation, among other technologies.

Applications

Examples of such real-world use cases include the following:

Speech recognition

Sentiment analysis

Face detection

Risk Assessment

Fraud Detection

Behavioral Recommendations

Text Mining

Perché il social media listening?

Text Mining: analisi semiautomatica di dati destrutturati.



Perché il social media listening?

“CAPIRE LA CONVERSAZIONE ONLINE PER IDENTIFICARE INSIGHT STRATEGICI”

- Con l' aumento di consumatori che utilizzano blog, forum e social network per scambiarsi opinioni su prodotti e servizi, i social media stanno diventando una fonte molto rilevante di informazioni per le moderne ricerche di mercato.
- L' accesso istantaneo alla “mente del consumatore” attraverso i social media può aiutare il business a rispondere rapidamente a cambiamenti del mercato; per il business farmaceutico sarà possibile non solo migliorare la qualità e l' impatto della comunicazione, ma anche capire necessità e preoccupazioni dei pazienti.

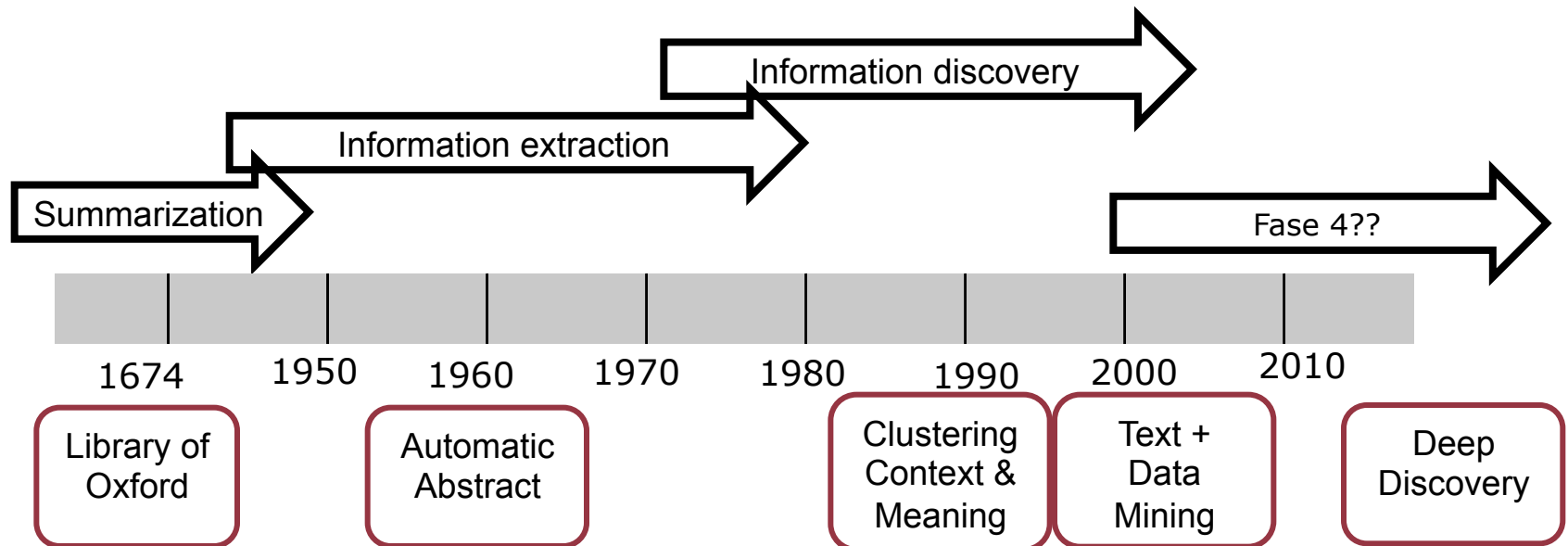


Text mining – Definizione e storia

Il text mining è:

“una forma particolare di data mining dove i dati consistono in testi in lingua naturale, in altre parole, documenti "destrutturati". Il text mining unisce la tecnologia della lingua con gli algoritmi del data mining.”

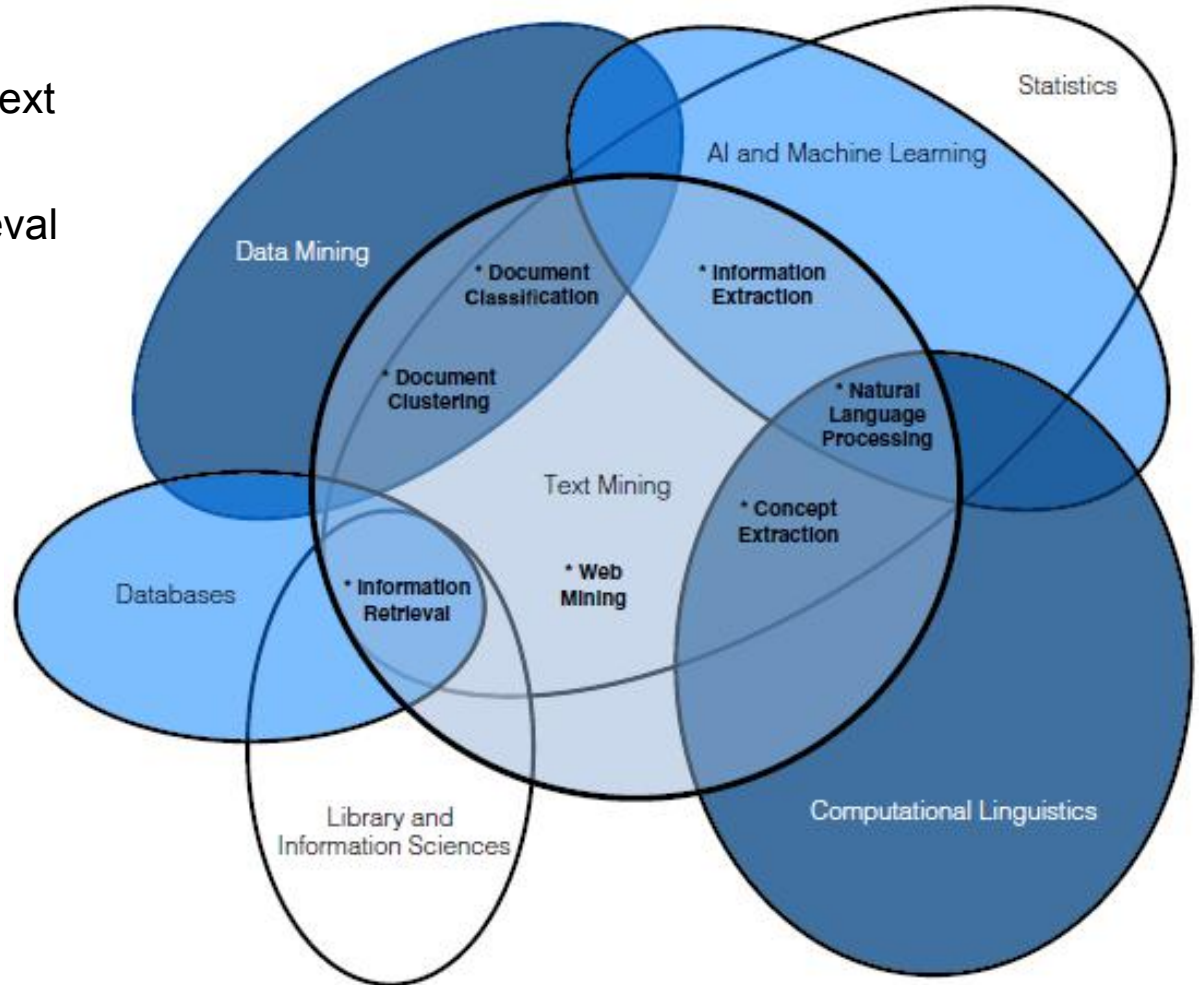
“l'analisi di dati testuali destrutturati tramite le tecniche del data mining per risolvere i problemi del business.”



Aree di applicazione e attività

Gli ambiti di applicazione del text mining ad oggi sono:

- Search and information retrieval
- Document clustering
- Document classification
- Web mining
- Information extraction
- Natural language processing
- Concept extraction



Aree di applicazione e attività

Attività:

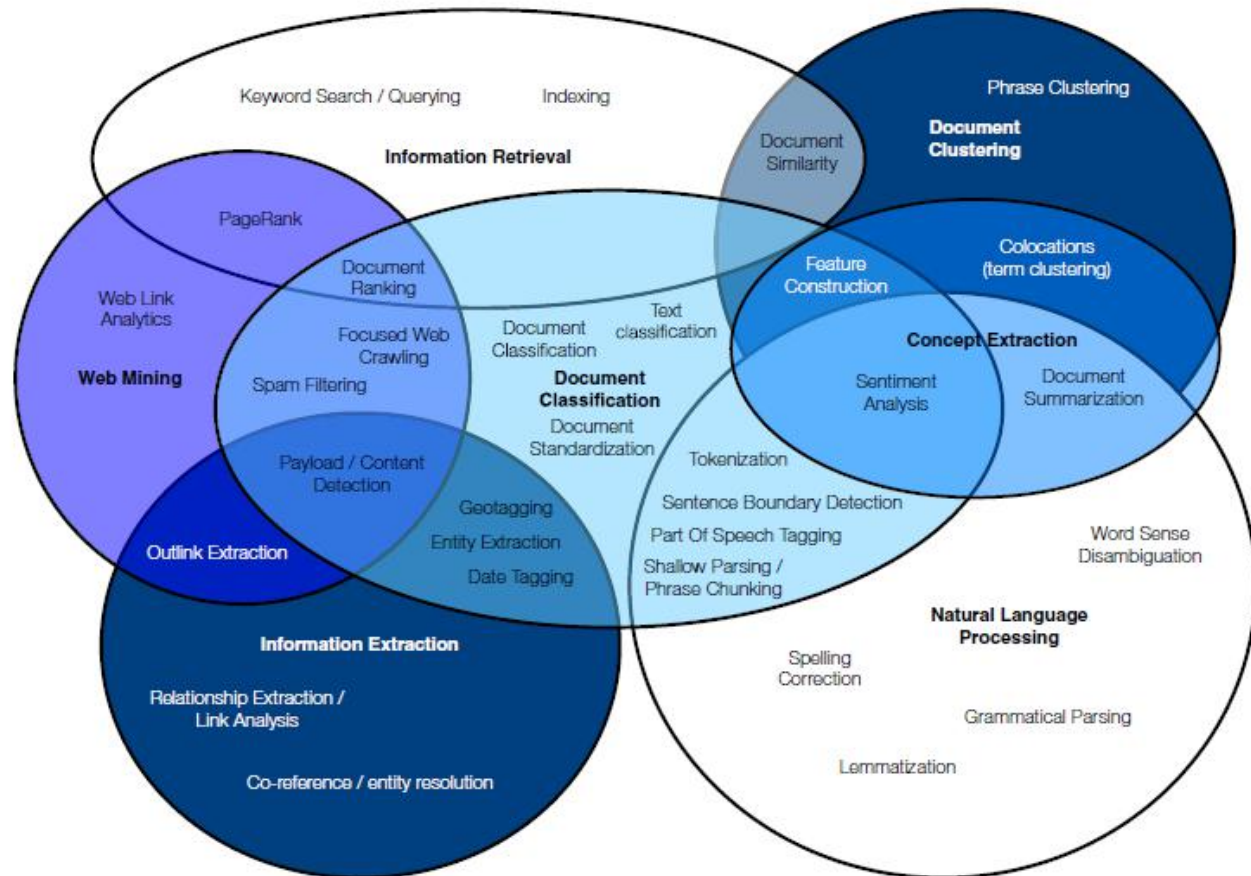
National Security/
Intelligence

Sentiment Analysis

Social Media
Monitoring

Automated Ads
Placement

Enterprise
Intelligence

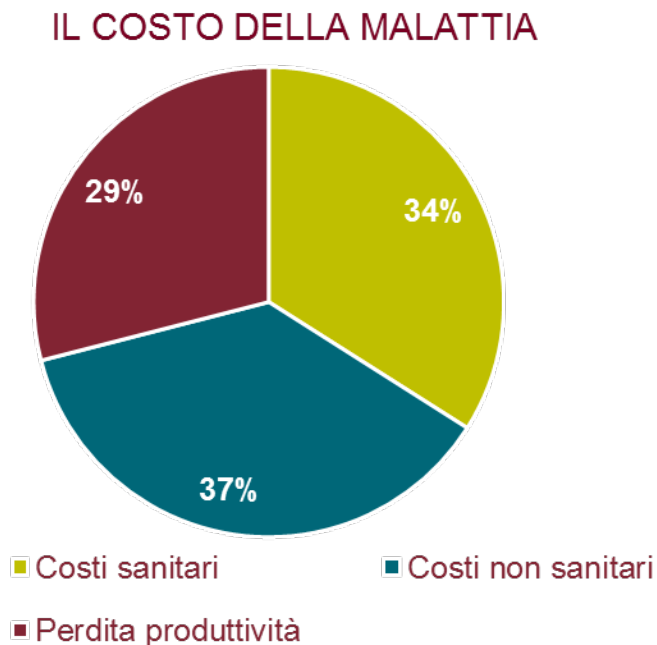
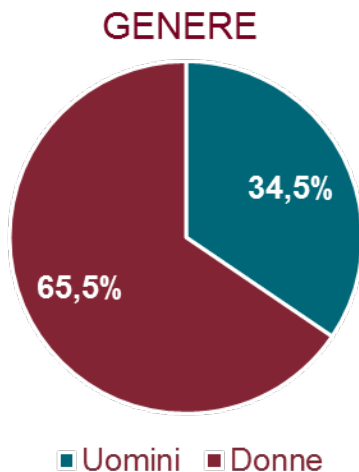


....

La sclerosi multipla

La sclerosi multipla è una malattia cronica progressiva che colpisce il sistema nervoso centrale (cervello, nervi ottici e midollo spinale), caratterizzata dalla distruzione della guaina mielinica che isola le fibre nervose all'interno del sistema nervoso centrale.

Stima totale pazienti 2015	Nuovi pazienti 2015	Età media pazienti
110.000	3.400	38 anni



Lo scenario di indagine

Il social media listening si pone quindi l'obiettivo di:

- identificare rapidamente cambiamenti di trend e scenario nella comunicazione tra pazienti nel tentativo di fornire sempre servizi migliori.
- far fronte a nuove richieste specifiche di cui si abbia evidenza sui canali social.
- sviluppare analisi quantitative e qualitative su tematiche chiave per l'azienda.

Le 3 fasi dell' indagine

Raccolta



Analisi



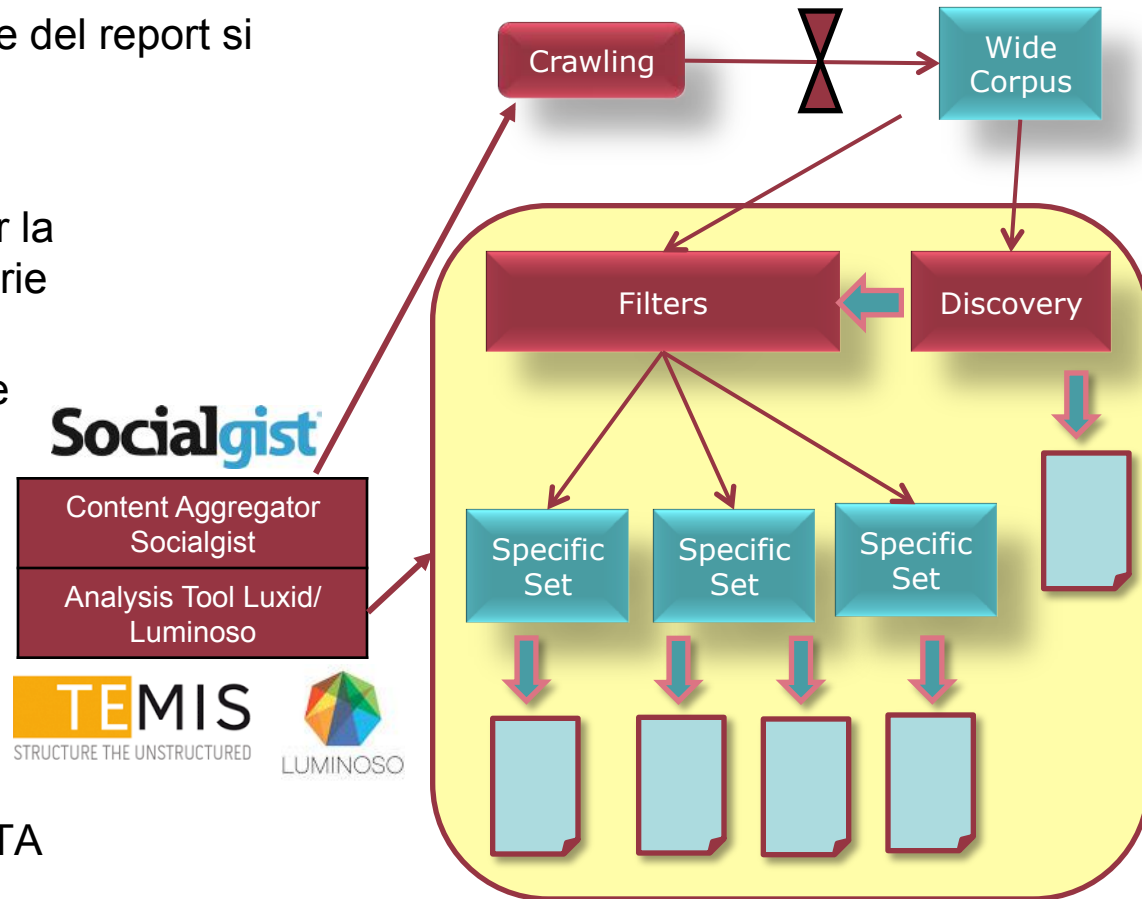
Presentazione



Il processo di indagine

Il processo per arrivare alla produzione del report si compone di varie fasi:

1. Interazione con il marketing delle singole Aree terapeutiche (TA) per la definizione dei temi e delle categorie (macrotrend) di indagine
2. Definizione della query di ricerca e implementazione delle fonti nel Content Aggregator (Socialgist)
3. Raccolta dei documenti e upload del corpus nei tool di text mining
4. Categorizzazione e analisi dei documenti
5. Produzione del report
6. Presentazione al marketing della TA



Le 4 fasi dell' analisi

1) Preprocessing

Insieme di tecniche per ricondurre il testo destrutturato ad una forma strutturata.

- Tokenization
- Stopping
- Lemmatization and PoS Tagging
- Spelling normalization
- Cases normalization

3) Estrazione di informazioni

L' obiettivo è ricavare informazioni strutturate da un testo semi o destrutturato.

- Named entity recognition
- Relation extraction

2) Vettorizzazione dei documenti

Trasformazione del testo in un vettore elaborabile dagli algoritmi di text mining.

- Binaria
- Integer
- TF-IDF

4) Categorizzazione dei documenti

L' obiettivo è dividere i documenti in due o più categorie.

- Decision Trees
- Rule-based Classifiers
- Support Vector Machine Classifiers
- Neural Network Classifiers
- Naive Bayes Classifiers

L'analisi tramite Luxid

L'analisi tramite Luxid si compone di diversi step:

1. Creazione delle SCA composte dalle entità riconoscibili dal sistema
2. Categorizzazione manuale di un training set di documenti sui quali creare un modello di classificazione
3. Applicazione del modello ai nuovi dati di test e assegnazione delle categorie
4. Analisi approfondita del dataset di documenti categorizzati
5. Creazione di grafici ad hoc per la rilevazione di picchi di conversazione
6. Analisi approfondita delle categorie in termini di entità più frequenti
7. Fase di discovery per rispondere alle domande del marketing

Label	ID	attribute
Copaxone	Copaxone	Teva
Daclizumab	Daclizumab	Biogen
Extavia	Extavia	Novartis
Gilenya	Gilenya	Novartis
Lemtrada	Lemtrada	Genzyme
Ocrelizumab	Ocrelizumab	Roche
Plegridy	Plegridy	Biogen
Rebif	Rebif	Merck
Tecfidera	Tecfidera	Biogen
Tysabri	Tysabri	Biogen

New Descriptor:

Edit

Label: Rebif

ID: Rebif

attribute: Merck

language: italian

variant:
febif
Rebiff22
Rebif
Rebif22
Rebiff
Rebif-44
febiff
Rebif-22
Rebif44

COPAXONE® - Foglietto Illustrativo - My Personal Trainer

www.my-personaltrainer.it › Foglietti illustrativi › Farmaci › Sclerosi multipla ▼

Copaxone: Per quali malattie si usa? A cosa Serve? Come si utilizza e quando non dev'essere usato. Avvertenze ed Effetti Collaterali.

Copaxone - Associazione Italiana Sclerosi Multipla

<https://www.aism.it/index.aspx?codpage=copaxone> ▼

Copaxone non è indicato in pazienti affetti da SM progressiva primaria o secondaria. Effetti collaterali comuni: reazioni sito iniezione, cefalea, malessere ...

Sclerosi multipla: cambio del trattamento da interferone a copaxone ...

https://www.aism.it/index.aspx?codpage=2015_02_ricerca_msif_sm...copaxone ▼

17 feb 2015 - Il passaggio del trattamento da interferone a copaxone nelle persone con ... trarre beneficio al passaggio con glatiramer acetato (Copaxone).

Copaxone | Scheda Tecnica del Farmaco - Torrinomedica

www.torinomedica.it › Indice dei Farmaci > > Lettera C > ▼

Copaxone 20 mg/ml soluzione iniettabile, siringa preriempita - Glatiramer Acetato - Consulta la Monografia del medicinale ed il foglietto illustrativo: ...

Effetti indesiderati - COPAXONE SC 28SIR 20MG/ML | Starbene.it

www.starbene.it/farmaci/copaxone-sc-28sir-20mg_ml.../effetti-indesiderati ▼

Copaxone è indicato per il trattamento di forme recidivanti di sclerosi multipla (SM) (vedere paragrafo 5.1 per informazioni importanti sulla popolazione in cui è ...

Effetti indesiderati - COPAXONE SC 12SIR 40MG/ML | Starbene.it

www.starbene.it/farmaci/copaxone-sc-12sir-40mg-ml.../effetti-indesiderati ▼

Copaxone è indicato per il trattamento di forme recidivanti di sclerosi multipla (SM) (vedere paragrafo 5.1 per informazioni importanti sulla popolazione in cui è ...

Il report

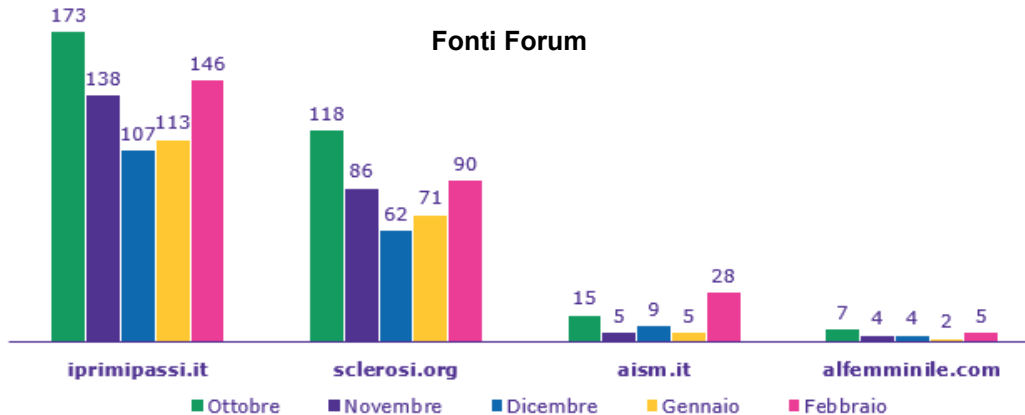
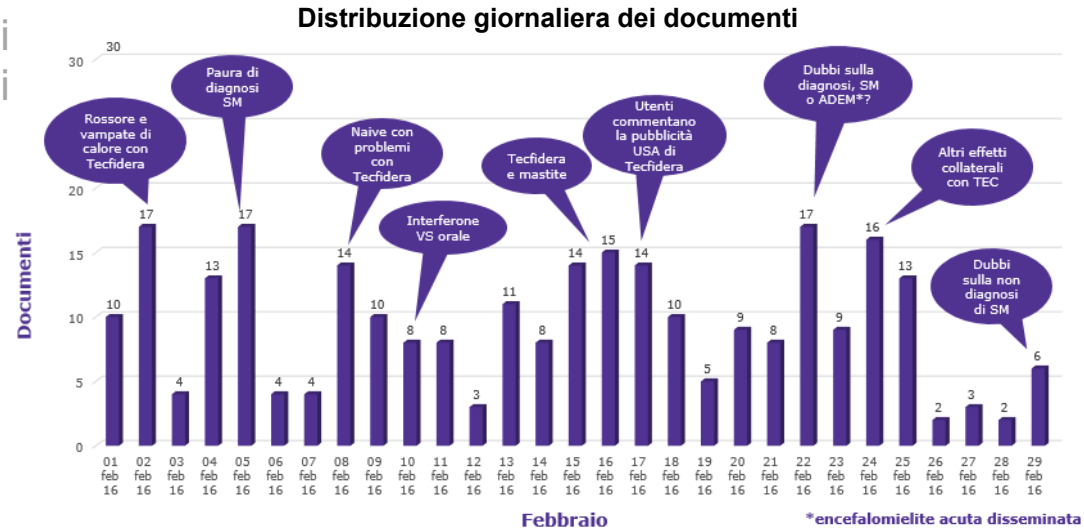
Il report discusso con il marketing di area terapeutica è diviso in tre sezioni chiave:

- Analisi quantitativa
- Focus macrotrend
- Discovery

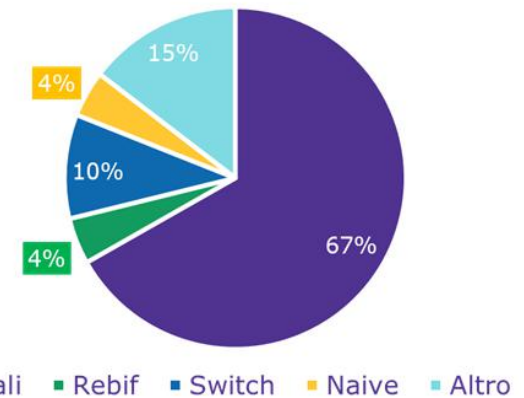
Il report – Analisi quantitativa

Il report discusso con il marketing di area terapeutica è diviso in tre sezioni chiave:

- Analisi quantitativa
- Focus macrotrend
- Discovery



Distribuzione conversazioni per MacroTrend



Il report – Focus macrotrend

Il report discusso con il marketing di area terapeutica è diviso in tre sezioni chiave:

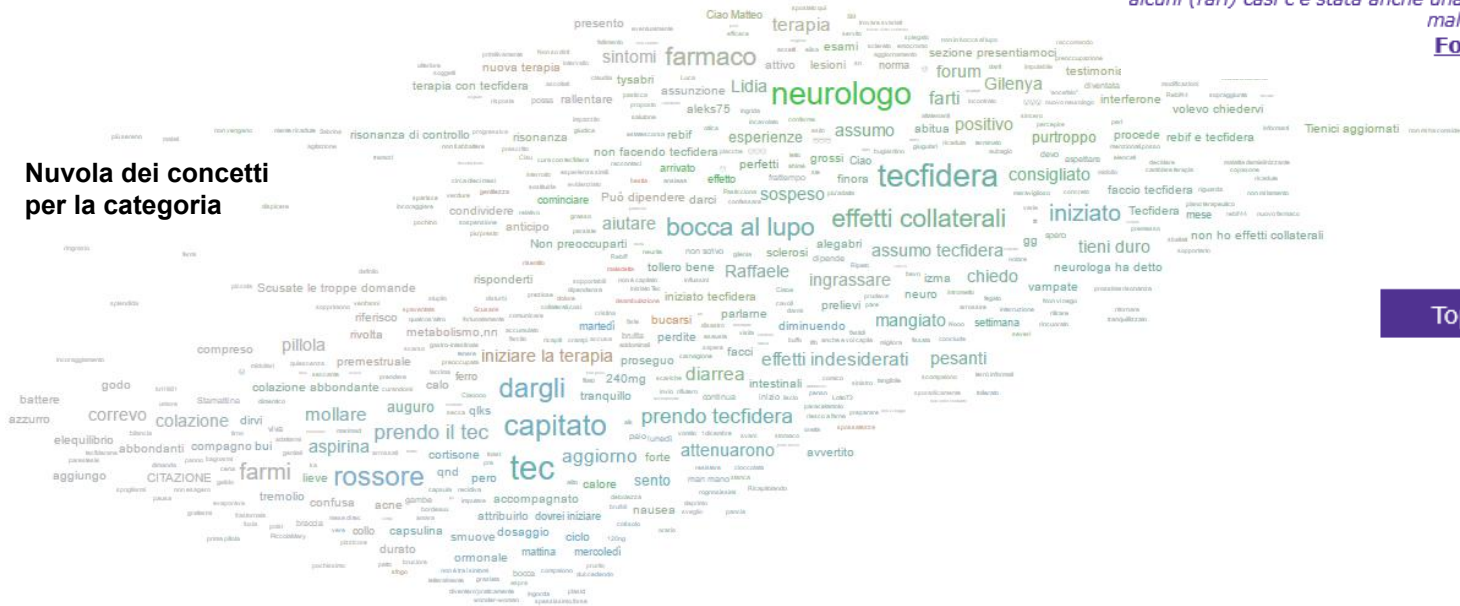
- Analisi quantitativa
 - Focus macrotrend
- Discovery

Verbatim

“Salve a tutti...faccio un veloce riassunto della mia situazione; primi sintomi della malattia 25 anni fa con varie neuriti ottiche, cortisone, alti e bassi e nel corso degli anni sensazione di "fasciatura" alla gamba destra e progressivo indebolimento della stessa. Tre anni, fa a seguito nuove lesioni, si decide di iniziare copaxone che, (...) nonostante le "bellissime" risonanze in realtà le cose non vanno poi così bene, la neurologa mi ha fatto l'esempio delle braci sotto la cenere dove se anche non si vede la fiamma in realtà sotto continua a bruciare, quindi ha stabilito che forse il copaxone non era sufficiente. Ora sospensione della terapia per una settimana poi **inizio con Tecfidera**, tutto sommato **non ne potevo più delle iniezioni giornaliere e gli effetti collaterali del nuovo farmaco non mi preoccupano più di tanto anzi...la colazione abbondante non mi disturba per niente e per quanto riguarda l'intestino sono stitica da sempre quindi...** Avendo letto in questa sezione che alcuni prendono Tecfidera già da tempo **mi piacerebbe sapere se avete avuto benefici a livello fisico** in quanto la mia neurologa mi ha detto che in America lo usano già da diversi anni e in alcuni (rari) casi c'è stata anche una leggera regressione dei sintomi della malattia.”

Fonte

Nuvola dei concetti per la categoria



Keywords

Top 5 parole	Occorrenze
Iniziare	91
Terapia	85
Anni	72
Sapere	62
Dovere	52

Il report - Discovery

Il report discusso con il marketing di area terapeutica è diviso in tre sezioni chiave:

- Analisi quantitativa
- Focus macrotrend
- Discovery

Sample di slide della sezione Discovery

Discovery Fonti dell'informazione – nei Forum

Da che fonti
traggono
informazioni
i pazienti?

VERBATIM

“ Con riferimento agli articoli usciti qualche tempo fa e alle interviste a Comi che sembrava entusiasta, **qualcuno sa qualcosa in più su questi farmaci "rivoluzionari"** che dovrebbero uscire nel 2016? Avete chiesto qualcosa ai vostri neurologi? Io no perchè essendo in cura al San Raffaele sono certa che nessuno direbbe qualcosa in più o di diverso dal Prof.

Il fatto che possano potenzialmente sanare le placche continua a sembrarmi un sogno...

<http://www.osservatoriomalattierare.it/sci...lla-storia>
Baci a tutti! ” **Fonte**

“ La fonte è Fondazione Serono.

<http://www.fondazione-serono.org/sclerosi-multipla/ultime-notizie-sclerosi-multipla/fingolimod-dimetilfumarato-laquinimod-aggiornamento-sicurezza>

Come si fa a rimanere tranquilli quando si leggono notizie del genere? la cosa che mi sciocca è che l'età di chi assume tali farmaci è sempre più bassa. ” **Fonte**

“ Ciao, su Internet si trovano notizie, io solitamente però **leggo solo i siti che ritengo affidabili!!!** in uno di questi c'è un articolo interessante:

<http://www.fondazione-serono.org/sclerosi-m...erosi-multipla/>

Per il resto ne so quanto te! Facci sapere però se ti va. Intanto in bocca al lupo ” **Fonte**

Dall'analisi
delle
conversazioni
su fonte Forum
solo l'1% dei
documenti (#12)
contiene un link
URL.

La principale fonte di informazione dei pazienti resta il neurologo, in 121 documenti da fonti Forum (26% del totale).

Dei 12 documenti con URL, 2 riportano un articolo di AISM.IT e 1, relativo ai nuovi trattamenti, fa riferimento a [osservatoriomalattierare.it](http://www.osservatoriomalattierare.it). Si evidenzia anche la presenza (in 2 casi) di [fondazione-serono.org](http://www.fondazione-serono.org) ritenuta fonte affidabile.

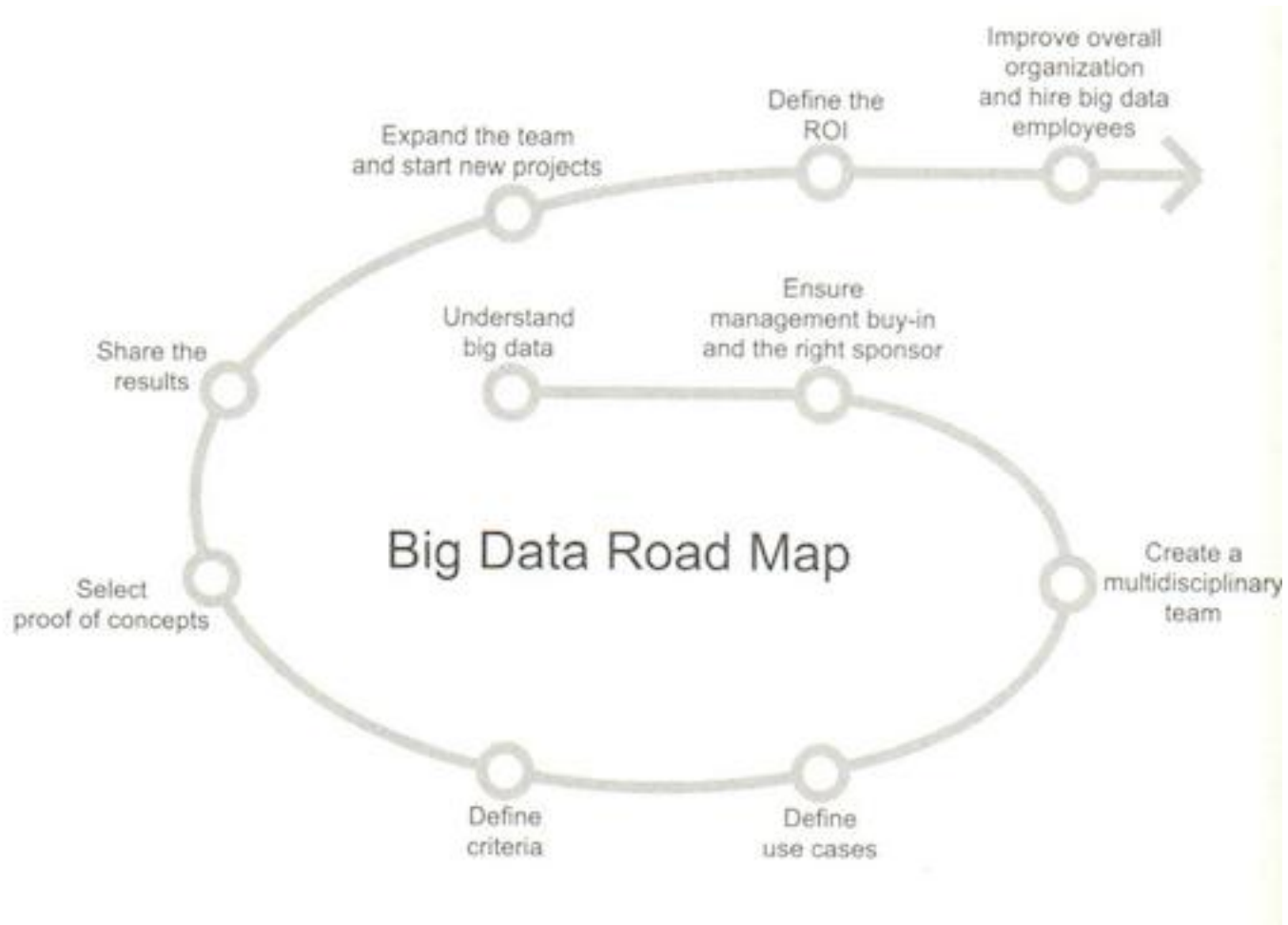
Nota

L'analisi **semiautomatica** di dati destrutturati permette di fornire un **quadro rappresentativo** della conversazione online in modo rapido e continuativo, utilizzando un **report sviluppato ad hoc**.

L'automatizzazione di tecniche come la categorizzazione dei documenti o l'estrazione di entità o di concetti permette all'analista di **ridurre i tempi di elaborazione**, sicuro di non perdere informazioni rilevanti e di approfondire a dovere le tematiche di maggiore interesse.

Ad oggi sono in valutazione step evolutivi del progetto di Social Media Listening; ad esempio il monitoraggio esteso alle news di settore o l'implementazione di reportistica ad hoc all'interno del CRM sono solo due dei possibili scenari futuri.

Cosa Fare: Road Map



Grazie per l'attenzione !

paolo.delloolmo@uniroma1.it