



Università di Pisa

Corso di Laurea in Informatica Umanistica  
**Specialistica in Editoria Elettronica**

Relazione

**LO SCENARIO COGNITIVO E  
CULTURALE DEI TESTI LETTERARI:  
DUE CASI DI STUDIO**

Candidata: *Silvia Greco*

---

Relatore: *Prof. Paolo Rossi*

---

**Anno Accademico 2010-2011**

*Ai sogni*

Ringrazio ... mia sorella Sabrina, il mio paziente compagno Guido, la mia amica di *sempre* Silvia, Nera, Gianni, Roberta, Tiziana, Iryna ... che, più di quanto non abbia mai fatto io stessa, hanno creduto che questo giorno sarebbe arrivato incoraggiandomi durante i miei innumerevoli momenti di scoraggiamento e di stanchezza;

... i miei “angeli” che mi hanno lasciata durante questo cammino, perché so che, da Lassù, hanno continuato a darmi la forza necessaria per non fermarmi;

... il corso di Informatica Umanistica, i compagni, i sapienti professori, “immensi orizzonti” dove far spaziare la mia mente curiosa, la loro professionalità e disponibilità continua.

Ringrazio colei ... che non ha mai creduto in me e nelle mie capacità, perché, senza volerlo, mi ha donato la caparbieta e la grinta necessaria.

Ringrazio anche ... me stessa, perché la volontà di farcela, malgrado tutto, è stata la mia più fedele alleata facendo prevalere, sempre e su tutto, la voglia di emozionarmi;

Ringrazio ... infine, il mio *c a s s e t t o* perché aprendosi mi ha mostrato un sogno bellissimo: *QUESTO!*

# Indice generale

<b>1</b>	<b>Introduzione.....</b>	<b>6</b>
<b>2</b>	<b>Definizione e concetti generali.....</b>	<b>8</b>
2.1	Il modello matematico	8
2.1.1	Interpolazione ed estrapolazione.....	9
2.2	La scienza statistica	10
2.2.1	Caratteri delle distribuzioni statistiche e teoria dell'errore.....	11
<b>3</b>	<b>L'analisi qualitativa dei testi.....</b>	<b>18</b>
3.1	Cenni	18
<b>4</b>	<b>Lo scenario cognitivo e culturale dei testi letterari.....</b>	<b>21</b>
4.1	L'analisi <i>quantitativa</i> del testo	21
4.2	Le metodologie fisiche per le scienze umane	22
4.3	Gli obiettivi prefissati	22
4.4	Procedure operative e concetti applicati	24
4.5	Il problema degli <i>oggetti estesi</i>	37
4.6	L'ellisse della percezione	38
<b>5</b>	<b>Il primo caso di studio: Il Decameron.....</b>	<b>44</b>
5.1	Giovanni Boccaccio	44
5.2	Decameron	47
5.2.1	Lo scenario cognitivo.....	52
5.2.2	Risultanze numeriche.....	57
5.2.3	Lo scenario culturale.....	60
5.2.4	Problematiche operative.....	63
5.3	Conclusioni	65
<b>6</b>	<b>Il caso di studio <i>originale</i> : Il Milione.....</b>	<b>66</b>
6.1	Marco Polo	66
6.2	Il Milione	69
6.2.1	L'ellisse della cognizione ne Il Milione.....	70
6.2.2	Gli orizzonti cognitivi di altri testi medievali.....	76
6.2.3	Risultanze numeriche.....	78
6.2.4	Lo scenario culturale.....	84
6.2.5	Problematiche operative.....	87
6.3	Conclusioni	90

<b>7 Il sito web: “Lo scenario cognitivo in letteratura” .....</b>	<b>91</b>
7.1 Struttura e funzionalità .....	91
<b>8 Le visualizzazioni .....</b>	<b>93</b>
8.1 Visualizzazioni per i testi .....	93
8.2 Le visualizzazioni della geografia cognitiva .....	94
8.2.1 Le visualizzazioni relative alle occorrenze .....	94
8.2.2 Le visualizzazioni relative agli oggetti geografici .....	97
<b>9 Riepilogo e sviluppi futuri .....</b>	<b>100</b>
9.1 Le tematiche trattate .....	100
9.2 Gli obiettivi raggiunti .....	101
9.3 Gli sviluppi futuri .....	102

# 1 Introduzione

**L**a tesi di laurea specialistica, dal titolo «Gli Scenari Cognitivi e Culturali dei testi letterari: due casi di studio» costituisce un'estensione dell'analisi illustrata nell'articolo [1] *Measuring Large Scale Space Perception in Literary Texts*, pubblicato sulla rivista scientifica «Physica A: Statistical Mechanics and its Applications» e curato dal professor Paolo Rossi, relatore di questa tesi e preside della facoltà di Scienze Matematiche, Fisiche e Naturali dell'Università di Pisa. Nel saggio del prof. Rossi si definivano, partendo dall'analisi di un testo letterario o storico, alcuni importanti concetti, che saranno fondamentali anche in questo studio. Citiamo, per esempio, il *centro* e il *raggio della percezione*, significati che saranno approfonditi anche focalizzando l'attenzione sull'applicazione di algoritmi formali per la loro determinazione e sull'associazione di tali parametri nell'analisi *quantitativa* dei testi letterari.

L'obiettivo di questo studio è costituito dalla rappresentazione delle percezioni umane insite negli artefatti culturali, tramite l'utilizzo di *parametri quantitativi* che, di norma, sono impiegati nell'analisi dei fenomeni legati alle *scienze esatte*.

L'intento è, in altre parole, quello di cogliere, studiare ed esporre il legame esistente tra discipline filosoficamente relegate ad ambiti separati e antistanti.

Per fare questo e per una migliore comprensione della materia trattata, saranno introdotti i *principi fisici di base* sui quali si fondano gli algoritmi formali adoperati nell'analisi *quantitativa* dei due casi di studio.

Sarà analizzato, nello specifico, il *Decameron* di Giovanni Boccaccio, all'interno del quale, si definiranno, utilizzando i principi della fisica, lo *spazio* e l'*orizzonte* della *cognizione*, costituito dal *centro* e dal *raggio della percezione*. L'opera di Boccaccio, per questo studio, costituisce il *caso base*, in quanto, per le sue caratteristiche la ricerca dei parametri necessari per l'analisi quantitativa del testo prevederà l'applicazione di algoritmi formali relativamente lineari. L'elemento *originale* sarà rappresentato dall'analisi quantitativa effettuata su *Il Milione* di Marco Polo e dall'ellisse della percezione che, dallo studio del testo, sarà tracciata.

Il Milione, per le risultanze cui dà origine, comporterà un'interessante evoluzione dei concetti *base* di centro e di raggio della percezione insiti nel Decameron. Infatti, per il resoconto del viaggio di Marco Polo sarà necessario l'impiego di calcoli trigonometrici più raffinati e complessi, al fine di ottenere quei parametri definiti *assi principali* della *percezione* (o direzioni privilegiate della cognizione) e, dai quali mostrare, numericamente e graficamente, l'ampiezza dell'orizzonte cognitivo definito in una forma ellittica.

I risultati, ottenuti da entrambi le analisi quantitative, costituiranno la base *decisionale* per una valutazione approfondita dello scenario storico-culturale originario dei testi; elemento, crediamo, indispensabile anche per confermare l'efficacia delle teorie fisiche e la fondatezza dei parametri applicati nell'indagine.

Inoltre, per dare completezza alla tematica trattata e sulla base delle risultanze ottenute, si presenteranno differenti tipologie di visualizzazione dei dati, di natura sia statica che dinamica, relative sia alle occorrenze rintracciate, che alla distribuzione degli oggetti geografici rinvenuti nei testi.

Infine, i punti cardine degli argomenti di studio, e le relative visualizzazioni, saranno raggiungibili accedendo alle pagine del sito web: *Scenari cognitivi dei testi letterari* creato allo scopo informativo e illustrativo. L'utente, accedendo alle pagine del sito web, sarà accolto da un'interfaccia semplice e gradevole e potrà conoscere i concetti trattati e le differenti tipologie di visualizzazione statiche e dinamiche, nella forma di galleria di immagini realizzate a corredo dell'esposizione dei risultati ottenuti. Il sito web è stato progettato e realizzato privilegiando il livello di accessibilità e utilizzando un codice (X)HTML e CSS valido, cioè, una codifica intesa nel rispetto degli standard definiti dal W3C.

Nella parte finale di questa relazione, si riepilogheranno i *punti principali* sui quali si basa lo studio, al fine di rendere fluida e ben articolata l'esposizione delle tematiche trattate. Si ipotizzeranno, infine, gli sviluppi che l'analisi *quantitativa* basata sulle metodologie fisiche applicate ai testi letterari, ci auguriamo possa avere in futuro.

## 2 Definizione e concetti generali

### 2.1 Il modello matematico

Si introducono i concetti matematici e statistici di base, da quello di *modello matematico*, per chiarire la materia che andremo a trattare, e per dimostrare la relazione esistente tra le scienze *fisiche* e quelle *umane*, nell'ambito dell'analisi dei testi scritti. Procediamo, quindi, con la definizione di *modello matematico*<sup>1</sup> intendendo una *versione semplificata* e immaginaria della *porzione di mondo* da studiare in cui è possibile e auspicabile eseguire calcoli esatti. Lo scopo della costruzione di un modello matematico è quello di rappresentare, nel modo più fedele possibile, un determinato fenomeno reale, al fine di poter fare, ad esempio, *previsioni* sullo stato in cui esso potrà trovarsi nel futuro. Inoltre, trattandosi di una versione semplificata del mondo reale, nel modello matematico è possibile considerare solamente quel particolare sottoinsieme di aspetti del fenomeno che siano d'interesse in una specifica situazione. In altre parole, un modello matematico descrive l'evoluzione di un fatto o di un sistema, fornendo dei valori in ingresso chiamati *dati di input* e restituendo dati in uscita denominati *dati di output*. Esso potrà dirsi *efficace*, se l'output è *prossimo* alle misurazioni effettuate nell'osservazione dell'evento reale. Questi strumenti sono spesso rappresentati da equazioni di varia natura risolte con metodi matematici conosciuti, che rappresentano le variabili, i parametri e le grandezze che non possono essere manipolate. Per quanto riguarda le tipologie di modelli matematici esiste una distinzione tra quelli *statici* e quelli *dinamici*. Nei modelli statici il fenomeno o il sistema è rappresentato in un determinato istante; in quelli dinamici, invece, si «(...) tiene in considerazione la variabilità di alcuni parametri del sistema (dette *variabili dipendenti*) quando sia nota o prevedibile la variazione di altri parametri (*variabili indipendenti*) (...)» (Rossi 2009).

---

<sup>1</sup> [http://it.wikipedia.org/wiki/Modello\\_matematico](http://it.wikipedia.org/wiki/Modello_matematico)



Riassumendo, con l'utilizzo dei modelli dinamici, è descritta la variabilità di un fenomeno o di un sistema fisico nel tempo. La caratteristica principale di un modello matematico, che si possa definire *efficace*, consiste nella sua capacità di incorporare e descrivere con ragionevole accuratezza, l'insieme di tutti i dati già conosciuti e che riguardano il fenomeno oggetto di studio.

### 2.1.1 Interpolazione ed estrapolazione

Dopo aver tracciato il concetto di modello matematico di un fenomeno, ci soffermiamo a descrivere le finalità descrittive e/o predittive della sua costruzione, illustrando le possibili limitazioni, materiali e concettuali, insite nella modellazione matematica. Nell'analisi quantitativa di un fenomeno, esistono precisi criteri, anch'essi di natura quantitativa, concepiti per misurare l'aderenza di un modello matematico ai dati disponibili. In matematica e, in particolare, in analisi numerica, per *interpolazione* si intende quel metodo che permette di individuare *nuovi punti* del piano cartesiano, da un insieme finito di punti dati. Questa definizione è valida nell'ipotesi in cui tutti i punti si possano riferire «a una funzione  $f(x)$  di una data famiglia di funzioni di una variabile reale»<sup>2</sup>. In altre parole, interpolazione è quando nelle attività scientifiche e tecnologiche, e in genere negli studi quantitativi di qualsiasi fenomeno, accade di avere un certo numero di punti del piano ottenuti tramite operazioni di campionamento o con apparecchiature di misura, e di voler individuare una funzione che passi per tutti i punti dati, o almeno nelle loro vicinanze. L'*estrapolazione*<sup>3</sup>, invece, è quel processo che permette di calcolare il valore di informazioni che sono *esterne* a un insieme discreto di dati noti. Nella pratica, dato un piano cartesiano sul quale siano stati tracciati i punti  $(x_i, y_i)$  corrispondenti all'insieme di valori noti, si vuole trovare il valore  $y_w$  corrispondente ad un valore  $x_w$  maggiore (o minore) di ciascun  $x_i$ .

---

<sup>2</sup> <http://it.wikipedia.org/wiki/Interpolazione>

<sup>3</sup> <http://it.wikipedia.org/wiki/Estrapolazione>

In definitiva, l'estrapolazione è un processo simile, per alcuni aspetti, all'interpolazione, infatti, quest'ultima costruisce nuovi punti all'interno di un insieme di punti noti; ma i risultati che si ottengono, nel caso della estrapolazione, hanno spesso minor significato poiché soggetti a un più elevato grado di incertezza<sup>4</sup>.

## 2.2 La scienza statistica

La statistica è una disciplina che ha come scopo lo studio quantitativo e qualitativo di un particolare fenomeno in condizioni di non determinismo o di incertezza, ovvero di non completa conoscenza di esso o di parte di esso. Essa studia «i modi (descritti attraverso formule matematiche) in cui una realtà fenomenica limitatamente ai fenomeni collettivi - può essere sintetizzata e quindi compresa»<sup>5</sup>. La scienza statistica studia, in particolare, come raccogliere i dati e analizzarli al fine di ottenere informazioni che consentano di avanzare nella conoscenza partendo dall'osservazione e dall'analisi della realtà<sup>6</sup>. In ambito statistico, è fondamentale il concetto di *distribuzione*<sup>7</sup> che si definisce come la rappresentazione del modo in cui le diverse modalità di un carattere si ripartiscono nelle unità statistiche che compongono il collettivo oggetto di studio. In altre parole, si descrive come distribuzioni di una certa variabile, all'interno di una *popolazione*, il numero delle ricorrenze all'interno della stessa, di ciascuno dei valori ammissibili e relativi alla variazione in esame. In statistica, con il termine *popolazione*, s'intende, qualsiasi insieme di elementi che siano oggetto di studio, ossia l'insieme delle unità (dette *unità statistiche*) sulle quali è effettuata la rilevazione di tutte le modalità di presentazione di un fenomeno studiato si presenta. Le unità statistiche hanno in comune almeno una caratteristica che è accuratamente definita al fine di delimitare il loro insieme.

---

<sup>4</sup> <http://it.wikipedia.org/wiki/Estrapolazione>

<sup>5</sup> [http://www.unipa.it/sorce/Materiali/diet0910/diet0910\\_04\\_Statistica.pdf](http://www.unipa.it/sorce/Materiali/diet0910/diet0910_04_Statistica.pdf)

<sup>6</sup> <http://it.wikipedia.org/wiki/Statistica>

<sup>7</sup> [http://it.wikipedia.org/wiki/Distribuzione\\_%28statistica%29](http://it.wikipedia.org/wiki/Distribuzione_%28statistica%29)

Le distribuzioni, in ambito statistico, possono essere *semplici* se è rilevato un solo carattere, oppure *multiple*, qualora si focalizzino più caratteri sullo stesso collettivo. Le due branche principali<sup>8</sup> in cui, comunemente, è suddivisa la scienza statistica sono la statistica *descrittiva* e quella *inferenziale*. Intendendo per *descrittiva* quella disciplina che studia i criteri di rilevazione, di classificazione e di sintesi, delle informazioni che si riferiscono a una «popolazione» oggetto di studio. In altre parole, essa ha come obiettivo principale, quello di organizzare, riassumere e presentare i dati in modo ordinato, permettendo la loro sintesi<sup>9</sup> al fine di ottenere informazioni efficaci per il problema studiato. Nella statistica *inferenziale*, l'inferenza statistica è, invece, il procedimento per cui si inducono le caratteristiche di una popolazione dall'osservazione di una parte di essa, detta campione. Esso è selezionato, solitamente, mediante un esperimento casuale, detto anche aleatorio<sup>10</sup> o stocastico e quindi, imprevedibile e incerto<sup>11</sup>.

### 2.2.1 Caratteri delle distribuzioni statistiche e teoria dell'errore

La statistica descrittiva ha come scopo la raccolta, la presentazione e l'analisi dei dati riguardanti l'oggetto di studio, e consiste nella costruzione delle relative distribuzioni e nella valutazione della *media*, della *varianza* e delle *correlazioni*, intese come le caratteristiche insite nelle distribuzioni stesse.

La *media*<sup>12</sup> (v.fig.1) è un valore numerico che descrive *sinteticamente* un insieme di dati.

Esistono varie tipologie di media, che possono essere utilizzate per rappresentare e studiare un determinato fenomeno, ma quelle più diffusamente impiegate sono: la media *aritmetica*, quella *geometrica*, e la media *armonica*.

---

<sup>8</sup> <http://it.wikipedia.org/wiki/Statistica>

<sup>9</sup> [http://it.wikipedia.org/wiki/Statistica\\_descrittiva](http://it.wikipedia.org/wiki/Statistica_descrittiva)

<sup>10</sup> [http://it.wikipedia.org/wiki/Statistica\\_inferenziale](http://it.wikipedia.org/wiki/Statistica_inferenziale)

<sup>11</sup> <http://it.wiktionary.org/wiki/aleatorio>

<sup>12</sup> [http://it.wikipedia.org/wiki/Media\\_%28statistica%29](http://it.wikipedia.org/wiki/Media_%28statistica%29)

La media aritmetica è la tipologia impiegata più comunemente ed è quella alla quale si fa in genere riferimento nel parlare comune.

Essa è usata per riassumere, con un *solo numero*, un insieme di dati che si riferiscono a un fenomeno misurabile, ad esempio nel caso dell'altezza media di una popolazione.

In pratica, la media aritmetica di una distribuzione  $\mathbf{P}(\mathbf{x})$  di una qualunque variabile numerica  $\mathbf{X}$  è calcolata con la seguente formula:

$$\langle \mathbf{x} \rangle = \Sigma \mathbf{x} \mathbf{P}(\mathbf{x}) / \Sigma \mathbf{P}(\mathbf{x})$$

dove  $\Sigma \mathbf{P}(\mathbf{x})$  è definito il numero totale degli elementi.

Dalla teoria della probabilità e dalla statistica, ci deriva l'importante concetto di varianza di una variabile aleatoria  $\mathbf{X}$  (e della distribuzione di probabilità che questa segue) intesa come quel valore, indicato con  $\mathbf{Var}(\mathbf{X})$ , che fornisce la misura di quanto siano *vari* i valori assunti dalla variabile stessa, ovvero, di quanto essi si discostino dalla media  $\mathbf{E}[\mathbf{X}]$ <sup>13</sup>. Intendendo con il numero  $\mathbf{E}[\mathbf{X}]$  (da *expected value* o *expectation* in inglese o dal francese *espérance*) il valore *atteso* (chiamato anche media, speranza o speranza matematica) di una variabile casuale  $\mathbf{X}$  che formalizza l'idea *euristica*<sup>14</sup> di *valore medio* di un fenomeno aleatorio<sup>15</sup>.

La varianza è anche definita come lo *scarto quadratico medio* della media, indica la *larghezza* della distribuzione ed è definita dalla relazione:

$$(\Delta \mathbf{x})^2 = \Sigma (\mathbf{x} - \langle \mathbf{x} \rangle)^2 \mathbf{P}(\mathbf{x}) / \Sigma \mathbf{P}(\mathbf{x})$$

<sup>13</sup> <http://it.wikipedia.org/wiki/Varianza>

<sup>14</sup> Si definisce *procedimento euristico*, un metodo di approccio alla soluzione dei problemi che non segue un chiaro percorso, ma che si affida all'intuito e allo stato temporaneo delle circostanze, al fine di generare nuova conoscenza

<sup>15</sup> [http://it.wikipedia.org/wiki/Valore\\_atteso](http://it.wikipedia.org/wiki/Valore_atteso)

Si definisce, *mediana* (v.fig.1), o valore mediano, nell'ambito della statistica descrittiva, quando, data una distribuzione  $X$ , avente carattere quantitativo, oppure qualitativo ordinabile, (ovvero le cui modalità possano essere ordinate in base a qualche criterio), il valore/modalità, o l'insieme di valori/modalità assunto dalle unità statistiche si trovano nel mezzo della distribuzione<sup>16</sup>. Se si procede all'ordinamento delle unità in base ai valori crescenti del carattere da esse detenuto, la mediana bipartisce la distribuzione in due sottodistribuzioni: la prima a sinistra della mediana è costituita dalla metà delle unità la cui modalità è minore o uguale alla mediana stessa; la seconda a destra della valore mediano è costituita dalla metà delle unità la cui modalità è maggiore o uguale alla mediana.

*Moda* (v.fig.1) è definito «(...) il valore più probabile di  $X$ , ovvero il valore per cui  $P(x)$  è massimo (...)» (Rossi 2009). In altre parole, moda<sup>17</sup> della distribuzione di frequenza  $X$  è la modalità (o la classe di modalità) caratterizzata dalla massima frequenza ed è spesso rappresentata con la simbologia  $v_0$  intendendo quel valore che compare più frequentemente. Una distribuzione è *unimodale* se ammette un solo valore modale, è *bimodale* se ne ammette due (ossia: se esistono due valori che compaiono entrambi con la frequenza massima nella distribuzione data), *trimodale* se ne ha tre, ecc. Per la determinazione della classe modale (o gruppo modale) è opportuno ricorrere all'istogramma<sup>18</sup>, individuando *l'intervallo di altezza massima* (v.fig.1), ovvero il *punto di massimo della curva*; la classe con la maggiore densità media (che corrisponde all'altezza dell'istogramma) è, appunto, quella modale.

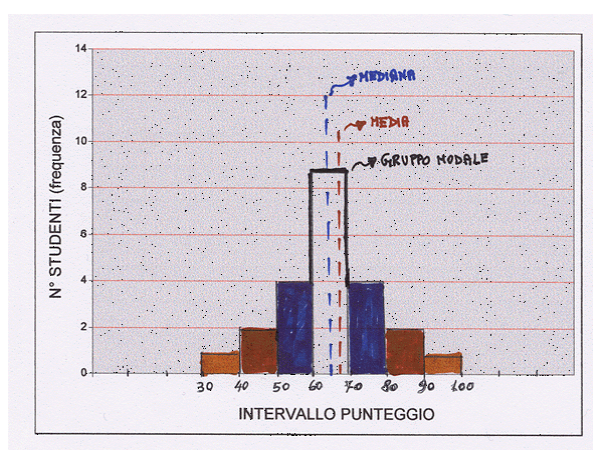


Figura 1: **Moda, mediana e media**

<sup>16</sup> [http://it.wikipedia.org/wiki/Mediana\\_%28statistica%29](http://it.wikipedia.org/wiki/Mediana_%28statistica%29)

<sup>17</sup> [http://it.wikipedia.org/wiki/Moda\\_%28statistica%29](http://it.wikipedia.org/wiki/Moda_%28statistica%29)

<sup>18</sup> la rappresentazione grafica di una distribuzione in classi di un carattere continuo

Nel caso particolare della distribuzione normale, detta anche *Gaussiana*, la moda coincide con la media e la mediana (v.fig.2)

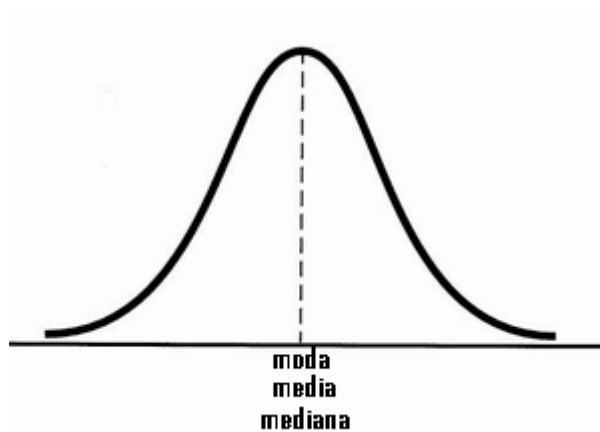


Figura 2: **Nella gaussiana la media, la moda e la mediana coincidono.**

Nella teoria della probabilità, s'intende per *distribuzione normale*<sup>19</sup> o Gaussiana, dal nome del matematico tedesco Carl Friederich Gauss, quella distribuzione di probabilità continua che è spesso usata come prima approssimazione per descrivere variabili casuali a valori reali che tendono a concentrarsi attorno a un singolo valore medio. Il grafico della funzione di densità di probabilità associata alla curva normale è a forma di campana, nota anche come *Campana di Gauss* (o anche come curva degli errori, curva a campana, ogiva) (v.fig.2)

---

<sup>19</sup> [http://it.wikipedia.org/wiki/Variabile\\_casuale\\_normale](http://it.wikipedia.org/wiki/Variabile_casuale_normale)

Nell'ambito della statistica e nella valutazione dei dati empirici è importante ricordare la *teoria dell'errore*. Nella maggior parte dei casi, durante le operazioni di valutazione delle entità quantitative, infatti, si producono inevitabilmente delle *inesattezze*. Esse possono essere causate, sia dalle imprecisioni degli strumenti di misurazione utilizzati (*errore sistematico*) che, più spesso, dalla limitatezza del campione utilizzato per l'esame (*errore casuale*).

Si parla di errore casuale quando, le ripetute misurazioni di una medesima quantità, danno risultati differenti. La mancata corrispondenza nelle stime di una stessa quantità costituiscono, esse stesse, una *distribuzione* che può essere studiata valutando gli stessi parametri utilizzati per il caso generale. Si avrà, quindi, che la *media della distribuzione* è il *valore atteso* della misura, ossia quello che appare più probabile per la quantità che realmente ci interessa.

L'errore da attribuire alla misura stessa, più precisamente l'intervallo di valori, entro il quale è plausibile che possa trovarsi il valore reale della quantità oggetto di interesse, risulta, invece, strettamente legato alla varianza della distribuzione e si considera misurato dal seguente rapporto:

$$\frac{\text{Deviazione standard}}{\sqrt{\text{numero totale misure effettuate}}}$$

Questa definizione è appropriata nel caso di un errore casuale, che è quello derivante dall'esito di operazioni di misurazione, poiché esse non influenzano il risultato, se non nella misura che può essere definita dal *calcolo delle probabilità*.

Un'altra tipologia di errore è quello sistematico, altrimenti detto *determinato*, che è definito come lo scostamento tra il valore ottenuto dalla misurazione e quello *reale* della grandezza studiata. L'errore sistematico rappresenta «l'indice di accuratezza dei dati» ed è chiamato in questo modo, perché è costante al ripetersi della misura.

Inoltre, l'errore sistematico non può essere eliminato con l'aumento delle misurazioni, così come avviene, invece, per l'errore statistico o casuale. Inoltre, per questa tipologia di errore, è particolarmente difficoltoso determinarne l'entità, se non ricorrendo a metodi di rilevamento alternativi o alla *valutazione di coerenza* e di qualità all'interno dei sistemi di dati. Tra gli errori sistematici vanno ricordati quello *strumentale, di metodo* e *l'errore personale*.

Le possibili cause dell'errore strumentale, ad esempio, sono rappresentate dall'utilizzo di uno strumento difettoso, alterato o non calibrato<sup>20</sup>.

Dalla teoria dell'errore deriva la nozione di *cifre significative*, considerando significative «(...) nella presentazione del risultato di una misura, soltanto le cifre fino a quella che può essere modificata restando all'interno dell'intervallo definito dalla stima dell'errore (...)» (Rossi 2009).

L'introduzione nella misura delle *cifre non significative* da una parte, non rende più precisa l'informazione e dall'altra, la rende ancor meno intelleggibile. Inoltre, un'importante caratteristica dell'errore nelle misure è definita nella sua proprietà di *propagazione*. Essa si verifica quando la stima di una quantità risulta determinata dalla sommatoria di dati differenti; in questo caso anche l'errore sarà in percentuale maggiore rispetto alla stima degli errori attribuiti separatamente ai singoli dati. In altre parole, la percentuale dell'errore «(...) finale sarà data dalla somma degli errori relativi parziali (...)» (Rossi 2009). Per tutti questi aspetti, propri degli errori, è opportuno, molte volte, diffidare della *falsa accuratezza* delle stime nel campo delle scienze umane.

Un'importante risultato della teoria dell'errore è il *teorema del limite centrale*, che stabilisce come la distribuzione dei risultati di una misura tenda, per ipotesi generali e quando il numero delle misure è molto elevato, ad assumere una forma di una curva a campana o curva di Gauss che matematicamente si esprime nel modo seguente:

$$P(x) = N/\sqrt{2\pi \Delta x^2} \exp\{ - (x-\langle x \rangle)^2/2 \Delta x^2 \}$$

dove  $N$  corrisponde al numero totale delle misure effettuate.

<sup>20</sup> [http://it.wikipedia.org/wiki/Errore\\_sistematico](http://it.wikipedia.org/wiki/Errore_sistematico)



Il principio mostrato nella formula presenta una grande rilevanza teorica e pratica poiché permette la formulazione di numerose proprietà delle distribuzioni dei risultati di misure che, a esse, sono collegate. Inoltre, è importante perché un scostamento significativo dalla distribuzione *normale* può denotare la presenza di errori sistematici.

La distribuzione normale è considerata il caso base delle distribuzioni di probabilità continue a causa del ruolo che ricopre nel teorema del limite centrale. Più specificamente, «(...) assumendo certe condizioni, la somma di  $n$  variabili casuali, con media e varianza finite, tende a una distribuzione normale al tendere di  $n$  all'infinito (...)». In virtù di questo teorema, la distribuzione normale si incontra spesso nelle applicazioni pratiche ed è usata, in statistica e nelle scienze naturali e sociali, come un semplice modello di riferimento per descrivere fenomeni complessi.

## 3 L'analisi qualitativa dei testi

### 3.1 Cenni

Con *analisi testuale* si intende, nell'ambito del settore della critica, l'insieme delle metodologie che eseguono l'esame qualitativo sui corpora. Tali procedure si basano sulla realizzazione di indici per individuare le caratterizzazioni sintattiche di un linguaggio. In altre parole, per realizzare l'analisi sui testi, si utilizzano procedure informatiche generate *ad hoc* in grado di compiere operazioni di *text retrieval*, cioè di reperimento e recupero delle informazioni nella forma di stringhe o di sequenze di caratteri, all'interno di un corpus testuale, consentendo la produzione di indici. In genere, il punto di partenza dell'analisi qualitativa di questo tipo è lo studio delle diverse parole che compongono il testo digitalizzato ordinate per valori decrescenti di frequenza (vocabolario). Tra le parole ad alta frequenza in un vocabolario, oltre a quelle con *contenuto strumentale*<sup>21</sup> (di, e, che, per, etc.) poco caratterizzanti e presenti in tutti i testi, è interessante eseguire l'analisi delle *parole tema* che presentando contenuti informativi rilevanti e un'elevata frequenza di comparizione. Esse consentono di cogliere, immediatamente, gli argomenti principali del testo e di formularne le peculiarità.

I task dell'analisi qualitativa che descrivono in modo semiautomatico il contenuto di un testo sono i seguenti:

1. analisi dei segmenti ripetuti;
2. analisi delle co-occorrenze;
3. analisi delle parole caratteristiche;
4. analisi del linguaggio peculiare;
5. analisi delle parole con caratteristiche grammaticali omogenee;
6. analisi delle concordanze;
7. analisi delle corrispondenze lessicali.

---

<sup>21</sup> [http://it.wikipedia.org/wiki/Analisi\\_testuale](http://it.wikipedia.org/wiki/Analisi_testuale)

- L'analisi dei *segmenti ripetuti* include l'individuazione di tutte quelle forme composte che sono formate dalle parole che compaiono nel *corpus* con la medesima sequenza. Alcuni segmenti sono particolarmente rilevanti dal punto di vista semantico, poiché consentono di delimitare i significati delle parole presenti nel testo (ad esempio come nel caso dell'analisi delle associazioni di termini come: carta di credito, Capo dello Stato, politica economica, guardia di finanza);
- l'analisi delle *co-occorrenze* consente di studiare le associazioni tra parole, individuando quei termini che compaiono più spesso vicini tra loro;
- l'analisi delle *parole caratteristiche* (o specifiche) permette, invece, di differenziare le diverse parti di un testo, evidenziando i termini che sono *sovrarappresentati* nel linguaggio di una certa categoria di autori; per esempio, nel caso della caratterizzazione del linguaggio femminile, oppure per analizzare la tipologia di articoli pubblicati su una determinata testata, ecc.;
- per l'analisi del *linguaggio peculiare* si utilizzano *lessici di frequenza*, costruiti, di solito, per rappresentare il linguaggio usato comunemente da una determinata comunità linguistica. In queste liste, composte da milioni di occorrenze e derivate da fonti diverse (stampa, linguaggio parlato, letteratura, etc.), a ciascuna parola è associata una frequenza che indica l'uso "atteso" di quel termine all'interno della comunità linguistica cui il lessico è riferito. Eseguendo il confronto tra il vocabolario del testo in esame e quello del lessico di frequenza prescelto, è possibile elaborare liste di parole sovrarappresentate e che corrispondono al linguaggio peculiare del testo stesso;
- anche l'insieme delle parole con *caratteristiche grammaticali omogenee* può essere utile per l'analisi qualitativa del testo. Ad esempio, l'insieme dei verbi ricondotti al lemma può fornire una graduatoria delle azioni menzionate nel testo. L'insieme degli aggettivi, invece, fornisce tutti gli elementi per valutare il tono di un testo e l'insieme e il tipo di pronomi, ad esempio, può indicare il tipo di interazione presente tra soggetti e che caratterizza il testo;

- l'analisi delle *concordanze* (Key word in context), invece, è una tecnica che consente di analizzare il *contesto d'uso* di una parola di interesse (o di gruppi di parole con la stessa radice) tramite la visualizzazione delle  $n$  parole precedenti e delle  $n$  parole successive alla parola  $x$  in analisi ed è effettuato per tutte le volte che la stessa parola compare nel *corpus*. L'analisi delle concordanze è indispensabile, sia per risolvere alcune ambiguità semantiche, che per ricostruire per ogni parola, i riferimenti tematici cui questa rinvia, tracciando così una mappa concettuale tra parole e temi affrontati.
- infine, l'analisi delle *corrispondenze lessicali* consente di sintetizzare l'informazione contenuta in una grossa matrice di dati testuali, visualizzando sul piano fattoriale l'associazione tra le forme (parole) all'interno del testo in analisi e cercando la migliore rappresentazione simultanea degli elementi di riga e colonna, in modo da studiarne l'interdipendenza tra caratteri.

L'input, su cui impiegare l'analisi qualitativa, è rappresentato dal corpus digitalizzato e preferibilmente *tokenizzato*<sup>22</sup>. S'intende per token l'unità di base per l'elaborazione linguistica e tokenizzare significa, in particolare, isolare ciascuna parola del corpus dando a essa un numero progressivo del tipo:

`<word id="w_00n"> parola </word>`

---

<sup>22</sup> [http://it.wikipedia.org/wiki/Token\\_%28testo%29](http://it.wikipedia.org/wiki/Token_%28testo%29)

## 4 Lo scenario cognitivo e culturale dei testi letterari

### 4.1 L'analisi *quantitativa* del testo

L'analisi del testo, presentata in questa tesi, rappresenta un'interessante *variante* dell'analisi qualitativa sui testi descritta al precedente paragrafo. Essa, infatti, si basa sull'impiego di metodologie *quantitative* fondate sui principi della fisica e applicate, nell'ambito tipico delle scienze umanistiche, allo studio dei testi scritti. Lo studio quantitativo sui testi letterari, in particolare, ha come obiettivo la definizione e lo studio della *geografia culturale* espressa nelle opere, definendo gli orizzonti cognitivi che da essa emergono e che caratterizzano l'autore al momento della loro scrittura. Gli orizzonti cognitivi delineati dall'analisi quantitativa sui testi, saranno inseriti nel loro contesto socio culturale di provenienza, sia per confermarne la fondatezza, che per riuscire a elaborare un'interpretazione documentata e storica del fenomeno evidenziato.

Gli studi qualitativi relativi ai linguaggi hanno una lunga storia e siamo consapevoli che ogni tentativo di definire con indicatori *quantitativi* le percezioni soggettive, oltre a presentare aspetti problematici, può dare origine a dubbi e reticenze filosofiche. Riteniamo che, quando si esaminano gli artefatti culturali in forma di testi scritti di qualsiasi natura, una *parziale semplificazione* possa essere rappresentata dalla caratteristica stessa del testo scritto; si tratta, infatti, di un insieme di informazioni codificate in un corpus che presentano un carattere, sostanzialmente, limitato e selezionato. Da questo presupposto deriva che, i parametri, ricavati dall'analisi quantitativa dei testi, possano essere valutati nella stessa maniera in cui si valuterebbero le grandezze in ambienti di laboratorio statici e controllati avulsi da qualsiasi contaminazione o modificazione che possa svilirne la scientificità. In altre parole, i testi scritti rappresentano, per il nostro studio, una *realtà fossilizzata* sulla quale realizzare uno studio scientifico, impiegando parametri quantitativi e i fondamenti derivanti dalle metodologie della fisica. Partendo, quindi, dall'analisi dei testi vedremo come sarà possibile evidenziare tutta una serie di indicatori per la determinazione della geografia cognitiva espressa dagli autori nei loro testi.

## 4.2 Le metodologie fisiche per le scienze umane

Negli ultimi decenni, a seguito della crescita della quantità dei dati aventi, molto spesso, una natura quantitativa, resa facilmente accessibile dal web e manipolabile grazie all'impiego di strumenti di calcolo elettronico, è apparso sempre più evidente come metodologie, nate nel contesto delle scienze fisiche, possano essere utilmente applicate anche in ambiti di ricerca apparentemente molto lontani. Gli esempi sono numerosi e riguardano i campi della linguistica, dell'arte, dell'archeologia, della filosofia, della storia, l'ambito economico o sociologico ecc. I casi di studio che andremo ad analizzare rappresentano un interessante esempio di impiego delle metodologie fisiche, dei relativi principi e algoritmi formali, nell'analisi quantitativa dei testi letterari del periodo del Medioevo.

## 4.3 Gli obiettivi prefissati

Gli obiettivi da realizzare sono rappresentati dalla volontà di proporre una *differente* analisi del testo basata sull'impiego di parametri quantitativi destinati, di solito, all'esame dei fenomeni appartenenti alle scienze *dure*. Le metodologie applicate si basano, infatti, sui fondamenti della fisica e della statistica e prevedono l'uso di algoritmi formali al fine di ottenere una serie di parametri numerici che tracciano lo scenario cognitivo in ambito letterario e sulla base dei quali realizzare interessanti interpretazioni e studi di carattere socio culturale.

In altre parole, si procederà all'illustrazione e alla dimostrazione del legame esistente tra i metodi qualitativi, propri delle scienze umanistiche, e le metodologie quantitative appartenenti alle scienze fisiche e statistiche. Si dimostrerà, attraverso l'analisi e l'illustrazione di due casi di studio, l'efficacia dell'applicazione delle metodologie fisiche impiegate nell'indagine qualitativa dello scenario cognitivo e la relativa rispondenza con le elaborazioni storico-culturali che faranno da cornice nella comprensione dei testi in esame.

L'opportunità di trasformare le informazioni *qualitative* tipiche dell'universo delle percezioni umane, in informazioni *quantitative*, intese come entità misurabili, è compresa dalla *meccanica classica* e, in particolare, dalle nozioni di *baricentro*, di *momento di inerzia* e di *assi principali*. Tali principi della fisica saranno impiegati nell'esame dei due casi di studio per implementare gli algoritmi formali che daranno come risultato i seguenti parametri numerici:

- il *Centro* di percezione, considerato come il luogo fisico o psicologico in cui, molto probabilmente, l'autore si trovava al momento della scrittura del testo;
- il *Raggio* di *percezione* calcolato come la distanza dal centro di percezione fino all'orizzonte cognitivo dell'autore;
- l'*Orizzonte della percezione* (che chiameremo anche scenario cognitivo) dell'autore che si definisce come la *regione spaziale* che traccia il *limite massimo* al di fuori della quale, gli eventi registrabili assumono un interesse limitato, se non addirittura trascurabile;
- gli *Assi principali* della *percezione* che identificano le *direzioni privilegiate* dell'autore all'interno dello spazio cognitivo. Essi saranno costituiti da un asse maggiore (più esteso) e uno minore ortogonale al primo. Gli assi principali formeranno l'ellisse della percezione, posizionata nello spazio a una certa angolazione rispetto al Polo Nord. I parametri che concorreranno a formare l'ellisse della percezione saranno, anch'essi le risultanze dell'applicazione degli algoritmi.

Riassumendo, questo studio assume l'interessante compito di determinare, utilizzando calcoli matematici e trigonometrici, la geografia cognitiva dei nostri casi di studio che coincide con la cognizione mentale e culturale dell'autore al momento della stesura dell'opera. Lo scenario cognitivo che disegneremo, nella sua connotazione numerica, e che è espresso nell'opera potrà essere la conseguenza di una scelta conscia o inconscia dell'autore e strettamente legata, in altre parole, alla sua limitata o illimitata conoscenza o attenzione nei confronti del mondo circostante.

## 4.4 Procedure operative e concetti applicati

Per eseguire il calcolo e la successiva estrapolazione dei parametri quantitativi riferiti al testo, e allo scopo di determinare la geografia cognitiva e culturale dell'autore, in concreto, occorre individuare, per prima cosa, all'interno del testo tutte le indicazioni geografiche e/o spaziali (oggetti geografici) presenti, e di seguito contarne le rispettive *occorrenze*. Intendendo per *occorrenza*<sup>23</sup> il numero complessivo delle *apparizioni* degli *n* oggetti geografici all'interno dell'artefatto culturale. Gli oggetti geografici possono essere stati menzionati dall'autore sia in maniera *esplicita* sia in quella *implicita*. Potranno essere, nel dettaglio, costituiti dai nomi di luoghi, regioni, monti, fiumi, edifici localizzabili in un territorio ecc., e facilmente rintracciabili nel corpus perché chiaramente nominati, oppure, in modo implicito, essere rappresentati dall'uso di perifrasi<sup>24</sup> e di allusioni (ad esempio, nel Decameron abbiamo incontrato citazioni come «il Santo Sepolcro» per riferirsi alla città di Gerusalemme, oppure «la nostra città» per sottintendere Firenze). S'intuisce come, in questi frequenti casi, l'operazione di rilevazione degli oggetti geografici e il relativo conteggio delle occorrenze rappresenti una problematica la cui risoluzione richieda maggiore attenzione e cura.

Approfondiremo e dettaglieremo le difficoltà incontrate, nell'esame di entrambi i due casi di studio, al par. 5.2.4 e 6.2.5 di questa relazione.

---

<sup>23</sup> [http://dizionari.corriere.it/dizionario\\_italiano/O/occorrenza\\_2.shtml](http://dizionari.corriere.it/dizionario_italiano/O/occorrenza_2.shtml)

<sup>24</sup> <http://www.wordreference.com/definizione/perifrasi>



Riepilogando, sono state definite le occorrenze rinvenute dall'esame di un testo. Si intende con il termine di «occorrenze» il numero complessivo delle comparse degli  $n$  oggetti geografici all'interno dell'artefatto culturale; «oggetto geografico», invece, è qualsiasi riferimento a un luogo, corso d'acqua, nazione, città, monte o castello ecc., presente in un corpus.

Dato un certo numero di occorrenze che rappresentiamo con il simbolo  $o(n)$ , per riferirsi all' $n$ -esimo oggetto geografico rintracciato nel testo, un parametro indispensabile nel calcolo è rappresentato dal *peso* dell'oggetto identificato applicando la formula:

$$w(n) = o(n) / \sum o(n)$$

Da notare che *la somma del peso di tutti gli oggetti geografici* deve essere uguale a 1.

Il primo elemento da computare e che costituisce lo scenario cognitivo dei corpora è rappresentato dal *centro della percezione* che definiamo concettualmente come il luogo fisico o psicologico in cui, molto probabilmente, l'autore si trovava al momento della scrittura del testo. Il centro della percezione è rappresentato da un punto nello spazio definito dalle sue coordinate *pesate* indicate con le incognite X, Y e Z e ricavate con l'ausilio della trigonometria sferica. In altre parole, sulla base delle coordinate geografiche degli oggetti rinvenuti nel testo e implementando una serie di algoritmi formali, saranno delineate le coordinate pesate del centro della percezione. Sulla base di queste vengono definite le coordinate geografiche indicate con le lettere A e B. Queste ultime rappresentano la posizione del centro della percezione nello spazio tradotto nella *latitudine* e nella *longitudine* e in formato decimale.

Si definisce latitudine di un punto, la *distanza angolare* di quel punto *dall'equatore* (v.fig.3).

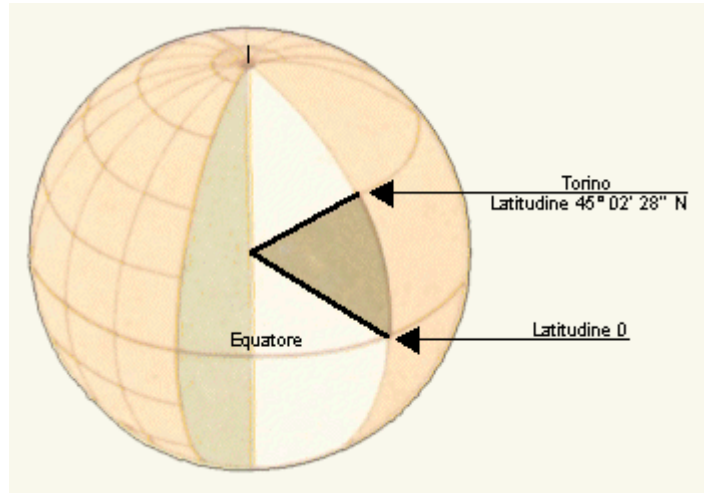


Figura 3: **Rappresentazione grafica della latitudine di Torino**

La longitudine dello stesso punto, invece, definisce la *distanza angolare* di quel punto dal *meridiano di Greenwich*.

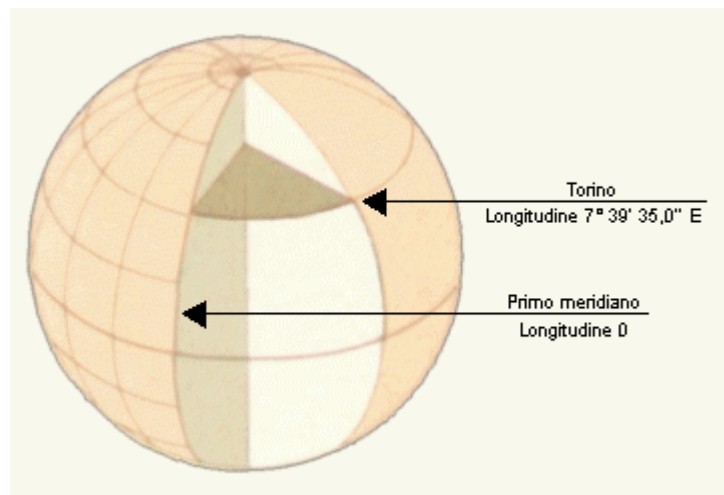


Figura 4: **Rappresentazione grafica della longitudine sempre di Torino**

Viene definita la posizione dell'oggetto geografico tramite l'individuazione di due *angoli*, misurati in gradi, *minuti di arco* e *secondi di arco* e, in particolare, sia la latitudine che la longitudine si indicano con i simboli ( ° , ' , " ). Ad esempio, un punto nello spazio espresso con il riferimento geografico pari a 35° 43' 9" indicherà un angolo di 35 gradi, 43 minuti e 9 secondi, considerando che 1 grado (°) è costituito da 60 minuti di arco, e che 1 minuto (') è formato da 60 secondi (") di arco.

Storicamente, l'ordine con cui si indicavano le coordinate è sempre lo stesso: per prima si indica la *latitudine* e poi la *longitudine*. Ma le espressioni che indicano i gradi possono avere diversi formati:

- *Gradi minuti secondi (DMS)*  
Esempio: 41° 53' 24" N, 12° 29' 32" E (le coordinate del Colosseo)
- *Gradi minuti decimali (DM)*  
Esempi: 49° 33.0', -123° 30.0' oppure 49d 30.0m, -123d 33.0m
- *Gradi decimali (DD)*  
di solito da 4 a 6 cifre decimali. Esempio: 49.5000°, -123.5000°

Per il nostro lavoro useremo questo tipo di indicazione *in gradi decimali* tenendo conto di 4 cifre decimali dopo la virgola.

Per rappresentare gli emisferi N (nord) / S (sud) e E (est) / O (ovest, spesso W per west) è possibile indicare il valore con il segno; in particolare, i valori negativi si riferiscono alle latitudini nell'emisfero Sud e alle longitudini che si trovano a Ovest del meridiano fondamentale. In un sistema di coordinate geografiche, il meridiano fondamentale a cui è assegnata la longitudine di 0°, è quello che viene preso come riferimento per la definizione della longitudine degli altri punti della superficie del corpo celeste<sup>25</sup>.

---

<sup>25</sup> [http://it.wikipedia.org/wiki/Meridiano\\_fondamentale](http://it.wikipedia.org/wiki/Meridiano_fondamentale)

Nel foglio elettronico progettato per contenere i parametri e gli algoritmi formali si implementeranno tutte le risultanze ricavate dall'analisi dei testi all'esame: verranno inseriti, per ogni oggetto geografico, i valori delle coordinate geografiche espressi in gradi decimali (e derivanti dalla fase denominata *georeferenziazione*<sup>26</sup>). In altre parole, ogni oggetto geografico verrà identificato con la propria latitudine e longitudine e, in base a questi valori, gli algoritmi effettueranno l'elaborazione del corrispondente valore espresso in *radianti*. Definendo radiante<sup>27</sup> l'angolo al centro di una circonferenza, di raggio arbitrario, che sottende un arco di lunghezza uguale al raggio stesso. Precisamente, i radianti si ottengono dal prodotto delle coordinate geografiche decimali (latitudine e longitudine) per il coefficiente  $\pi/180$ .

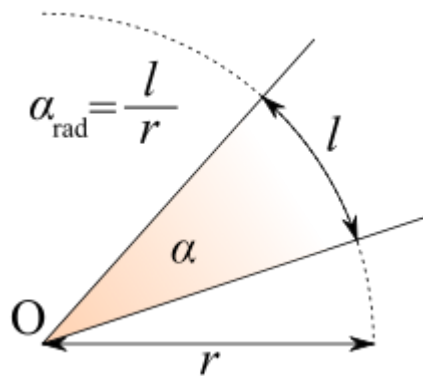


Figura 5: Un angolo misurato in radianti<sup>28</sup>

Sia la latitudine che la longitudine, che individuano i punti corrispondenti alla posizione degli oggetti geografici, sono coordinate definite sullo *sferoide* terrestre; entrambe rappresentano angoli definiti in gradi minuti e secondi d'arco.

<sup>26</sup> <http://it.wikipedia.org/wiki/Georeferenziazione>

<sup>27</sup> <http://progettomatematica.dm.unibo.it/Trigonometria/archi.htm>

<sup>28</sup> <http://en.wikipedia.org/wiki/Radian>

Si nota che  $n$  come grado di latitudine corrisponde sulla superficie terrestre a circa 111 km, così come lo si può calcolare dalla proporzione:

$$1^\circ : 360^\circ = d : 2\pi R$$

dove  $d$  è l'equivalente in km di 1 grado di latitudine e  $R$  è il raggio terrestre

Mentre un grado, in termini di longitudine, sulla superficie terrestre, corrisponde a 72,5 km. Nel foglio elettronico, entrambi i coefficienti sono definiti con le costanti in km rappresentate dai valori 111 e 72,5 utilizzati nella trasformazione dei gradi decimali che identificano la latitudine e la longitudine in chilometri, per poter esprimere le distanze terrestri.

Il *raggio della Terra*<sup>29</sup> (che abbiamo indicato nella precedente formula con  $R$ ) può essere definito con la considerazione che il nostro pianeta non è una sfera perfetta ma un ellissoide appiattito in corrispondenza dei poli Nord e Sud, per questo è chiamato anche sferoide oblato. Per questa caratteristica dello sferoide non esiste un singolo valore che possa rappresentare il suo raggio che, di solito, è una caratteristica delle sfere perfette. In questo caso, con raggio della Terra viene intesa la distanza tra un «alcuni "centri"» (en.Wikipedia, voce *Earth radius*) della Terra e un punto sulla superficie di questa, oppure su una superficie ideale che modella la Terra. Può anche essere intesa, più chiaramente, come una sorta di media di tali distanze. In altre parole, tale definizione sta a rappresentare il raggio di una sfera la cui curvatura corrisponde alla curvatura del modello ellissoidale della Terra in un dato punto. Il raggio della Terra si collega al concetto di *raggio equatoriale* (definito come il raggio della circonferenza immaginaria quale l'equatore) che è, approssimativamente, stimato pari a 6378,388 km<sup>30</sup>.

La *distanza euclidea*, compare anch'essa come risultanza degli algoritmi che andremo a definire; essa è considerata come la *comune distanza* tra due punti. Essa rappresenta la distanza tra due oggetti geografici come risulterebbe dall'ideale misurazione su una superficie piana facendo uso di una normale riga.

---

<sup>29</sup> [http://en.wikipedia.org/wiki/Earth\\_radius](http://en.wikipedia.org/wiki/Earth_radius)

<sup>30</sup> [http://it.wikipedia.org/wiki/Raggio\\_terrestre](http://it.wikipedia.org/wiki/Raggio_terrestre)

Inoltre, la distanza euclidea è ottenuta dall'applicazione ripetuta del *Teorema di Pitagora*<sup>31</sup> ed è per questa ragione che la letteratura tradizionale si riferisce ad essa come alla *metrica pitagorica*.

La formula espressa di seguito, ha come risultanza la metrica pitagorica; lo spazio euclideo diventa, a tutti gli effetti, uno spazio metrico.

Per due punti **P** e **Q** con le rispettive coordinate:

$$P = (p_x, p_y) \text{ e } Q = (q_x, q_y)$$

avremo che la distanza euclidea è calcolata come segue:

$$\sqrt{(p_x - q_x)^2 + (p_y - q_y)^2}$$

---

<sup>31</sup> [http://it.wikipedia.org/wiki/Distanza\\_euclidea](http://it.wikipedia.org/wiki/Distanza_euclidea)

Si definiscono, i concetti di *seno*<sup>32</sup> (abbreviato *sin* o *sen*) e di *coseno*<sup>33</sup> (abbreviato *cos*) che si determineranno per tutti gli oggetti geografici ricavati dai testi in esame.

Dato un triangolo rettangolo, il *seno* di uno dei due angoli interni adiacenti all'ipotenusa, si definisce come il rapporto tra le lunghezze del cateto *opposto* all'angolo e dell'ipotenusa.

*Coseno* invece, dato, ancora, un triangolo rettangolo, è il rapporto tra le lunghezze del cateto adiacente all'angolo e dell'ipotenusa.

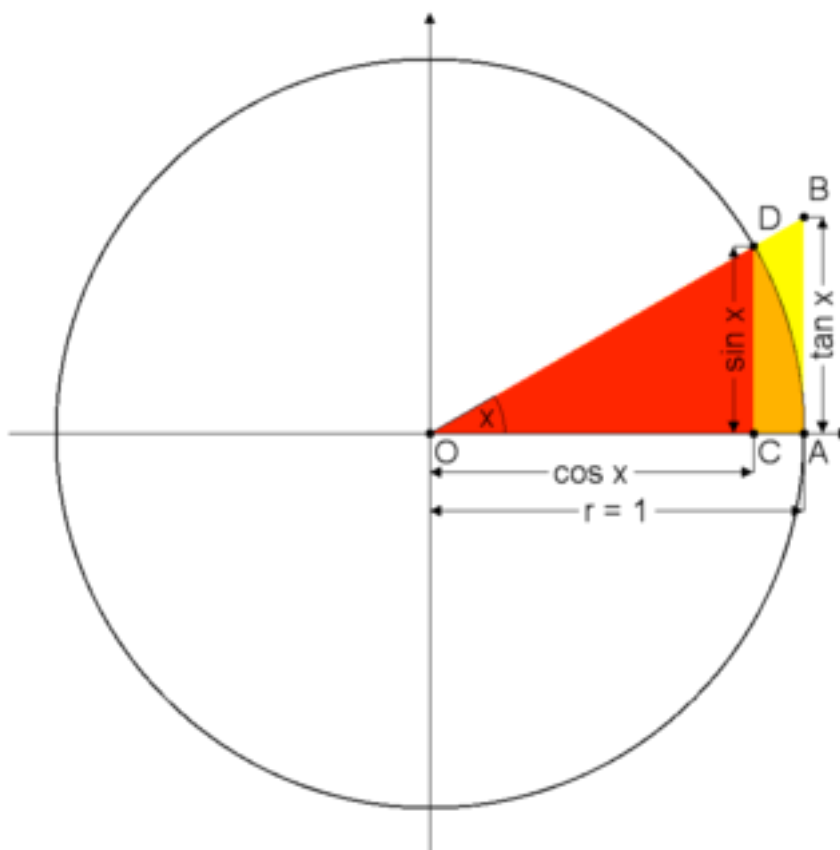


Figura 6: **Rappresentazione grafica del seno e del coseno**

<sup>32</sup> [http://it.wikipedia.org/wiki/Seno\\_%28matematica%29](http://it.wikipedia.org/wiki/Seno_%28matematica%29)

<sup>33</sup> <http://it.wikipedia.org/wiki/Coseno>

Dopo aver annotato, all'interno del foglio elettronico, entrambe le coordinate decimali di tutti gli  $n$  oggetti geografici, con l'ausilio degli algoritmi formali ispirati ai principi della trigonometria sferica, vengono calcolate le *coordinate pesate*  $X, Y, Z$  applicando le seguenti formule:

$$X = \Sigma w(n) \cos a(n) \cos b(n)$$

$$Y = \Sigma w(n) \cos a(n) \sin b(n)$$

$$Z = \Sigma w(n) \sin a(n)$$

Il concetto trigonometrico di *arcotangente*, che andremo a calcolare, determinerà i parametri delle coordinate geografiche decimali del centro di percezione.

Si dice arcotangente, l'inverso della *funzione tangente* nell'intervallo:

$$\left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \subset \mathbb{R}$$

$$\arctan \alpha = \cos \alpha / \sin \alpha$$

Intendendo per *tangente* di un angolo il rapporto tra il sin e il cos dell'angolo stesso<sup>34</sup>.

La tangente dell'angolo  $\alpha$  si esprime con il rapporto tra il seno e il coseno dell'angolo stesso, in questo modo:

$$\tan \alpha = \sin \alpha / \cos \alpha$$

<sup>34</sup> <http://edu.os3.it/html/manual/trigonometria.pdf>



Il termine *arcotangente* deriva dalla locuzione «*uno degli archi la cui tangente*» è la misura dell'angolo; infatti i radianti corrispondono alla lunghezza dell'arco di una circonferenza di raggio unitario.

Con maggior precisione, si può affermare che l'arcotangente di  $x$  è l'angolo di valore assoluto minore, la cui tangente è  $x$ .

Dalle coordinate pesate  $X$ ,  $Y$  e  $Z$  si sono ricavate le coordinate geografiche  $A$  e  $B$  del centro di percezione e risultanti rispettivamente dalle formule:

$$A = \arctan \frac{Z}{\sqrt{X^2+Y^2}}$$

$$B = \arctan \frac{Y}{X}$$

La formula matematica descritta è basata sul concetto di *baricentro* (a volte chiamato centroide) che, di una figura  $X$  *n-dimensionale* in uno spazio euclideo *n-dimensionale*, si definisce come l'intersezione di tutti gli iperpiani che dividono  $X$  in due parti di misura identica; in modo più informale, si può definire come il calcolo della *media* di tutti i punti di  $X$ . I risultati evidenziati, mediante l'applicazione di questa formula, permette di valutare le distanze *in linea d'aria* e si applicano a tutti quei casi in cui diventerebbe difficoltoso, se non addirittura impossibile, valutarle *dal terreno*.

La distanza indicata con  $d(n)$  dell'ennesimo oggetto geografico calcolato dal centro di percezione, definito, quest'ultimo, con le coordinate decimali  $A$  e  $B$ , viene calcolata con la formula:

$$d(n) = T \arccos[\cos a(n) \cos A \cos (b(n)-B) + \sin a(n) \sin A]$$

identificando con  $T$  il raggio terrestre (calcolato in 6.366 km )

Si definisce il *raggio di percezione*  $R$  come la *media pesata* delle distanze degli oggetti geografici presenti nel testo e si ottiene con la seguente formula:

$$R = \sum_n w(n) d(n)$$

Per la maggior parte delle distribuzioni si può registrare la presenza *di più della metà* delle citazioni geografiche di un testo *all'interno del raggio di percezione* rappresentato con il simbolo  $R$  e calcolato con la formula appena riportata. Questo particolare aspetto evidenzia l'importanza che riveste questo parametro nella determinazione dello scenario cognitivo dei testi esaminati. Più precisamente, se si volesse calcolare l'incidenza delle distribuzioni all'interno del raggio di percezione, ipotizzando il caso di una distribuzione gaussiana *isotropa*, intendendo con l'aggettivo isotropa<sup>35</sup> la proprietà d'*indipendenza* della stessa distribuzione dalla direzione, la frazione esatta di citazioni incluse nel raggio di percezione sarebbe determinata dalla formula:

$$1 - e^{-\pi/4} = 0,544$$

---

<sup>35</sup> <http://it.wikipedia.org/wiki/Isotropia>

Per testare l'efficacia della formula appena descritta, nell'analisi quantitativa dei testi, si deve procedere ordinando gli oggetti geografici in maniera crescente, sulla base delle loro distanze dal centro di percezione. Quindi, si deve calcolare la somma dei rispettivi pesi, partendo dal peso del centro di percezione fino ad arrivare al valore di 2/3 del totale.

La distanza, entro la quale sono racchiusi i 2/3 delle citazioni geografiche in un testo, può essere considerata una *valida alternativa* al raggio di percezione ma solo nel caso si tratti di una distribuzione gaussiana isotropa.

In alcuni testi non sarebbe, in realtà, necessario procedere al calcolo del centro di percezione in quanto il corpus stesso ci offre un'indicazione empirica o intuitiva per la sua definizione. In questi casi basterà operare con la *formula della distanza* (o con le distanze empiriche note) per determinare il valore del raggio e quindi l'ampiezza dello spazio cognitivo.

Anche in questi casi, e cioè per questi testi in cui si intuisce quale possa essere il centro della percezione, perché specificato ad esempio chiaramente dall'autore, l'applicazione dei parametri relativi all'analisi quantitativa del testo rimane, comunque, un'importante occasione di verifica e di riscontro della fondatezza delle teorie fisiche applicate.

Riepiloghiamo le *fasi operative* che caratterizzano l'analisi quantitativa sui testi oggetto dei casi di studio e l'implementazione delle celle del foglio elettronico che verranno dettagliate al par. 5.2.1 e 6.2.1 per entrambi i casi di studio:

- abbiamo rintracciato separatamente, all'interno di ciascun corpus, le informazioni geografiche presenti, associandole al totale delle rispettive occorrenze. In altre parole, si è proceduto al *materiale conteggio*, all'interno del testo, per ciascun oggetto geografico (indicando nel foglio elettronico e in colonne separate: il nome originale utilizzato nel testo e il moderno toponimo corrispondente) del numero delle volte in cui esso è apparso all'interno del corpus;
- ogni oggetto geografico è stato *georeferenziato* con le proprie coordinate geografiche espresse in gradi decimali;

- tramite l'applicazione degli algoritmi basati sui principi della trigonometria sferica si sono calcolate le *coordinate pesate* X, Y e Z del centro di percezione;
- e da queste ultime, le corrispondenti coordinate geografiche A e B espresse in gradi decimali;
  
- è stata calcolata la *distanza euclidea* di ogni oggetto geografico dal centro di percezione, che era stato definito nella fase precedente dalle coordinate decimali A e B, tenendo conto del valore rappresentato dal raggio terrestre, assunto come costante (e calcolato in 6.366 km );
  
- infine, si è definito il *raggio di percezione R* come la *media pesata* delle distanze degli oggetti geografici presenti nel testo.

## 4.5 Il problema degli *oggetti estesi*

L'analisi di entrambi i casi di studio ha comportato una serie di problematiche che è stato necessario pianificare a inizio analisi e risolvere in corso d'opera, al fine di realizzare un'indagine che desse luogo al minor numero possibile di errori materiali e di imprecisioni. Una prima difficoltà incontrata nei due casi di studio e che è comune, riteniamo, alla gran parte dei testi sui quali si volesse effettuare un'analisi quantitativa, è rappresentata dal *trattamento* dei riferimenti a *oggetti geografici estesi*; oggetti cioè caratterizzati da una superficie geografica molto estesa e per questo, difficoltosi da georeferenziare. Nel dettaglio, ci riferiamo ai nomi di superfici molto estese nominate nei testi, come nel caso di regioni, catene montuose, fiumi, ecc. In tutti i questi casi, si intuisce come l'attribuzione delle coordinate geografiche relative agli oggetti rinvenuti non risulta un'operazione né semplice né intuitiva. E' possibile risolvere il problema adottando, in alternativa, due soluzioni che empiricamente portano al medesimo risultato.

Una prima soluzione consiste nel considerare per l'oggetto geografico esteso il suo *baricentro* definito dalle sue coordinate. In altre parole, la latitudine e la longitudine del baricentro saranno rappresentative dell'intero oggetto geografico esteso.

In alternativa, per considerare correttamente gli oggetti estesi, è possibile dare rilevanza, all'interno degli stessi, alle coordinate di elementi *particolari* che, essendo caratterizzanti dell'oggetto stesso, sono indicati per rappresentarlo. Per esempio nel caso di nomi di nazioni è corretto considerare le coordinate di città particolarmente importanti dal punto di vista geografico, ad esempio le capitali, oppure dare rilevanza a particolari città perché centrali nella biografia dell'autore.

Nel caso di nomi di corsi d'acqua rintracciati nel testo, è possibile «*eleggere*» come riferimento rappresentativo il ponte principale su quel fiume, identificato dalle sue coordinate. Infine, nel caso di nomi riferiti a catene montuose, è corretto considerare le coordinate dei valichi principali, ecc.

In tutti questi casi, la latitudine e la longitudine dell'elemento stabilito come il più rappresentativo, sarà il riferimento *assoluto* per indicare l'intero oggetto geografico esteso.

All'interno di queste casistiche, il raggio di percezione che verrà calcolato, può essere considerato come la reale *scala del sistema in esame* e risulterà in stretta correlazione non tanto con la media della distribuzione  $w(\mathbf{n})$ , quanto piuttosto con la sua *varianza*. Le ragioni di questa affermazione si devono ricercare nella definizione di *varianza*<sup>36</sup>, intesa come la variabile aleatoria  $X$  (e della distribuzione di probabilità che questa segue) che è definita da un numero indicato con  $\mathbf{Var}(X)$  e che fornisce una misura di quanto siano *vari* i valori assunti dalla variabile stessa; in altre parole, ci fornisce la misura di quanto essi si discostino dalla media  $E[X]$ .

## 4.6 L'ellisse della percezione

Nel secondo caso di studio: Il Milione di Marco Polo i concetti di centro e di raggio della percezione sono sostituiti dalla nozione più sofisticata e complessa di ellisse della percezione che costituisce l'aspetto *originale* di questa tesi.

Il termine "ellisse" deriva dal greco «ἔλλειψις *elleipsis*, nel senso di *manca*, poichè l'ellisse è un *cerchio* imperfetto»<sup>37</sup>. L'interessante concetto di ellisse della percezione viene impiegato nell'analisi di testi che presentano caratteristiche peculiari tali che il concetto di centro e di raggio della percezione non risultano indicativi per rappresentarlo fedelmente. Per realizzare l'analisi quantitativa di testi con caratterizzazioni particolari che, per esempio, narrano di itinerari di viaggio verso territori molto distanti dal punto di partenza, il centro e il raggio della percezione non sono indicati, perché poco rappresentativi, a definire lo scenario cognitivo dell'opera. E' per questa ragione che essi devono necessariamente subire l'evoluzione che è rappresentata dal calcolo dell'ellisse della percezione. Vedremo l'applicazione pratica di questo concetto nell'analisi su Il Milione di Marco Polo, il secondo caso di studio.

---

<sup>36</sup> <http://it.wikipedia.org/wiki/Varianza>

<sup>37</sup> <http://it.wikipedia.org/wiki/Ellisse>

Nell'impiego degli algoritmi per il calcolo del centro della percezione e del raggio dell'area cognitiva dell'opera di Boccaccio, infatti, vedremo come essi evidenzieranno risultanze non indicate a rappresentare, nel contempo, un orizzonte della cognizione così singolare come quello dell'opera di Marco Polo. Per questa ragione verrà impiegata la versione *evoluta* di tali algoritmi che andrà a calcolare il parametro dell'ellisse della percezione.

L'ellisse della cognizione, nell'ambito dell'analisi numerica dei testi letterari, si fonda sul presupposto che, all'interno dell'area cognitiva compresa nel corpus, esistono *direzioni preferenziali di osservazione* dell'autore che presentano caratteri di *anisotropia*<sup>38</sup>, definendo tale concetto fisico come quella proprietà per cui, «il valore d'una grandezza (velocità di accrescimento, ...) dipende dalla direzione che si considera» (Treccani voce *Anisotropia*).

Nel dettaglio, l'algoritmo avanzato per la definizione delle direzioni preferenziali, all'interno dell'ellisse della percezione, fornisce, a fronte dei parametri di input che abbiamo dettagliato al par. 4.4, i seguenti valori di output:

- △ il quadrato dell'asse maggiore identificata con **a** e dell'asse minore **b**, tra loro ortogonali (perpendicolari) e che insieme formano l'ellisse della percezione;
- △ il valore dell'angolo  $\Phi'$  che determina l'orientamento dell'ellisse della percezione e, quindi, dei suoi assi con l'asse polare.

Si chiamano ortogonali o, in modo equivalente perpendicolari, quelle rette che, incontrandosi in un punto, formano quattro angoli uguali di  $90^\circ$  e che si dicono retti)<sup>39</sup>.

L'asse *maggiore* **a** rappresenta la direzione *più lontana* mentre l'asse minore **b** costituisce quella più vicina verso le quali l'autore ha diretto il suo sguardo al momento della stesura dell'opera.

---

<sup>38</sup> <http://it.wikipedia.org/wiki/Anisotropia>

<sup>39</sup> <http://it.wikipedia.org/wiki/Ortogonalità>

Nell'analisi di testi con caratteristiche peculiari, come quelle che riscontreremo nel secondo caso di studio, i concetti di assi principali e di ellisse della percezione sono computati attraverso il riferimento a un differente sistema di *coordinate sferiche* (dette coordinate polari) tali che il centro di percezione, per il nuovo sistema, risulti incentrato nel Polo Nord.

Otterremo, con l'applicazione delle seguenti formule, un nuovo sistema di coordinate sferiche  $\Theta'_n$  e  $\varphi'_n$  (che rappresentano rispettivamente la latitudine e la longitudine):

$$\Theta'_n = d_n / R_E$$

$$\varphi'_n = \arctan \left( \frac{\cos \theta_n \sin \varphi_n - \Phi}{\cos \theta_n \sin \theta \cos (\varphi_n - \Phi) - \sin \theta_n \cos \theta} \right)$$

Si introducono, quindi, i principi scientifici di *sistema di coordinate polari planari*, le cui coordinate (latitudine e longitudine) vengono rappresentate con i simboli  $\Theta'_n$  e  $\varphi'_n$ , e di *tensori*<sup>40</sup> a supporto della comprensione delle fasi che caratterizzeranno la definizione dell'algoritmo complesso per il calcolo dell'ellisse della percezione che adopereremo nell'analisi del testo de Il Milione.

In matematica, si definisce un *sistema di coordinate polari* come quella configurazione di coordinate bidimensionale nella quale ogni punto del piano è identificato da un *angolo* e da una *distanza* da un punto fisso detto *polo*. Il sistema di coordinate bidimensionali risulta in corrispondenza biunivoca con quello di coordinate cartesiane; infatti si sostiene che a un vettore di coordinate cartesiane ne corrisponda uno e uno solo in coordinate polari. Il sistema di coordinate polari può essere esteso anche alle *tre dimensioni*, utilizzando due differenti sistemi di coordinate, quelle *cilindriche* e quelle *sferiche*; entrambi, comunque, richiedono coordinate polari planari o bidimensionali come base.

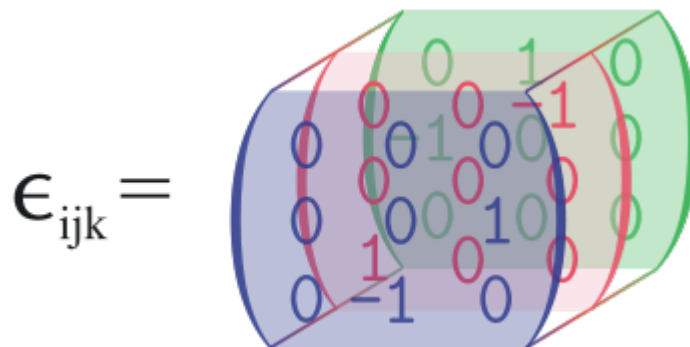
---

<sup>40</sup> <http://it.wikipedia.org/wiki/Tensore>



In altre parole, il sistema cilindrico<sup>41</sup> è la *naturale* espansione del sistema polare nelle tre dimensioni. In questo caso le coordinate sono  $\rho$ ,  $\varphi$  e  $z$ ; il sistema sferico<sup>42</sup> è un ulteriore assetto di riferimento che può essere utilizzato per orientarsi nello spazio. Esso è formato da tre coordinate:  $\rho$ ,  $\theta$  e  $\varphi$ . Nel sistema sferico si considera sempre un generico punto  $P$  e la sua proiezione sul piano  $XY$  chiamata  $Q$ .

In matematica, la nozione di *tenore* generalizza tutte le strutture definite usualmente in algebra lineare a partire da un singolo spazio vettoriale. Nell'ambito delle scienze fisiche, un tenore può essere descritto informalmente come una “*matrice a più dimensioni*” contenente valori arbitrari. Nella fig.7 è mostrata una matrice tridimensionale (un cubo) contenente  $3^3 = 27$  numeri che rappresenta il «tenore di Levi-Civita». Similmente a quanto accade per la matrice associata a una applicazione lineare, una descrizione di questo tipo dipende però fortemente dalla scelta di un sistema di riferimento, ovvero di una base.



**Figura 7:** La rappresentazione grafica di un tenore

<sup>41</sup> [http://it.wikipedia.org/wiki/Sistema\\_di\\_riferimento#Il\\_sistema\\_cilindrico](http://it.wikipedia.org/wiki/Sistema_di_riferimento#Il_sistema_cilindrico)

<sup>42</sup> [http://it.wikipedia.org/wiki/Sistema\\_di\\_riferimento#Il\\_sistema\\_sferico](http://it.wikipedia.org/wiki/Sistema_di_riferimento#Il_sistema_sferico)

Tenendo presente i concetti appena esposti, è necessario, nella applicazione all'analisi dei testi, sia reinterpretare i parametri  $\mathbf{d}(\mathbf{n})$  e  $\varphi'_n$  in funzione delle *coordinate polari planari*<sup>43</sup>, che costruire un *tensore*<sup>44</sup> simmetrico i cui componenti sono:

$$\mathbf{I}_{11} \equiv \mathbf{I}_+ + \mathbf{I}_-$$

$$\mathbf{I}_{22} \equiv \mathbf{I}_+ + \mathbf{I}_-$$

$$\mathbf{I}_{12} = \mathbf{I}_{21}$$

Tenendo presente la notazione che esprime come  $\mathbf{I}_+$  e  $\mathbf{I}_{12}$  non subiscono variazioni dovute alle rotazione, si definiscono le seguenti componenti:

$$\mathbf{I}_+ = \frac{1}{2} \sum \mathbf{w}(\mathbf{n}) \mathbf{d}^2(\mathbf{n})$$

$$\mathbf{I}_- = \frac{1}{2} \sum \mathbf{w}(\mathbf{n}) \mathbf{d}^2(\mathbf{n}) \cos 2\varphi'_n$$

$$\mathbf{I}_{12} = \frac{1}{2} \sum \mathbf{w}(\mathbf{n}) \mathbf{d}^2(\mathbf{n}) \sin 2\varphi'_n$$

Nello specifico, anticipiamo ciò che sarà una parte della risultanza dell'analisi sul testo di studio: l'area cognitiva che emergerà dal testo di Marco Polo presenterà una forma di un'ellisse orientata di un angolo  $\Phi'$  e che sarà costituita dai due assi principali  $\mathbf{a}$  e  $\mathbf{b}$ , uno maggiore rispetto all'altro e ortogonali fra loro. L'ellisse della percezione, che si formerà con gli assi, verrà compresa in un nuovo sistema di coordinate polari sarà costituita dalle due direzioni preferenziali proprie dell'autore ed espresse nel Milione. Tale ellisse della percezione avrà un'inclinazione, rispetto all'asse polare, definita dall'angolo  $\Phi'$  e identificato dalla relazione:

$$\Phi' = \frac{1}{2} \arctan \mathbf{I}_{12} / \mathbf{I}_-$$

<sup>43</sup> [http://it.wikipedia.org/wiki/Sistema\\_di\\_coordinate\\_polari](http://it.wikipedia.org/wiki/Sistema_di_coordinate_polari)

<sup>44</sup> <http://it.wikipedia.org/wiki/Tensore>

Le lunghezze degli assi principali **a** e **b** che hanno la caratteristica di non subire variazioni dovute alla rotazione, vengono definiti rispettivamente dalle formule:

$$\mathbf{a}^2 = \mathbf{I}_+ + \sqrt{\mathbf{I}_- + \mathbf{I}_{12}^2}$$

$$\mathbf{a} = \sqrt{\mathbf{a}^2}$$

$$\mathbf{b}^2 = \sqrt{\mathbf{I}_+ - \mathbf{I}_{12}^2}$$

$$\mathbf{b} = \sqrt{\mathbf{b}^2}$$

Il concetto di ellisse della percezione offre significative informazioni che approfondiscono l'analisi dei caratteri della geografia percettiva dei testi letterari. Essa, infatti, rivela le direzioni privilegiate dell'autore e i suoi orientamenti cognitivi, definiti sia in termine di "lunghezza" che di "larghezza" ed espressi al momento della stesura della sua narrazione. Il risultato che si otterrà, in termini informativi, sarà tanto più apprezzabile quando le misure relative alle lunghezze dei due assi ortogonali **a** e **b** presenteranno considerevoli differenze.

## 5 Il primo caso di studio: Il Decameron

### 5.1 Giovanni Boccaccio<sup>i</sup>

Nato a Certaldo (o a Firenze) nel giugno/luglio 1313 e deceduto a Certaldo il 21 dicembre 1375, Giovanni Boccaccio è ritenuto dalla critica e da Vittore Branca<sup>45</sup> in particolare, uno fra i maggiori narratori italiani ed europei del XIV secolo. La nota peculiare di Giovanni Boccaccio, che si riscopre soprattutto nella sua opera più famosa: Decameron, è quella della logica del mondo borghese-mercantile<sup>46</sup> e dell'utile che diventa, per il poeta, il vero elemento costituzionale. Se ne appura la ragione analizzandone la vita: il padre e gli zii avevano per decenni seguito il traffico mercantile europeo a Firenze, Napoli, Parigi, come agenti della compagnia dei Bardi alleata con i potenti Peruzzi e Acciaiuoli. A Napoli Boccaccio per sei anni fece pratica di banco (ricevette clienti, tenne i libri «della ragione», «dell'asse», «della cassa», «delle tratte», di «compere e vendite», di revisione e di «saldamento di ragione») come «discepolo», acquistando una profonda esperienza dei punti di vendita d'Italia e dell'Europa anche attraverso il suo contatto con i corrieri del tempo. Dai protagonisti dei commerci, egli, ancora giovane, imparò a conoscere il dominio del denaro, la spietatezza della convenienza economica chiusa al sentimento, le leggi ferree della ragione di mercatura, i rapporti dei mercanti con gli usurai, con i banchieri, con i politici e con i regnanti. Conobbe, inoltre, lo stato della mercatura nelle Fiandre, in Provenza, in Inghilterra, i rischi, i pericoli e le avventure narrate da naviganti e viaggiatori.

La ricca vita interiore e la fantasia del poeta erano proiettate, grazie a questa sua esperienza di vita, verso un mondo esterno la cui matrice era fornita, proprio, dalle esperienze della vita commerciale. In realtà, il giovane Boccaccio non amava dedicarsi alla vita di mercante né agli studi giuridici impostigli dal padre («tanto l'animo tutto era preso dall'affetto agli studi poetici» il poeta scriverà di sé stesso).

---

<sup>45</sup> (Savona, 9 luglio 1913 – Venezia, 28 maggio 2004) è stato un filologo e critico letterario italiano

<sup>46</sup> <http://www.storiadellaletteratura.it/main.php?cap=4&par=5#it>

Agli studi letterari, in realtà, si avvicinò anni dopo, quando poté seguire la sua vera inclinazione frequentando i giovani dell'aristocrazia napoletana che avevano accesso alla corte di Roberto d'Angiò. Nella capitale di un regno, presso la corte cosmopolita del dotto sovrano, fu avviato allo studio dell'antichità classica (il suo poeta preferito era Ovidio) ma, soprattutto, dei poemi cortesi medievali e alle esperienze di amore passionale di una società gaia e spregiudicata. Per questa libertà insita nella vita aristocratica, Napoli venne definita dal poeta «*lieta, pacifica, abbondevole, magnifica*» e rimase per lui un luogo di fascino e di splendore. I suoi rapporti con l'ambiente aristocratico gli determinarono una vera ammirazione per il mondo feudale cortese e una visione gerarchica della vita che egli proiettò nei confronti della società borghese-mercantile in decadenza. Nel 1340, in seguito al fallimento della banca dei Bardi, dovette ritornare, per gestire gli affari con il padre, nella Firenze che poi definirà «triste, grigia e noiosa». I suoi concittadini, ai quali rivolse parole quali «quella gente superba e avara che “bada solo a se stessa”», negli anni seguenti gli affidarono importanti incarichi e ambascerie.

L'intermezzo forlivese, così verrà chiamato il periodo dal 1347 al 1348, gli diede l'occasione di seguire le tracce di Dante e di essere ospite di Francesco Ordelaffi il Grande<sup>47</sup>, nonché, di frequentare poeti illustri del tempo come Nereo Morandi<sup>48</sup> e Francesco Miletto de Rossi<sup>49</sup>, detto Checco, col quale mantenne una lunga e amichevole corrispondenza.

Tra i testi di questo periodo vi fu l'egloga *Faunus*, in cui Boccaccio ricordò il passaggio a Forlì di Luigi d'Ungheria (*Titiro*, nell'egloga) e a cui si unì Francesco Ordelaffi (*Fauno*, appunto), diretto verso Napoli. Questo componimento venne incluso dallo stesso poeta nella raccolta *Bucolicum Carmen* del 1349-1367. Nel 1351, fu invitato dai fiorentini a Padova per offrire al Petrarca una cattedra (che quest'ultimo non accettò) nello Studio fiorentino.

---

<sup>47</sup> (1310? – 1373) , signore di Forlì nel XIV secolo, della nobile famiglia degli Ordelaffi [http://it.wikipedia.org/wiki/Francesco\\_II\\_Ordelaffi](http://it.wikipedia.org/wiki/Francesco_II_Ordelaffi)

<sup>48</sup> (Forlì, ... – ...) è stato un poeta italiano del XIV secolo, amico di Francesco Petrarca che lo rammenta nelle epistole. Operò alla corte di Francesco II Ordelaffi insieme al concittadino Checco Miletto de Rossi, a sua volta in rapporto di amicizia con il Petrarca e il Boccaccio [http://it.wikipedia.org/wiki/Nereo\\_Morandi](http://it.wikipedia.org/wiki/Nereo_Morandi)

<sup>49</sup> (Forlì, 1320 – 1363), è stato un poeta italiano

Del Petrarca rimase, comunque, amico affettuoso subendone l'influenza negli ultimi anni in cui si dedicò agli studi umanistici e alla lettura della *Commedia* che eseguì nella chiesa di S. Stefano di Badia. Il suo ritorno a Firenze fu la fine della vita amorosa e cortese, tanto che diede voce al rimpianto di quel periodo nel *Ninfale d'Amleto* o *Commedia delle Ninfe fiorentine* e dove manifestò un confronto dai toni duri e impietosi tra l'ambiente di corte e quello domestico borghese.

## 5.2 Decameron

«con le sue storie e i suoi poliedrici personaggi, non è semplicemente una raccolta di racconti uno diverso dall'altro. Nel suo insieme è il luogo in cui si muove un intero cosmo, descritto dal Boccaccio attraverso brevi storie, fruibile dal lettore attraverso una “navigazione” tutta personale. Non a caso è stato oggetto di uno dei primi esperimenti di ipertesto letterario sul Web »<sup>50</sup>

«(...) Comincia il libro chiamato Decameron, cognominato Prencipe Galeotto, nel quale si contengono cento novelle in dieci dì dette da sette donne e da tre giovani uomini (...)» G. Boccaccio, *Decamerone*, incipit proemio)



Figura 8: John William Waterhouse, *The Decameron*, 1916 (Liverpool, Lady Lever Art Gallery)

Il *Decamerone* o *Decameron*<sup>ii</sup> (dal greco antico, δέκα, *déka* “dieci”, e ἡμέρα, *hēméra* “giorno”: *opera di dieci giorni*) è una raccolta di cento novelle scritta da Giovanni Boccaccio, probabilmente tra il 1349 ed il 1351, per celebrare il *mondo borghese*.

<sup>50</sup> <http://www.controversi.org/decamerone.pdf>

Quella borghesia mercantile<sup>51</sup> che non era riuscita a creare una realtà politica unitaria ma che è ideologicamente la protagonista del *Decameron*. Lo è come cetto economico ma non politico perché, quando Boccaccio scrisse il suo capolavoro, essa si avviava verso la decadenza. Sono ormai lontani i tempi in cui Bindo Bonichi<sup>52</sup> scriveva «Tutti sem d'una massa - e l'uno all'altro eguale, - parlando generale», per esaltare l'affermazione sociale dell'uomo attraverso le capacità personali.

Tuttavia, i valori dell'attività mercantile furono pienamente assorbiti dal poeta come realtà e necessità della vita. La logica del mondo borghese e l'utile divennero un elemento fondamentale della struttura narrativa della sua produzione matura. Infatti, il lirismo autobiografico giovanile è scomparso e Boccaccio rappresenta artisticamente, in modo oggettivo, il *suo* mondo perché ha preso coscienza di alcune costanti della realtà. Le leggi del guadagno sono inesorabili, passano su tutto e a esse soggiacciono i sentimenti. Al di fuori del misticismo, il poeta riconobbe che l'uomo non può andare contro le leggi della natura e si prende gioco di coloro i quali, ipocritamente, fingono la costanza della virtù. I luoghi di osservazione della vita, in verità, erano stati per Boccaccio i porti di mare, i fondachi di commercio dove convenivano gruppi di persone esperte della vita di diversi Paesi e ambiti. La cornice del *Decameron* è quella della peste a Firenze del 1348, durante la quale sette giovani donne e tre giovani uomini si incontrano nella chiesa di S. Maria Novella e decidono, per evitare il contagio, di ritirarsi in una villa di Fiesole. Qui trascorrono delle settimane raccontando a turno delle novelle, una ciascuno, in dieci giorni. Alla fine di ogni giornata, la radunanza canta una ballata che, talvolta, accompagna con il suono di strumenti e con la danza. Nella prima e nona giornata ognuno è libero di narrare ciò che gli è più gradito, nelle altre, il *re* o la *regina* che presiede, propone un tema con esito a lieto fine oltre la speranza, amanti infelici, motti leggiadri e pronte risposte, azioni liberali, burle di buontemponi e inganni di ciurmatori ecc.

Le fonti consultate dall'autore, per la stesura delle novelle, sono comprese nei *fabliaux* francesi, i racconti orientali, le raccolte novellistiche italiane ed europee, nonché la loro tradizione orale che segue le vie dei mercanti.

---

<sup>51</sup> <http://www.storiadellaletteratura.it/main.php?cap=4&par=6#it>

<sup>52</sup> (Siena, 1260 – Siena, 3 gennaio 1337) è stato un religioso e poeta italiano.



Alcune novelle hanno, in particolare, origine in terre lontane come l'Indi, la Persia e la Spagna, altre sono vecchie di secoli, come quella tratta dal racconto di Andreuccio da Perugia (II,5) nato a Efeso nel secondo secolo. La struttura narrativa del telaio (ma non i personaggi o la trama) nasce dalla *Pachatantra*<sup>53</sup>, scritta in sanscrito prima del 500 d.C. e appresa da Boccaccio attraverso una catena di traduzioni che comprendeva l'antico persiano, l'arabo, l'ebraico e il latino. Anche la descrizione della peste nera, testimoniata puntualmente dall'autore nell'opera, è frutto dell'intreccio delle sue conoscenze personali con l'*Historia Langobardorum*<sup>iii</sup> di Paolo Diacono<sup>54</sup>, vissuto nell'ottavo secolo.

Se il mondo della borghesia, protesa attivamente verso i guadagni e aperta a modi di comportamento non contaminati dall'ipocrisia, rappresenta le strutture ideologiche del *Decameron*, l'elemento cortese è vivamente presente nell'opera e ne costituisce l'atmosfera di generosità umana. Anche questo aspetto ritrae l'esperienza di vita di Boccaccio che si era svolta, come si è detto, tra l'aristocrazia della borghesia, dotata di mezzi e di stampo liberale, e la società feudale della corte napoletana; per questa ragione gli ideali del poeta sembrano sempre essere accampati in queste due sfere. L'atmosfera di nobiltà morale che da esse promana, ha sotto di sé la varietà della commedia umana in cui si esercitano le forze degli individui.

Il mondo ideale di Boccaccio è una società staticamente divisa in classi: la borghesia antif feudale fiorentina è contraria ad avvicinare le classi inferiori che premono per ottenere la partecipazione politica. L'idealizzazione stessa della società cortese, inoltre, porta il poeta alla contemplazione della realtà e a un arretramento politico, nel senso di un più debole inserimento nel processo storico. In sostanza, il giudizio boccacciano sul mondo è quello del ceto, della radunanza, del gruppo, al quale egli appartiene o crede di appartenere (quando egli è a Napoli «intra nobili giovani meco in età convenienti» osserva che costoro «quantunque nobili d'entrare in casa mia né di me visitare non si vergnavano»).

---

<sup>53</sup> è una antica raccolta indiana di favole in prosa e versi

<sup>54</sup> in latino Paulus Diaconus, pseudonimo di Paul Warnefried o anche Paolo di Varnefrido è stato un monaco, storico, poeta e scrittore longobardo di espressione latina

Nelle cento novelle, inoltre, il rapporto *uomo - fortuna* ha largo campo tanto che si nota come i protagonisti mettano in opera le capacità di ingegno e di previsione dell'etica borghese, la quale, d'altra parte, nella polemica contro i privilegi feudali, sostiene l'uguaglianza fra gli uomini anche se è un'uguaglianza fra i più dotati, forniti di mezzi e di cultura e meno soggetti ai condizionamenti. La legge universale è, come abbiamo visto, il naturalismo per cui contrastare le leggi «della natura troppe gran forze bisognano, e spesse volte non solamente invano ma con grandissimo danno del faticante s'adoperano». Boccaccio, in realtà, nello scrivere il *Decameron* non si rivolge ai dotti ma a un pubblico vasto e di donne (come egli stesso disse per difendersi dalle accuse di «troppa licenzia usata» fin da quando fu pubblicata una parte dell'opera), «per cacciar la malinconia» loro e con libertà di artista. Le novelle non sono raccontate in chiesa, aggiungerà, «ma tra' giardini, in luogo di sollazzo, tra persone giovani» e non inclinabili al male; del resto, «niuna corrotta mente intese mai sanamente parola» né le parole non troppo oneste possono guastare un animo sano. L'ispirazione amorosa che soggiace al *Decameron*, dunque, non è compiacimento dell'osceno ma interesse per il modo in cui si intrecciano istinti e sentimenti. Anche nella prosa (Boccaccio, fu detto, concepisce come Plauto e scrive come Cicerone) la retorica che presiede all'organizzazione stilistica si modella elegantemente nelle scene tragiche o comiche, di beffa o di arguzia, ed è consona agli atteggiamenti interiori della brigata novellatrice. E' chiaro quindi, il continuo e puntuale riferimento dell'autore alla Firenze contaminata dalla peste, che si contrappone alla vita *regolata* della brigata.

Proprio per sottolineare le due realtà contrapposte il poeta, all'inizio dell'opera, ci presenta una lunga e dettagliata descrizione della calamità che colpì Firenze nel 1348 e che, oltre a decimare la popolazione, distrusse tutte quelle norme sociali, quegli usi e quei costumi che tanto gli erano a cari. In parallelo alla sciagura della peste, il gruppo, la brigata dei giovani, creatori di una sorta di realtà quasi perfetta; essi dimostrano come l'uomo, grazie alle proprie forze e alla propria intelligenza, sia in grado di contrastare gli eventi della vita e di dare un ordine alle cose.

Principio questo, ribadito spesso nel Decameron e che costituirà uno dei temi fondamentali dell'Umanesimo: la capacità dell'uomo di cambiare il corso degli eventi, imponendosi con volontà e determinazione. Una facoltà insita nell'umanità chiamata *industria* che Boccaccio associa sempre alla classe emergente (mercanti e nuovi borghesi), della quale sente di far parte. Tutto ciò, quindi, fa del Decameron, un'opera unica, non una semplice raccolta di novelle ma una storia variegata unita da un filo conduttore che, attraverso la cornice narrativa, forma un vero e proprio romanzo.

Per concludere, nelle novelle racchiuse nel Decameron, si ritrova non solo il puro intrattenimento, spesso umoristico e con numerosi richiami all'erotismo bucolico del tempo, ma anche un importante messaggio umanistico «*Tributo all'umano ingegno, epico capolavoro di una società mercantile in dinamica ascesa, la cui ricerca del piacere è minacciata dal rischio di una subitanea estinzione, il Decameron può essere letto tanto come un manuale di comportamento trasgressivo quanto come un breviario di mondana sopravvivenza, inteso per un lettore laico e spregiudicato*», così, è come dell'opera ha scritto Guido Almansi<sup>55</sup>.

Il *Decameron* è quindi un testo all'apparenza scandalosamente amorale ma che, al contrario, può prestarsi a una lettura e a una meditazione etica. E proprio per questi connotati che *le cento novelle* hanno dato vita nei secoli al dibattito sulla natura dell'arte narrativa, sulla morale medievale confrontata con quella moderna nonché, sul valore sociale ed educativo di qualsiasi forma di espressione artistica e letteraria. Una sorta, di enciclopedia della vita, summa della cultura tardo medievale e cittadina, il Decameron contiene anche un repertorio universale di situazioni e dilemmi squisitamente umani.

---

<sup>55</sup> critico letterario, scrittore e traduttore italiano naturalizzato britannico ([http://it.wikipedia.org/wiki/Guido\\_Almansi](http://it.wikipedia.org/wiki/Guido_Almansi))

### 5.2.1 Lo scenario cognitivo

Il Decameron di Giovanni Boccaccio costituisce il primo caso di studio che affrontiamo adoperando l'*analisi quantitativa* del testo e l'impiego di formalismi delle scienze fisiche, al fine di determinare i parametri necessari per il calcolo dello scenario cognitivo dell'opera. Più precisamente, si determinano i parametri indicativi del *centro della percezione* che coincide con il luogo fisico e/o mentale nel quale Boccaccio, molto probabilmente, si trovava al momento della stesura della raccolta delle novelle, e del *raggio della percezione* che rappresenta la distanza massima oltre la quale gli eventi, per l'autore, assumono un significato marginale o inesistente.

L'identificazione dei valori numerici, relativi al centro e al raggio della percezione, ci offre l'interessante opportunità di trasformare le informazioni di carattere *qualitativo*, delle quali abbiamo parlato al par. 4.1 e che «naturalmente» sono insite nei testi perché strettamente legate alla loro sintassi, in informazioni *quantitative*. Il centro e il raggio della percezione che risulteranno, dall'impiego degli algoritmi formali definiscono in termini scientifici la *percezione spaziale* dell'autore; in altre parole, la geografia cognitiva espressa nella sua opera.

Per la realizzazione dell'analisi del primo caso di studio le fonti, consultate nella loro versione *cartacea*, sono rappresentate dall'opera *Decameron* di Giovanni Boccaccio (sesta edizione riveduta, corretta e aggiornata nelle bibliografie al 1991 di Giulio Einaudi editore) nell'edizione curata da Vittore Branca<sup>56</sup>, uno tra i più accreditati studiosi contemporanei del Boccaccio, filologo e critico letterario italiano nonché emerito professore di Letteratura italiana presso l'Università degli Studi di Padova deceduto nel 2004. Inoltre, è stata consultata una fonte digitale rappresentata dal *testo originale* del *Decameron* nella corrente edizione critica, a cura di Vittore Branca, digitata da Nanda Cremascoli e dai suoi studenti, e resa disponibile in rete da *Liber Liber*<sup>57</sup>.

---

<sup>56</sup> [http://it.wikipedia.org/wiki/Vittore\\_Branca](http://it.wikipedia.org/wiki/Vittore_Branca)

<sup>57</sup> <http://www.liberliber.it/biblioteca/b/boccaccio/>

Quest'ultimo testo è comprensivo delle aggiunte riguardanti la prima traduzione dall'inglese (XVII secolo) e consultabile in rete a opera dell'Oxford Text Archive e codificata in XML<sup>58</sup>.

Il ricorso in parallelo, sia al supporto cartaceo e *tradizionale* dell'opera, che all'uso *creativo della tecnologia* in ambito letterario, costituito dai testi digitali codificati con il linguaggio di marcatura xml, ha reso possibile, a nostro avviso, un lavoro di analisi più accurato e rigoroso, in quanto ci ha fornito la possibilità di effettuare un continuo riscontro sulle rilevazioni effettuate dalla consultazione della fonte cartacea.

L'elenco delle fasi operative che hanno contraddistinto l'analisi quantitativa del testo del Decameron di Giovanni Boccaccio vengono così dettagliate:

- *Il primo stadio* è stato rappresentato dall'individuazione, all'interno del corpus del *Decameron*, di tutte le *indicazioni geografiche*, o più propriamente, degli *oggetti geografici* menzionati dall'autore e dal relativo conteggio delle *occorrenze*. L'annotazione e il conteggio ha compreso il totale delle indicazioni spaziali e geografiche del testo sia di tipo *esplicito* che *implicito*. Esse sono costituite sia dal nome dei luoghi (monasteri, chiese, piazze, ecc.), delle regioni, dei monti, dei fiumi e degli edifici localizzabili in un determinato territorio ecc., e, quindi, facilmente rintracciabili nel testo, oppure sottintesi, in modo implicito, tramite l'uso di perifrasi<sup>59</sup> e di allusioni (citazioni come «il Santo Sepolcro» per riferirsi alla città di Gerusalemme, oppure «la nostra città» per sottintendere Firenze). Si intuisce come, per questi frequenti casi, l'operazione di rilevazione degli oggetti geografici e il relativo conteggio delle occorrenze abbia rappresentato una problematica la cui risoluzione abbia richiesto particolare attenzione;

---

<sup>58</sup> [http://www.brown.edu/Departments/Italian\\_Studies/dweb/texts/DecIndex.php?lang=it](http://www.brown.edu/Departments/Italian_Studies/dweb/texts/DecIndex.php?lang=it)

<sup>59</sup> <http://www.wordreference.com/definizione/perifrasi>

- si è passati al *geocoding*<sup>60</sup> o *georeferenziazione*<sup>61</sup>. Questa fase ha previsto l'attribuzione dell'informazione relativa alla precisa posizione dell'oggetto geografico nel sistema di coordinate decimali. La latitudine e la longitudine per ogni oggetti geografico, espresse in gradi decimali (si sono considerate 4 cifre decimali), sono state *tutte* rintracciate, per esigenze di *uniformità*, con la consultazione delle pagine web del portale *geohack*<sup>62</sup>. In altre parole, l'informazione utile, in questa fase del lavoro, è stata quella espressa nel particolare *sistema geodetico*<sup>63</sup> di riferimento (longitudine e latitudine) al fine di descrivere la posizione dell'oggetto geografico sulla superficie terrestre;
- le informazioni ottenute dalla prima fase hanno costituito i *dati di input* e sono state implementate in un foglio elettronico finalizzato alla descrizione del modello matematico. E' stata indicato, in due colonne differenti (colonna A e colonna E) sia il toponimo antico che quello moderno rilevato, molto spesso, dalla consultazione delle note del testo oppure dei siti web dei Comuni che riportano una parte dedicata alla toponomastica.

Il modello matematico riassunto nel foglio elettronico è stato implementato sulla base degli algoritmi formali ispirati ai principi delle scienze fisiche.

I *dati di input* inseriti sono i seguenti:

- gli oggetti geografici identificati con il toponimo antico;
- il corrispondente toponimo moderno;
- la latitudine e la longitudine espresse in gradi decimali;

I *dati di output* sono rappresentati dai valori attesi utili per l'analisi quantitativa del corpus. Per il Decameron, in particolare, sono definiti dalle coordinate pesate e in quelle decimali del *centro della percezione*, inoltre, nell'estensione espressa in chilometri sia del *raggio della percezione* che della *distanza euclidea*.

<sup>60</sup> <http://en.wikipedia.org/wiki/Geocoding>

<sup>61</sup> <http://it.wikipedia.org/wiki/Georeferenziazione>

<sup>62</sup> [http://onestepahead.de/index.php?option=com\\_content&view=article&id=151:wikipedia-geohack&catid=34:news&Itemid=14](http://onestepahead.de/index.php?option=com_content&view=article&id=151:wikipedia-geohack&catid=34:news&Itemid=14)

<sup>63</sup> [http://it.wikipedia.org/wiki/Sistema\\_geodetico\\_di\\_riferimento](http://it.wikipedia.org/wiki/Sistema_geodetico_di_riferimento)

Nel dettaglio, il foglio elettronico contiene le seguenti informazioni:

<i>Colonna A</i>	Oggetti geografici rilevati dal testo
<i>Colonna C</i>	Numero di occorrenze di ogni singolo oggetto geografico
<i>Colonna E</i>	Toponimo moderno corrispondente
<i>Colonna F</i>	Latitudine espressa in gradi decimali (DD)
<i>Colonna G</i>	Longitudine espressa in gradi decimali (DD)
<i>Colonna I</i>	Latitudine espressa in radianti
<i>Colonna J</i>	Longitudine espressa in radianti
<i>Colonna K</i>	Coseno della latitudine
<i>Colonna L</i>	Seno della latitudine
<i>Colonna M</i>	Coseno della longitudine
<i>Colonna N</i>	Seno della longitudine
<i>Colonna O</i>	Differenza con la latitudine del centro di percezione (Firenze)
<i>Colonna P</i>	Differenza con la longitudine del centro di percezione (Firenze)
<i>Colonna Q</i>	Prodotto della colonna O per la costante 111 per ottenere dai gradi di latitudine i corrispondenti km
<i>Colonna R</i>	Prodotto della colonna P per la costante 72,5 per ottenere dai gradi di longitudine i corrispondenti km
<i>Colonna S</i>	Peso dell'oggetto geografico n
<i>Colonna U</i>	Distanza euclidea di ogni oggetto geografico dal centro di percezione precedentemente ricavato (dalla radice quadrata della somma dei quadrati delle differenze delle coordinate omonime dei punti)
<i>Colonna V</i>	Calcolo della distanza $d(n)$ sulla superficie terrestre dell'ennesimo oggetto geografico dal centro di percezione
<i>Colonna W</i>	Calcolo della coordinata sferica $\varphi'_n$
<i>Colonna X</i>	Controllo congruità dei dati della colonna V (per evitare cifre negative)
<i>Colonna Y</i>	Coseno del quadrato della colonna X

<i>Colonna Z</i>	Seno del quadrato della colonna X
<i>Cella C<sub>n</sub></i>	Sommatoria della colonna C
<i>Cella F<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna F
<i>Cella G<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna G
<i>Cella I<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna I
<i>Cella J<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna J;
<i>Cella L<sub>n</sub></i>	Coordinata pesata Z
<i>Cella L<sub>n+1</sub></i>	Coordinata A;
<i>Cella L<sub>n+2</sub></i>	Latitudine del <i>centro di percezione</i>
<i>Cella M<sub>n</sub></i>	Coordinata pesata X
<i>Cella N<sub>n</sub></i>	Coordinata pesata Y
<i>Cella N<sub>n+1</sub></i>	Coordinata B
<i>Cella N<sub>n+2</sub></i>	Longitudine del <i>centro di percezione</i>
<i>Cella O<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna O
<i>Cella P<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna P
<i>Cella Q<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna Q
<i>Cella R<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna R
<i>Cella S<sub>n</sub></i>	Somma dei pesi degli n oggetti geografici = 1
<i>Cella U<sub>n</sub></i>	Distanza Euclidea
<i>Cella V<sub>n</sub></i>	Raggio di percezione



## 5.2.2 Risultanze numeriche

Nel Decameron, il numero totale degli oggetti geografici citati è risultato n. 161 per un totale di n. 2063 occorrenze. Sono stati riuniti nel medesimo oggetto geografici i toponimi nonché i nomi di quartieri, rioni ecc. che si riferiscono al medesimo oggetto geografico. Allo stesso modo, si è fatta particolare attenzione, nel caso di uso di parafrasi nel testo a identificare correttamente l'oggetto geografico corrispondente.

Per gli oggetti estesi trattati al par 4.5 si è adottata la prima soluzione che consiste nel considerare per l'oggetto il suo *baricentro* che viene definito dalle sue coordinate e che diventa il riferimento assoluto dell'intero oggetto. In altre parole, la latitudine e la longitudine del baricentro diventano gli elementi rappresentativi dell'intero oggetto geografico esteso.

La gestione delle problematiche, incontrate nell'analisi del Decameron, sono state dettagliate insieme alle soluzioni adottate al par. 5.2.4.

Il *centro di percezione* del Decameron espresso rispettivamente dalla sua latitudine e longitudine in gradi decimali e che risulta dall'applicazione dell'algoritmo formale è il seguente:

43,2884°
----------

12,5983°
----------

Esso corrisponde (con esiguo scostamento) alle *coordinate di Firenze* (latitudine 43,7687°, longitudine 11,2569°).

Un risultato, che potremmo definire *intuitivo* per il testo di Boccaccio. Infatti, già dalla lettura dei caratteri dell'opera del Decameron, che abbiamo riportato al par. 5.1, era prevedibile che, il *luogo* fisico e/o *mentale* di Boccaccio, al momento della scrittura dell'opera, fosse la città fiorentina. E' stato interessante, anche per questo particolare aspetto, verificare che la risultanza degli algoritmi formali fosse in grado di dare una inequivocabile conferma a ciò che l'indagine storica e culturale dell'opera ci aveva suggerito chiaramente. Riteniamo, pertanto che l'intento sia stato raggiunto, in quanto il risultato formale dei calcoli ha delineato le coordinate di Firenze; città alla quale Boccaccio rivolge la sua opera nell'ottica delle sue motivazioni socio-culturali e nella cornice storica del Trecento come dettagliato al par. 5.2.

Nel prospetto riportato a fig. 9 vengono elencate le prime località (in ordine decrescente di numero i occorrenze) rinvenute nel testo, il numero delle rispettive occorrenze nonché la percentuale di incidenza rispetto al totale generale delle occorrenze: (*Notiamo la rilevante presenza di Firenze (centro di percezione) dell'autore e del testo, con 254 occorrenze su un totale di 2063*)

<b>Localita'</b>	<b>O ccorrenze</b>	<b>% del Totale</b>
<i>Firenze Toscana</i>	<b>254</b>	<b>12,31</b>
<i>Palermo Sicilia</i>	<b>88</b>	<b>4,27</b>
<i>Marsiglia Francia</i>	<b>77</b>	<b>3,73</b>
<i>Stra Lombardia</i>	<b>70</b>	<b>3,39</b>
<i>Bologna Emilia Romagna</i>	<b>66</b>	<b>3,20</b>
<i>Anversa Belgio settentrionale</i>	<b>55</b>	<b>2,67</b>
<i>Saludeco Rimini Emilia Romagna</i>	<b>49</b>	<b>2,38</b>
<i>Procida Napoli Campania</i>	<b>48</b>	<b>2,33</b>
<i>Val di Magra Toscana/Liguria</i>	<b>48</b>	<b>2,33</b>
Altre Località	<b>1308</b>	<b>63,40</b>
<b><i>Totale Complessivo</i></b>	<b>2063</b>	<b>100</b>

Figura 9: Elaborazione della tabella Pivot Microsoft® Excel per le occorrenze del Decameron<sup>64</sup>

<sup>64</sup> Estratto da Foglio elettronico Decameron di Giovanni Boccaccio.xls presentato in allegato a questa relazione

Nella fig. 9 viene rappresentato il diagramma *pie chart* (diagramma a torta) statico di Microsoft® Excel che evidenzia le percentuali di incidenza delle singole occorrenze in relazione al totale registrato delle stesse:

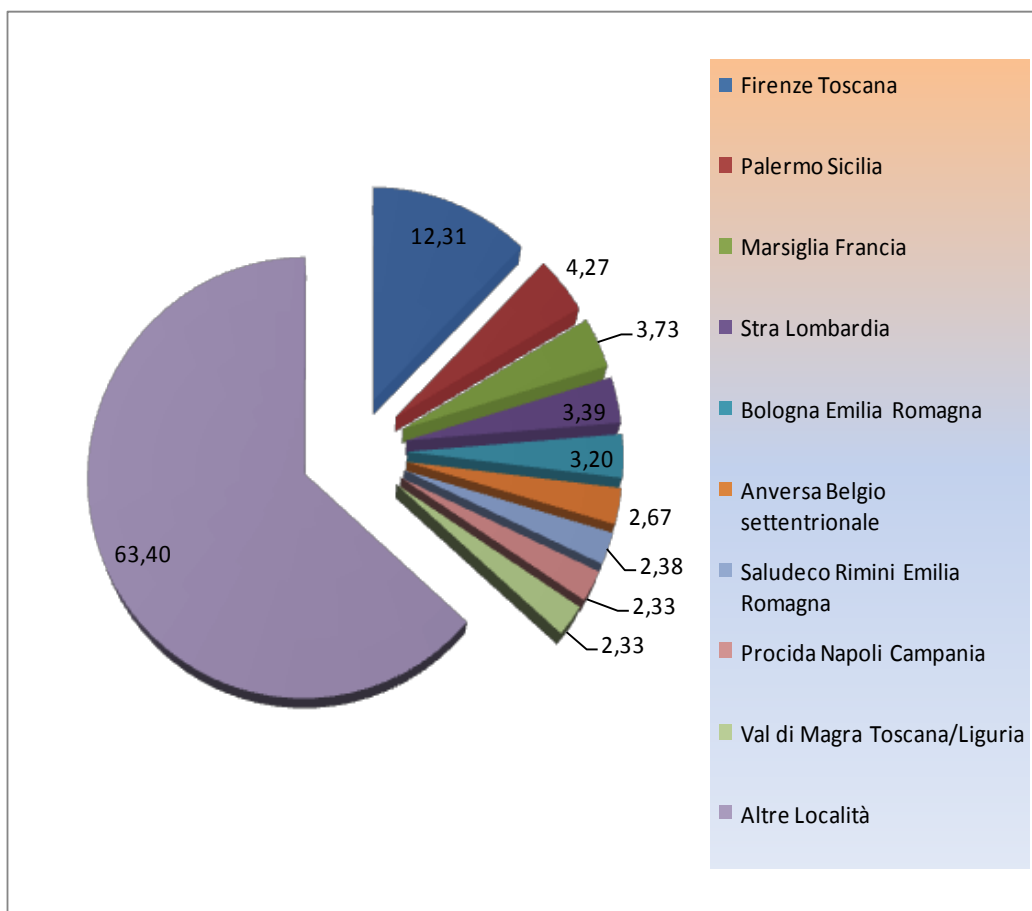


Figura 10: **Pie Chart delle risultanze precedenti**<sup>65</sup>

<sup>65</sup> Estratto da Foglio elettronico Decameron di Giovanni Boccaccio.xls presentato in allegato a questa relazione

### 5.2.3 Lo scenario culturale

Il calcolo formale del raggio di percezione del Decameron produce un risultato di 550 km (e una distanza euclidea di 508 km) che possiamo definire *molto esteso* per almeno due ordini di ragioni:

- nel Medioevo per percorrere qualche chilometro erano necessari svariati giorni; si comprenderà, di conseguenza, la singolarità del risultato conseguito;
- se facciamo un parallelo con il risultato medio del raggio della percezione di altre opere medievali, con un riscontro di 200 km per le opere storiografiche del X secolo francese degli autori Flodoard, Richer e Gerbert, e di 437 km per la *Commedia* e il *De Vulgari Eloquentia* di Dante (v.par.6.2.2), possiamo intuire l'interesse che la risultanza del caso di studio può rappresentare.

Inoltre, la caratterizzazione socio culturale dell'opera del Decameron, il particolare periodo storico e le frequentazioni di Giovanni Boccaccio durante la sua vita, si collocano perfettamente in sintonia con il risultato ottenuto dalla nostra indagine, e confermano la considerevole *ampiezza mentale e cognitiva* di Giovanni Boccaccio aiutandoci a capirne le possibili motivazioni.

Vi sono, a nostro avviso, diversi aspetti che è interessante descrivere e che si ricavano dall'analisi della biografia dell'autore, nonché dai caratteri della sua opera: *in primo luogo* riguardano i *destinatari* dell'opera. Giovanni Boccaccio, dedica il Decameron alla donne, per riscattarle dalle pene d'amore dalle quali sono tormentate e perché custodi della *cultura* dell'epoca e quindi anche in grado di comprendere il *valore* del contenuto delle novelle. Ma, soprattutto, Boccaccio, con il Decameron, si rivolge alla *borghesia commerciale* in declino che, grazie alle cento novelle, doveva ricevere un incentivo per *reagire* agli eventi sfavorevoli di quel particolare periodo storico: primo fra tutti, la peste nera che incombeva sulla Firenze dell'epoca. Ed è proprio in quest'ottica largamente progressista che si colloca un centro cognitivo nella città fiorentina, sede dei suoi interessi al momento della scrittura del Decameron, e un raggio della percezione che comprende le mete dei commerci e delle navigazioni dell'epoca;

*un secondo ordine* di motivazioni lo ritroviamo analizzando la biografica dell'autore, in particolare soffermandoci sulle *frequentazioni* del poeta durante gli anni della giovinezza. Fu intorno al 1331 che il padre del poeta, Boccaccino da Chellino, decise di mandare il giovane a Napoli per introdurlo al diritto canonico e, di conseguenza, ad un mestiere redditizio. Fu proprio in questi anni che Boccaccio ebbe l'opportunità di frequentare la Corte angioina, noto bacino dell'*intelligencijs* progressista del tempo, nonché testimone dei più variegati e suggestivi orizzonti culturali; ecco che ritroviamo in Boccaccio un'ampiezza cognitiva fuori dal comune per il suo tempo;

*un terzo aspetto* che non deve essere trascurato e che contribuisce, in modo ulteriore, a motivare un orizzonte così esteso, come quello delineato, riguarda le fonti consultate dal poeta per la stesura delle cento novelle racchiuse nel Decameron. Come abbiamo dettagliato al par. 5.2, Boccaccio per la stesura del Decameron si è ispirato alle *fabliaux* francesi, ai racconti orientali, alle raccolte novellistiche italiane ed europee, nonché alla tradizione orale che seguiva le vie dei mercanti e con i quali egli, per ragioni commerciali, era in contatto. Alcune novelle, infatti, portano i caratteri originari di terre lontane come l'Indi, la Persia e la Spagna, altre sono vecchie di secoli. La struttura narrativa del telaio (ma non i personaggi o la trama) nasce dalla *Pachatantra*<sup>66</sup>, scritta in sanscrito prima del 500 d.C. e appresa da Boccaccio attraverso una catena di traduzioni che comprendeva l'antico persiano, l'arabo, l'ebraico e il latino;

infine, *un altro aspetto* che ha condizionato un orizzonte cognitivo così geograficamente rilevante è rappresentato dalla connotazione geografica della Corte d'Angiò<sup>67</sup> con la quale era in contatto; ritroviamo, infatti, nella sua opera la quasi totalità dei domini angioini. Il *bacino geografico* della Corte era, infatti, costituito dal Lazio meridionale e reatino (Cittaducale, Leonessa, Amatrice), dall'Abruzzo, Molise, Campania, Puglia, Basilicata e dalla Calabria. Inoltre, era rappresentato dagli originari territori dell'Italia centro meridionale e molti furono i tentativi di espansione nell'Italia centro settentrionale e oltremare in Tunisia e nei Balcani; in particolare, con gli Aragonesi, Napoli divenne la *capitale de facto* della vasta Corona d'Aragona nel XIV e XV secolo.

---

<sup>66</sup> è una antica raccolta indiana di favole in prosa e versi

<sup>67</sup> [http://it.wikipedia.org/wiki/Regno\\_di\\_Napoli](http://it.wikipedia.org/wiki/Regno_di_Napoli)

I commerci che interessavano la Corte angioina riguardavano, oltre gli stati italiani ed europei del Mediterraneo, ad esempio l'Inghilterra. Le vivaci frequentazioni del poeta non possono che aver contribuito positivamente a formare in lui e, di conseguenza, nelle sue opere un ampio orizzonte cognitivo e culturale.

Per concludere, lo scenario cognitivo di Giovanni Boccaccio è anche e soprattutto uno *scenario culturale* e costituisce una tangibile testimonianza di un interesse del poeta per «il mondo» che esula dai normali confini dell'epoca medievale.

Le risultanze numeriche ottenute dall'applicazione delle metodologie fisiche sul Decameron di Giovanni Boccaccio, confermano in modo inequivocabile i caratteri che, dall'indagine socio culturale, abbiamo potuto appurare e che sono in relazione alla formazione del poeta.

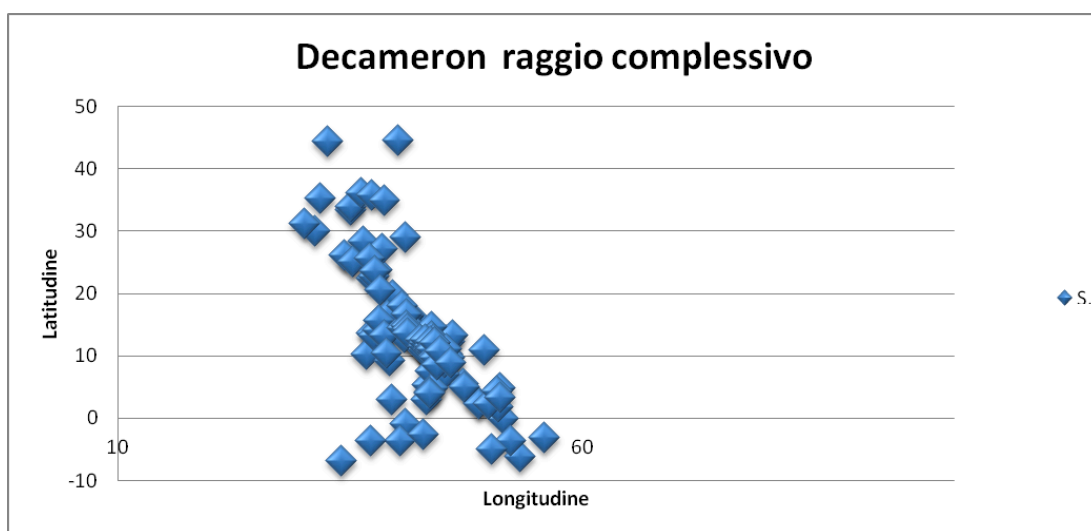


Figura 11: Scatter Plot del Decameron<sup>68</sup>

Nella fig.11 è stato elaborato un grafico Microsoft® Excel del tipo *scatter-plot*<sup>69</sup> (diagramma di dispersione) dove si raffigurano su un piano le georeferenze degli oggetti rinvenuti nel Decameron.

<sup>68</sup> Estratto da Foglio elettronico Decameron Boccaccio.xls presentato nelle Tavole di questa tesi

<sup>69</sup> Grafico di dispersione o grafico a dispersione o scatter plot o scatter graph è un tipo di grafico in cui due variabili di un set di dati sono riportate su uno spazio cartesiano.

E' possibile notare, piuttosto chiaramente, la configurazione della Toscana, dell'intera penisola italiana, con Roma, Napoli, Genova, la Sicilia e le sue coste, oltre la presenza della Francia, di Rodi, Cipro, Alessandria d'Egitto, Gerusalemme, e della Turchia ecc., le rotte dei commerci dell'epoca, nonché i luoghi geografici all'interno del raggio della percezione riscontrabili nella biografia e nelle frequentazioni socio-culturali del poeta.

Si rimanda alle tav. dalle 8 alla 12 nelle quali si presenteranno altre tipologie di visualizzazione che illustrano le coordinate geografiche degli oggetti rinvenuti nel Decameron e che delineano, sulla mappa terracquea, le entità della distribuzione rappresentate degli oggetti geografici rinvenuti nell'opera boccacciana.

#### **5.2.4 Problematiche operative**

- Un prima problematica che è stata trattata al par. 4.5 e accennata al par. 5.2.2 riguarda la presenza, nel Decameron, di riferimenti a *luoghi geograficamente molto estesi*. Per la gestione di questi particolari oggetti geografici è stata adottata la soluzione dettata dalla teoria e che consiste nel considerare il *baricentro* dell'oggetto definito dalle sue coordinate. In altre parole, la latitudine e la longitudine del baricentro sono state ritenute rappresentative dell'intero oggetto geografico esteso. In pratica, per le nazioni menzionate nel corpus, ad esempio, sono state considerate le coordinate delle capitali, mentre per le regioni sono state considerate quelle dei capoluoghi. Infine, per i corsi d'acqua e i fiumi è stato calcolato il baricentro del corso stesso (dalle coordinate dell'origine a quelle della foce) e le relative coordinate geografiche hanno assunto un valore assoluto per rappresentare l'intero oggetto;

- la questione degli oggetti geografici menzionati nel testo tramite l'uso di riferimenti *impliciti* è stata affrontata con l'interpretazione fornita dalle fonti e delle note al testo stesso. L'esempio più chiaro è rappresentato dall'attribuzione all'oggetto che rappresenta Firenze il totale delle occorrenze relative a tutte le allusioni alla città fiorentina espresse nei diversi modi, ad esempio, tramite il riferimento al «la nostra città», «Firenza», «la città dei fiorentini». Nel caso di Gerusalemme sono state sommate le occorrenze risultanti dalle espressioni riferite a «il Santo Sepolcro» o alla «Terra Santa»;
- è stato effettuato il raggruppamento nell'oggetto geografico di riferimento (es. Firenze) di tutti i nomi di quartieri, delle chiese (es. Santa Maria Novella nel caso di Firenze), dei conventi (es. San Gallo), delle piazze (es. ancora Santa Maria Novella), delle contrade (es. Garbo, Parione) nominate nel testo; analogamente si è operato nel caso di quartieri o frazioni dislocate nelle immediate vicinanze delle città principali (ad es. Settignano nel caso di Firenze) ad imputarle al conteggio per la città fiorentina;
- un'altra problematica è stata la presenza di oggetti geografici identificati con toponimi antichi e desueti, a volte di difficile riconoscimento. Uno dei numerosi esempi è costituito Ajeroli per riferirsi a Agerola in Campania. È stato utile, in questo caso, utilizzare le ricerche in Rete accedendo, spesso, ai siti web dei Comuni nei quali vi sono pagine web dedicate alla toponomastica<sup>70</sup> e all'evoluzione dei toponimi<sup>71</sup>;
- infine è stato attribuito l'oggetto geografico corrispondente ad alcuni nomi di personaggi o localizzazioni che, con l'oggetto geografico, presentavano un'attinenza geografica o contestuale. Elenchiamo per maggiore chiarezza alcuni dei casi incontrati nel testo:
  - Castello Cignano: Palermo – Italia;
  - Castello del Rossiglione e del Guardastagno: Marsiglia – Francia;
  - Currado Malaspina: Val di Magra – Italia

<sup>70</sup> <http://it.wikipedia.org/wiki/Toponomastica>

<sup>71</sup> dal greco ἔθνος *éthnos*, "tribù", + ὄνομα *ónyma*, variante dorico-eolica dell'attico ὄνομα "nome"- nome di un popolo



L'esempio fornito descrive tre castelli che sono stati attribuiti all'oggetto geografico che rappresenta la zona in cui si erigevano, in questo caso: Palermo, Marsiglia e la Val di Magra; essi, nel testo sono stati nominati utilizzando i nomi dei signori che li abitavano;

All'interno del foglio elettronico, per chiarezza, a colonna A sono state riportate le indicazioni di tutte le località menzionate nel testo e che sono state riunite nel medesimo oggetto geografico e a colonna C è stata riportata la somma delle occorrenze di ciascuna indicazione; infine a colonna E è stato riportato il toponimo moderno di tutte le località nominate dall'autore.

### **5.3 Conclusioni**

Riassumendo nel Decameron, sono stati rinvenuti nel testo n. *161* oggetti geografici per un totale di n. *2063* occorrenze. Il calcolo formale del raggio di percezione produce un risultato molto ampio (definito in km *550*) e una distanza euclidea altrettanto rilevante di *508* km. Il centro di percezione è stato localizzato (con esiguo scarto) nella città di Firenze, centro degli interessi economici e culturali di Giovanni Boccaccio al momento della stesura del Decameron. E' stato interessante appurare la rispondenza delle risultanze, ottenute dall'applicazione delle teorie fisiche, con le intuizioni che dalla lettura del testo derivano e che riguardavano la centrale importanza della città fiorentina.

E' stato, inoltre, interessante inserire le risultanze numeriche ottenute, dall'applicazione degli algoritmi formali, nell'originario contesto storico sociale e culturale, al fine di proporre una logica serie di motivazioni che potessero spiegare i parametri numerici relativi a uno scenario cognitivo e culturale sorprendentemente vasto.

## 6 Il caso di studio *originale* : Il Milione

### 6.1 Marco Polo

Marco Polo<sup>72</sup> nacque nel 1254 a Venezia da una famiglia patrizia di facoltosi mercanti, originaria di Sebenico in Dalmazia. Più o meno in quegli anni, non si sa con certezza se prima o dopo la sua nascita, il padre Niccolò e lo zio Matteo partirono per un viaggio commerciale in Oriente. Si stabilirono, dapprima, nella capitale dell'Impero latino, Costantinopoli, poi a Soldaia in Crimea, dove intorno al 1280 Marco il Vecchio (il fratello maggiore) insieme con Matteo e Niccolò, fondarono una compagnia di affari. Nel loro viaggio, i fratelli Polo si spinsero fino alla corte del grande Qubilai, il conquistatore e unificatore della Cina, il più illustre discendente del Gengis Can, e durante questo loro primo soggiorno (nel 1265) ottennero importanti privilegi e probabilmente anche la dignità nobiliare mongola. Nel 1269, quando il padre e lo zio fecero ritorno a Venezia, Marco ancora giovinetto, probabilmente nella primavera o nell'estate del 1271, decise di partire insieme con loro per la Cina, dove rimarrà per circa venticinque anni. Dopo aver lasciato, nel novembre di quello stesso anno, San Giovanni d'Acari, i Polo giunsero alla corte di Qubilai nel maggio del 1275. Qui Marco, dopo aver assolto l'incarico affidatogli dall'imperatore di ispezionare le regioni al confine del Tibet e lo Yün-nan, venne elevato alla dignità di «messere», titolo che lo legò direttamente alla figura del sovrano di cui divenne informatore e ambasciatore personale presso tutti i popoli dell'impero.

---

<sup>72</sup> [http://www.italialibri.net/autori/polom.html#Top\\_of\\_Page](http://www.italialibri.net/autori/polom.html#Top_of_Page)

E' questo il titolo con il quale Marco viene menzionato ne *Il Milione*. Durante tutta la sua permanenza presso la corte mongola, per conto del Gran Khan, Marco Polo svolse attività amministrative, delicate ambascerie e incarichi diplomatici di prestigio, compiendo a tal fine diversi viaggi. Tra i tanti riconoscimenti a lui dedicati, nel 1278, vi fu la nomina a governatore di Hang-chou, già capitale del reame dei Mangi sotto la dinastia dei Sung. Nel 1292 i Polo salparono dal porto di Zaitun e iniziarono per mare il viaggio di ritorno in patria che si concluderà nel 1295. In quello stesso anno, poco dopo, in una delle tante battaglie navali che a quel tempo colpirono il Mediterraneo orientale e i mari italiani, tra veneziani e genovesi, che si ipotizza possa essere quella di Curzola, Marco cadde prigioniero dei genovesi. E, fra il 1298 e il 1299, nelle carceri di Genova dettò al compagno di prigionia, Rustichello da Pisa<sup>73</sup>, quello che, per la letteratura italiana, fu il primo esemplare di *resoconto di viaggio: Le Divisament dou Monde* scritto, nella redazione originale, in antico francese (precisamente in lingua *d'oïl*<sup>74</sup>). Il libro fu ben presto noto con il titolo de *Il Milione*. Il nome sembra derivare dal soprannome di tutta la stirpe dei Polo e, più precisamente, per aferesi da *Emilione*, nome di un antenato della famiglia; oppure, secondo un'altra ipotesi, per il fatto che Marco Polo abbia usato, più volte, questo termine per descrivere la quantità di beni amministrata dal Kublai Khan<sup>75</sup>. Ratificata, il primo luglio 1299, la pace tra veneziani e genovesi, Marco fu di nuovo libero e fece ritorno a Venezia, dove sposò Donata (probabilmente della famiglia Loredano), da cui ebbe tre figlie. Fino alla morte, il viaggiatore veneziano si occupò, con lo zio Matteo, di affari e commercio oltre della diffusione del suo libro. Nell'agosto del 1307 consegnò una copia de *Il Milione* a Thibault de Cepoy, affinché la recapitasse a Carlo di Valois<sup>76</sup>, fratello del re di Francia, Filippo il Bello<sup>77</sup>.

---

<sup>73</sup> Rustichello da Pisa è ricordato, oltre che come compilatore de *Le Livre de messer Marco Polo citoyen de Venise, appelé Milion, où sont décrites les merveilles du monde*, anche come uno dei volgarizzatori del ciclo bretone compose il *Meliadus*, romanzo in prosa così intitolato dal nome del padre di Tristano

<sup>74</sup> è una lingua romanza medievale derivante dal gallo-romano; è la lingua da cui si è sviluppata nel corso dei secoli l'attuale lingua francese, imposta come lingua ufficiale dell'antica Gallia

<sup>75</sup> (28 agosto 1215 – 18 febbraio 1294) è stato un condottiero mongolo, Khagan (il quinto, dal 1260 al 1294), nonché fondatore del primo Impero cinese della Dinastia Yuan

<sup>76</sup> Anche detto Carlo di Francia, in francese *Charles de Valois* fu conte di Valois dal 1286

<sup>77</sup> Filippo IV di Francia detto il Bello, in francese *Philippe le Bel*, fu re di Francia dal 1285 alla sua morte. Membro della dinastia dei Capetingi fu soprannominato il Bello per la sua bella apparenza

Oltre a quest'ultimo, se ne procurano trascrizioni, nel 1428 l'infante di Portogallo, don Pedro<sup>78</sup> e numerosi nobili e principi. Il libro fu ben presto *volgarizzato* per circolare in versioni toscane più o meno fedeli, e riscosse, fin dai primi del Trecento, un notevole successo. Il 9 gennaio 1324 Marco firmò il suo testamento che, insieme ad altri documenti, rivelò come le proprietà dei Polo fossero in realtà più limitate rispetto alle meravigliose ricchezze che venivano loro attribuite. La modesta casa della famiglia, nell'odierna Corte del Milion, ne dà conferma. Il «*nobilis vir Marchus Paulo Milioni*», così come l'illustre viaggiatore è nominato in un documento del 1305, morì a Venezia nel 1324.

---

<sup>78</sup> Pietro di Aviz o principe Pietro delle Sette Parti (del Mondo) (in portoghese: Pedro, infante de Portugal, 1° duque de Coimbra o Infante D. Pedro das Sete Partidas [do Mundo]).

## 6.2 Il Milione

«[...] E questo vi conterà il libro ordinatamente siccome messere Marco Polo, savio e nobile cittadino di Vinegia, le conta in questo libro e egli medesimo le vide. Ma ancora v'à di quelle cose le quali elli non vide, ma udille da persone degne di fede, e però le cose vedute dirà di veduta e l'altre per udita, acciò che 'l nostro libro sia veritieri e senza niuna menzogna. [...]» (Proemio)



Figura 12: Il testo latino di Marco Polo (forse tradotto da Fra' Pipino da Bologna all'inizio del XIV secolo)

Il Milione viene definita un'opera saggistico-biografica che narra dei viaggi di Marco Polo nell'allora sconosciuto Estremo Oriente. In seguito alla partecipazione ad una battaglia nel 1298, probabilmente a quella di Curzola, combattuta dalla Repubblica di Genova contro la Repubblica di Venezia, Marco Polo fu tenuto in carcere per alcuni mesi. In quel periodo dettò in lingua d'oïl a Rustichello da Pisa, anch'egli prigioniero dei genovesi, *Le Divisament dou Monde* (La Descrizione del Mondo), conosciuto poi con il nome de Il Milione.

In seguito, la versione originale del libro fu rimaneggiata da autori francesi, i quali apportarono delle correzioni personali e delle modifiche linguistiche, sia durante che dopo il periodo del Rinascimento, aggiungendo, ad esempio, qualche pittura miniaturizzata che, se da una parte ebbe lo scopo di abbellire l'opera rendendola più gradevole, dall'altra la impoverì sul piano qualitativo. Contribuì, infatti, a attribuirle l'immagine di uno scritto denso di fantasticherie e relativo a un mondo inesistente o immaginario, genere letterario molto diffuso in quell'epoca. Solamente durante il periodo dell'Illuminismo, *Il Milione* nella sua versione originale, venne rivalutato e inserito nell'ambito delle scoperte ed esplorazioni.

Nel libro si narra del viaggio dei Polo attraverso l'Anatolia, l'Armenia e della loro discesa fino al Tigri, toccando probabilmente Mossul e Bagdad. Si presume, essi giunsero poi, fino al porto di Ormuz con l'intenzione di proseguire via mare, mentre, in realtà, continuarono seguendo la via terrestre e, attraverso la Persia e il Khorasan, raggiunsero Balkh e il Badakhshan. Superarono, in quaranta giorni, il Pamir e scesero verso il bacino del Tarim e, attraverso il deserto dei Gobi giunsero ai confini del Catai, nel Tangut: la provincia più occidentale della Cina. Quindi proseguirono lungo la parte settentrionale dell'ansa del Fiume Giallo, arrivando a Khanbaliq, dopo un viaggio durato tre anni e mezzo. Al loro arrivo nel Catai, Marco Polo ottenne, in breve tempo, i favori di Kubilai Khan, tanto da essere nominato suo consigliere e quindi, ambasciatore. Kubilai Khan, infatti, gli conferì incarichi relativi a missioni, tra le quali quella a Sayangfu, dimostrando a Marco e alla sua famiglia riconoscimenti in diverse occasioni. Dopo 17 anni, i Polo ripartirono per fare ritorno a Venezia e fu durante quel viaggio, che accompagnarono la principessa Kocacin, nipote di Kubilai Khan, dal re di Persia, suo promesso sposo.

### **6.2.1 L'ellisse della cognizione ne *Il Milione***

Per realizzare l'analisi quantitativa de *Il Milione*: secondo e *originale* caso di studio di questa tesi, sono state utilizzate differenti fonti [3][4][5][6][7], tra le quali il supporto *cartaceo* dell'opera, nella versione *Il Milione* di Messere Marco.

- Le fasi operative del progetto sono iniziate con l'individuazione degli oggetti geografici e delle relative *occorrenze* all'interno del testo. La rilevazione dei riferimenti geografici corretti è stata possibile grazie alla consultazione di differenti versioni dell'opera e delle relative note, al fine di rintracciare il maggior numero di indicazioni spaziali e geografiche espresse nel corpus in forma *esplicita, implicita*, con nomi desueti e di fantasia. In questa delicata fase, è stato determinante poter consultare un numero cospicuo di *note*, presenti nelle differenti versioni del testo, per poter effettuare una interpretazione corretta dei toponimi corrispondenti agli oggetti geografici nominati;
- si è passati, poi, alla fase del *geocoding*<sup>79</sup> o *georeferenziazione*<sup>80</sup> che ha previsto l'attribuzione dell'informazione relativa alla precisa posizione dell'oggetto geografico nel sistema di coordinate decimali. In altre parole, è stata rintracciata, e per rispondere a un'esigenza di uniformità con la consultazione esclusiva delle sole pagine web del portale *geohack*<sup>81</sup>, la latitudine e la longitudine espresse in gradi decimali (si sono considerate 4 cifre decimali) di ogni oggetto geografico rinvenuto nella fase precedente. L'informazione utile, a questo punto del nostro lavoro, è stata quella espressa nel particolare *sistema geodetico*<sup>82</sup> di riferimento (longitudine e latitudine) al fine di descrivere la posizione dell'oggetto geografico sulla superficie terracquea;
- le informazioni ottenute, nelle prime fasi dell'analisi, hanno costituito i *dati di input* che sono stati implementati in un foglio elettronico finalizzato alla descrizione del modello matematico. È stato indicato, in due colonne differenti: la colonna A e la colonna E sia il toponimo antico che quello moderno, rilevato molto spesso dalla consultazione delle note;

<sup>79</sup> <http://en.wikipedia.org/wiki/Geocoding>

<sup>80</sup> <http://it.wikipedia.org/wiki/Georeferenziazione>

<sup>81</sup> [http://onestepahead.de/index.php?option=com\\_content&view=article&id=151:wikipedia-geohack&catid=34:news&Itemid=14](http://onestepahead.de/index.php?option=com_content&view=article&id=151:wikipedia-geohack&catid=34:news&Itemid=14)

<sup>82</sup> [http://it.wikipedia.org/wiki/Sistema\\_geodetico\\_di\\_riferimento](http://it.wikipedia.org/wiki/Sistema_geodetico_di_riferimento)

Il modello matematico riassunto nel foglio elettronico (che si allega alla relazione nel file *Il Milione di Marco Polo.xls*) è stato implementato sulla base di algoritmi formali ispirati ai principi delle scienze fisiche e della matematica trigonometrica.

Riassumendo, i *dati di input* inseriti nel foglio elettronico sono i seguenti:

- gli oggetti geografici identificati con il toponimo antico;
- il corrispondente toponimo moderno;
- la latitudine e la longitudine espresse in gradi decimali;

I *dati di output* sono rappresentati dai valori attesi utili per l'analisi quantitativa del corpus. In particolare, il caso *originale* che Il Milione rappresenta, comporta l'introduzione di parametrizzazioni complesse che non erano previste per l'analisi quantitativa del testo del Decameron.

Il foglio elettronico predisposto ha in comune, quindi, molti dei campi utilizzati per il Decameron; alcuni, al contrario sono *originali* e sono finalizzati al calcolo degli assi della percezione e dell'angolo  $\Phi'$ . Specifichiamo le colonne e le celle che saranno contenute nel foglio elettronico:

<i>Colonna A</i>	Oggetti geografici rilevati dal testo
<i>Colonna C</i>	Numero di occorrenze di ogni singolo oggetto geografico
<i>Colonna E</i>	Toponimo moderno corrispondente
<i>Colonna F</i>	Latitudine espressa in gradi decimali (DD)
<i>Cella F<sub>3</sub></i>	Latitudine del centro matematico
<i>Colonna G</i>	Longitudine espressa in gradi decimali (DD)
<i>Cella G<sub>3</sub></i>	Longitudine del centro matematico
<i>Colonna I</i>	Latitudine espressa in radianti
<i>Cella I<sub>3</sub></i>	Latitudine in radianti del centro matematico
<i>Colonna J</i>	Longitudine espressa in radianti
<i>Cella J<sub>3</sub></i>	Longitudine in radianti del centro matematico
<i>Colonna K</i>	Coseno della latitudine
<i>Cella K<sub>3</sub></i>	Coseno della latitudine del centro matematico



<i>Colonna L</i>	Seno della latitudine
<i>Cella L<sub>3</sub></i>	Seno della latitudine del centro matematico
<i>Colonna M</i>	Coseno della longitudine
<i>Cella M<sub>3</sub></i>	Coseno della longitudine del centro matematico
<i>Colonna N</i>	Seno della longitudine
<i>Cella N<sub>3</sub></i>	Seno della longitudine del centro matematico
<i>Colonna O</i>	Differenza con la latitudine del centro matematico
<i>Colonna P</i>	Differenza con la longitudine del centro matematico
<i>Colonna Q</i>	Prodotto della colonna O per la costante 111 per ottenere dai gradi di latitudine i corrispondenti km
<i>Colonna R</i>	Prodotto della colonna P per la costante 72,5 per ottenere dai gradi di longitudine i corrispondenti km
<i>Colonna S</i>	Peso dell'oggetto geografico n
<i>Colonna U</i>	Distanza euclidea di ogni oggetto geografico dal centro di percezione precedentemente ricavato (dalla radice quadrata della somma dei quadrati delle differenze delle coordinate omonime dei punti)
<i>Colonna V</i>	Calcolo della distanza $d(n)$ sulla superficie terrestre dell'ennesimo oggetto geografico dal centro di percezione
<i>Colonna W</i>	Calcolo della coordinata sferica $\varphi'_n$
<i>Colonna X</i>	Controllo congruità dei dati della colonna V (per evitare cifre negative)
<i>Colonna Y</i>	Coseno del quadrato della colonna X
<i>Colonna Z</i>	Seno del quadrato della colonna X
<i>Cella C<sub>n</sub></i>	Sommatoria della colonna C
<i>Cella F<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna F
<i>Cella G<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna G
<i>Cella I<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna I
<i>Cella J<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna J;
<i>Cella L<sub>n</sub></i>	Coordinata pesata Z
<i>Cella L<sub>n+1</sub></i>	Coordinata A;
<i>Cella L<sub>n+2</sub></i>	Latitudine del <i>centro matematico</i>
<i>Cella M<sub>n</sub></i>	Coordinata pesata X
<i>Cella N<sub>n</sub></i>	Coordinata pesata Y
<i>Cella N<sub>n+1</sub></i>	Coordinata B

<i>Cella N<sub>n+2</sub></i>	Longitudine del <i>centro matematico</i>
<i>Cella O<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna O
<i>Cella P<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna P
<i>Cella Q<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi Colonna C * colonna Q
<i>Cella R<sub>n</sub></i>	Matrice Somma Prodotto (pesata) n-elementi colonna C * colonna R
<i>Cella S<sub>n</sub></i>	Somma dei pesi degli n oggetti geografici = 1
<i>Cella U<sub>n</sub></i>	Distanza Euclidea
<i>Cella W<sub>n</sub></i>	Coordinata polare planare I <sub>+</sub>
<i>Cella Y<sub>n</sub></i>	Coordinata polare planare I
<i>Cella Z<sub>n</sub></i>	Coordinata polare planare I <sub>12</sub>
<i>Cella Z<sub>n+1</sub></i>	Angolo $\Phi$
<i>Cella Z<sub>n+2</sub></i>	Quadrato dell'asse della percezione maggiore <b>a</b>
<i>Cella Z<sub>n+3</sub></i>	Quadrato dell'asse della percezione minore <b>b</b>

Uno degli aspetti più interessanti, di questo secondo caso di studio, è rappresentato dal fatto che l'applicazione dell'algoritmo formale, per il calcolo del centro di percezione, definisce, in realtà, non un centro della percezione ma un centro teorico-matematico. Tale centro risulta localizzato nella zona del Tibet ed è identificato con una latitudine di  $33,8836^\circ$  e una longitudine di  $80,2528^\circ$ . Non essendo un luogo di specifico interesse per la narrazione non può definirsi centro di percezione ma, è altrettanto importante in quanto risulta il punto centrale di *un'area di interesse*, che si è delineata in termini spaziali, molto più vasta, anche in senso *trasversale*, di quanto ci saremmo potuti aspettare. A conferma della correttezza dei valori che contraddistinguono il centro matematico possiamo affermare che il punto geografico localizzato nel Tibet si trova a una trascurabile distanza dalla traiettoria rappresentata dall'itinerario di Marco Polo. Il raggio della percezione, per Il Milione, è stato calcolato in 3.172 km per una distanza euclidea di 2.709 km. Dalle risultanze si nota come il raggio e la distanza euclidea risultino *molto estesi* in termini di chilometri.

Una tipologia di parametri come questa, connotati in questo modo, per questo studio, non ci dovrebbero sorprendere, in quanto, essi sono rivelatori di un genere letterario come quello della *letteratura di viaggio* e, in particolare, adeguati a una narrazione e che si estende dalla città di Venezia fino in Cina. Ciò che, al contrario, deve destare la nostra attenzione è la «larghezza» della distribuzione geografica degli oggetti rinvenuti nel testo (v.tav.20-24); in altre parole, una distribuzione che si presenta ampia non solo in termini *longitudinali* ma anche, e soprattutto, in quelli *latitudinali* rappresentando una considerevole «larghezza» cognitiva (che dal punto di vista geometrico sarà rappresentata da un'ellisse non troppo schiacciata, come in questi casi sarebbe prevedibile). Per poterne studiare i caratteri si definiscono, mediante l'applicazione di algoritmi complessi e sofisticati, una serie di parametri *originali* che saranno identificati negli *assi principali* della *percezione*, nell'*ellisse* della *percezione* e nell'*angolo  $\Phi$* '. In altre parole, ne *Il Milione*, viene utilizzata una versione più *evoluta* e *inconsueta* degli algoritmi fisici applicati per l'analisi del *Decameron*. Non è sufficiente, infatti, per questo caso di studio, calcolare il centro e il raggio della percezione, in quanto tali parametri non risulterebbero perfettamente rispondenti alle caratteristiche intrinseche dell'opera di Marco Polo. Per *Il Milione*, al contrario, viene definito un centro *teorico-matematico* che, dai calcoli, risulta posizionato nella zona del Tibet (con latitudine di  $33,8836^\circ$  e longitudine di  $80,2528^\circ$ ) in un punto non troppo distante dalla traiettoria del viaggio di Polo; e questo a conferma della correttezza dei calcoli impiegati. Inoltre, per delineare l'area percettiva dell'opera sarà necessario definire i due assi della percezione rappresentativi delle direzioni privilegiate dell'autore al momento della scrittura del suo resoconto di viaggio. La geometria approssimativamente circolare che era stata delineata dalla distribuzione degli oggetti geografici nell'analisi dell'opera di Boccaccio (v.tav.8-12), per questo specifico studio, risulterà deformata e schiacciata riducendosi a una forma ellittica che sarà incentrata nella zona del Tibet. Dalla posizione degli assi principali che formano l'ellisse della percezione si definisce, in particolare, l'area dell'Asia sud-orientale quale *reale* centro psicologico e cognitivo di Marco Polo (v.tav.24) in quanto, in questa zona ha luogo la massima concentrazione degli oggetti geografici rinvenuti nel testo.

L'elemento, da sottolineare, per la configurazione topografica degli oggetti, è rappresentato dalla definizione della zona indocinese (v.tav.24) come *reale* centro dell'interesse cognitivo e culturale di Marco Polo; una zona «largamente» al di fuori dell'itinerario del suo viaggio. In altre parole, si denota che l'area cognitiva di Marco Polo non strettamente collegata al solo percorso del suo viaggio ma rivelatrice, al contrario, di un'attenzione per il «*mondo intero*». Concludiamo affermando come, dalle risultanze numeriche appena delineate, sia possibile ritenere come «parziale e riduttiva» la semplice qualificazione culturale data a Il Milione di «resoconto di viaggio», in quanto, esso, come dimostrano le teorie fisiche applicate all'analisi del testo, fu redatto in un'ottica *enciclopedica* manifestando un chiaro interesse di Marco Polo per ogni luogo «altro».

## 6.2.2 Gli orizzonti cognitivi di altri testi medievali

A conferma della *singolarità* delle risultanze ottenute nell'esame del Decameron e de Il Milione, eseguiamo un parallelismo con l'esame quantitativo che ha per oggetto altre opere del periodo del Medioevo illustrato nello studio del prof. Paolo Rossi[1]. Ci sembrato opportuno riportare questo studio perché, solamente analizzando i risultati derivanti dal calcolo del centro e del raggio della percezione di altri testi medievali, possiamo interpretare correttamente la straordinarietà dell'ampiezza cognitiva risultante dai nostri casi di studio.

Le opere sono testimoni della storiografia del X secolo della Francia settentrionale e, risultano collegati fra loro per contenuti e per origine, provengono infatti tutti e tre dalla città di Reims<sup>83</sup>, centro episcopale di notevole importanza in quel periodo storico. Gli autori dell'analisi sono il canonico Flododard 893/4-966)<sup>84</sup> autore di *Annales*, il monaco di Saint Remi, Richer<sup>85</sup> (del quale non si conoscono i dati biografici) con *Historiae o Historiarum libri IV* e il monaco benedettino poi Papa (999-1093) Silvestro III, Gerbert d'Aurillac<sup>86</sup> con *Epistolae*.

---

<sup>83</sup> <http://it.wikipedia.org/wiki/Reims>

<sup>84</sup> Flodoard, *Annali* (919-966), Introduzione, traduzione, note ed excursus di P. Rossi, PLUS, Pisa 2007

<sup>85</sup> Richer di Saint-Remi, *I quattro libri delle Storie* (888-998), Introd. e note di P. Rossi, PLUS, Pisa 2008

<sup>86</sup> Gerbert d'Aurillac, *Lettere*(883-897), Introduzione, trad. e note P. Rossi, PLUS Pisa 2009

Nella tabella sottostante, si riassumono le risultanze dell'esame sui tre dei testi:  
 Centro delle percezione per i testi: *Reims (centro di provenienza degli autori)*

<i>Autore</i>	<i>N. località citate</i>	<i>Raggio della percezione</i>
Flodoard	181 su 800 citazioni	123 km
Richer	230 su 1300 citazioni	200 km
Gerbert	160 su 540 citazioni	280 km

Dal prospetto si deducono chiaramente i «parametri-chiave» dell'analisi effettuata sui testi. La *media del raggio* della percezione complessivo per le opere francesi è di 200 km. Il raggio della percezione è l'importante indicatore della distanza tra il centro episcopale e metropolitano di Reims e l'orizzonte della cognizione, oltre il quale limite i fatti, per l'autore, assumono una scarsa o irrilevante importanza. Un raggio della percezione così definito risulta *ampio* se lo si inserisce nel contesto storico del Medioevo e se si considerano le difficoltà di spostamento dell'epoca.

Un ulteriore studio riportato nello stesso testo del prof. Rossi[1] riguarda i corpora danteschi della *Commedia* e del *De Vulgari Eloquentia* (DVE). I risultati conseguiti e che definiscono il raggio della percezione risultano interessanti e ancora *più ampi* per estensione rispetto agli scrittori francesi:

Centro di percezione: *Firenze*

<i>Titolo</i>	<i>N. località citate</i>	<i>Raggio della percezione</i>
Commedia	364 su 743 citazioni	673 km
DVE	n.p.	200 km

Nei testi danteschi si nota che, a fronte di un centro della percezione delineato nella città fiorentina, la media complessiva del raggio della percezione è notevolmente definita in 437 km circa.

### 6.2.3 Risultanze numeriche

I valori ottenuti dall'analisi quantitativa de Il Milione tracciano un'area della percezione dalla forma ellittica e un orizzonte cognitivo *molto esteso* geograficamente. Possiamo valutare l'ampiezza dell'elisse della percezione risultante analizzando la rappresentazione con ARCGIS on line (v.tav.24), con GeoCommons (v.tav.20-21-22), infine con i prodotti del portale di Google (Google Earth tav.23 e Google Map tav.23\_bis). Nelle mappe topografiche si evidenziano, non solamente il valore della longitudine considerevolmente elevato ma comune risultanza per testi appartenenti al genere letterario dei racconti di viaggio verso mete molto lontane, ma anche quello della latitudine che testimonia una *larghezza* dell'area cognitiva e un interesse enciclopedico per il *mondo intero* da parte di Marco Polo. Uno scenario cognitivo che possiamo definire, senza dubbio, «*fuori dal comune*» nel Medioevo. Questa valutazione è suffragata anche dal parallelismo con i testi medievali dei quali si è riportata l'analisi al par.6.2.2.

Ne Il Milione, il numero totale degli oggetti geografici corrisponde a 169 per un totale di 1105 occorrenze.

Così come è avvenuto per il Decameron, è stato effettuato il raggruppamento, nell'oggetto geografico di principale, dei nomi di quartieri, di chiese, di piazze, di contrade nominate nel testo e ubicati in zona; analogamente si è operato nel caso di quartieri o frazioni dislocate nelle immediate vicinanze delle città principali. Inoltre, è stata dedicata particolare attenzione a identificare correttamente gli oggetti geografici che nel testo sono stati oggetto di *sottintesi* oppure nominati con l'utilizzo di parafrasi nel testo. Per la gestione degli oggetti estesi e la risoluzione delle problematiche incontrate nell'analisi si rimanda rispettivamente ai par. 4.5 e 6.2.5.

Il centro teorico-matematico è localizzato nella zona del Tibet (precisamente nelle *immediate* vicinanze di Rutog County nel Tibet, longitudine 33.913333° e latitudine 80.750278°) definito dalle seguenti coordinate geografiche:

33,8836°

80,2528°

Un centro matematico così connotato ci appare particolarmente interessante e singolare. Esso non deve essere interpretato come un luogo di specifico interesse per la narrazione (che si trova a una irrilevante distanza, peraltro, dalla linea che, dell'itinerario Venezia – Cina, idealmente possiamo tracciare e che non si configura come un centro della percezione) ma come il centro matematico di un'estensione molto più vasta, dal punto di vista topografico, non solo in senso longitudinale ma anche e soprattutto trasversalmente per quanto concerne la sua latitudine, di quella che, teoricamente, al momento della pianificazione dell'analisi del testo, ci aspettavamo di riscontrare.

Sulla base del *centro matematico*, definito mediante l'applicazione di algoritmi formali avanzati e complessi, è stato identificato un nuovo sistema di coordinate sferiche  $\theta'_n$  e  $\varphi'_n$  (che rispettivamente identificano la latitudine e la longitudine) (v. par. 4.6) sulla base delle quale calcolare gli elementi che costituiscono il tensore  $I_+$ ,  $I_-$  e  $I_{12}$  (v. par. 4.6) che si quantificano nel seguente modo:

$$I_+ = 5.818.527$$

$$I_- = -2.493.807$$

$$I_{12} = 507.890$$

Gli assi della percezione che identificano le direzioni preferenziali ortogonali fra loro, sono definiti dai parametri corrispondenti al quadrato della lunghezza degli assi  $a$  e  $b$  e dall'angolo  $\Phi'$ , rispettivamente con le seguenti risultanze numeriche:

$$a^2 = 8.363.527 \text{ da cui } a = 2.892 \text{ km}$$

$$b^2 = 3.273.527 \text{ da cui } b = 1.809 \text{ km}$$

$$\Phi = -0,100456^\circ$$

Gli assi principali e precisamente l'asse principale  $a$  e l'asse secondario  $b$  costituiscono l'ellisse della percezione che risulta orientata nello spazio rispetto all'asse polare in ragione dell'angolo  $\Phi$ .

Si riporta uno *schizzo informale* e elementare volto a chiarire la conformazione geometrica dell'ellisse della percezione risultante dai due assi ortogonali  $a$  e  $b$  e dall'angolo  $\Phi$ .

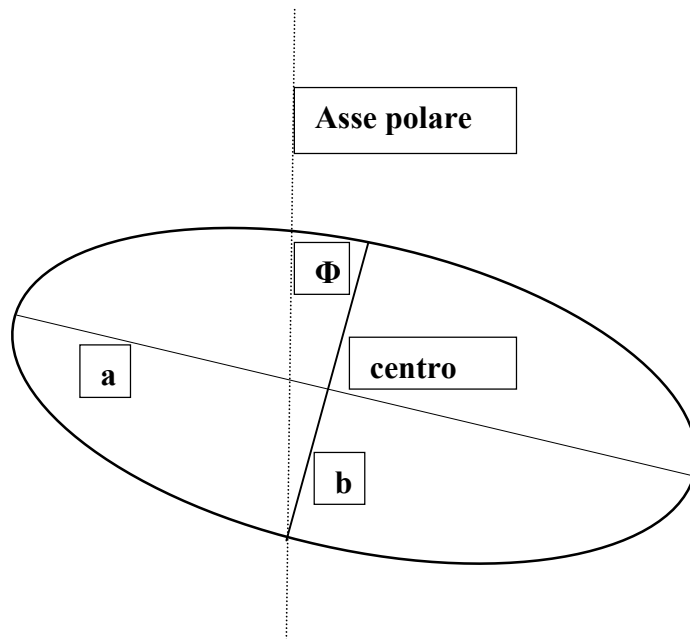


Figura12\_bis: **Riproduzione dello schizzo informale, in quanto non corrispondente nelle misure alle risultanze, della raffigurazione e del posizionamento degli elementi costitutivi dell'ellisse della percezione:**



Di seguito, nel diagramma di tipo *scatter plot* (diagramma di dispersione), si riproduce la distribuzione, nello spazio bidimensionale, degli oggetti geografici rinvenuti rappresentati in base alle coordinate geografiche decimali:

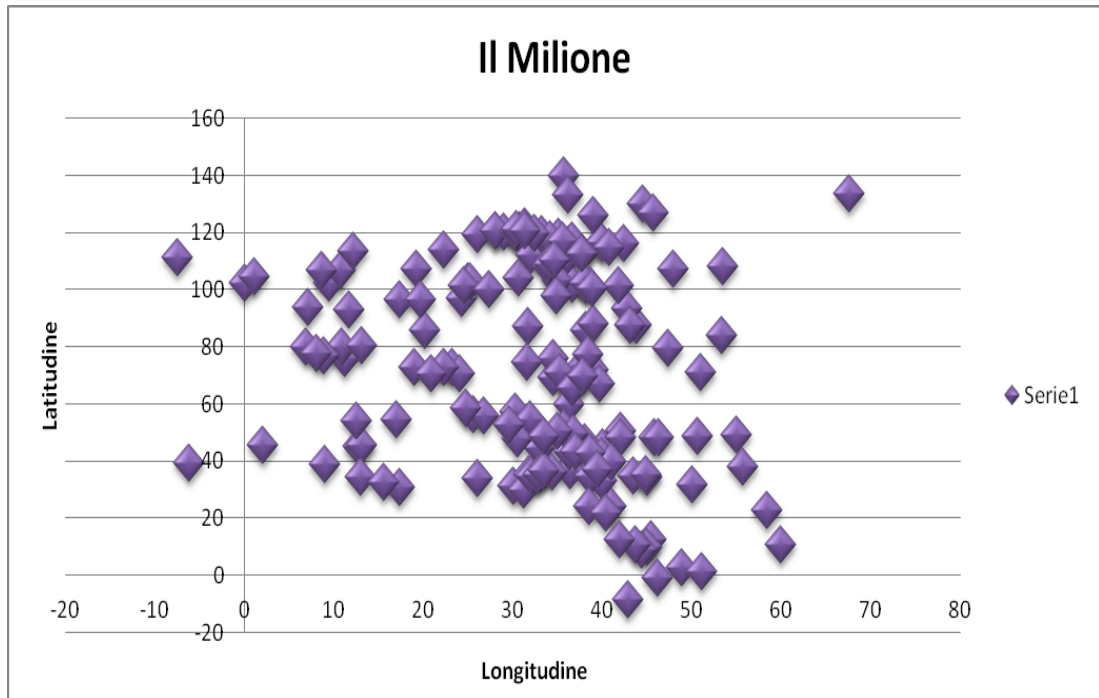


Figura 13: **Scatter Plot de Il Milione**

La connotazione geografica della forma assunta dall'ellisse della percezione costituita: dagli oggetti geografici rinvenuti nel testo posizionati in relazione alla relative georeferenze, dagli assi principali della percezione e orientata in ragione dell'angolo  $\Phi'$  e i cui valori sono stati calcolati dall'applicazione degli algoritmi formali complessi, può essere visivamente identificata sia nella fig.13 che, in modalità più efficace, le tav.20-21-22-23-23\_bis-24 allegate alla relazione. Le tav. 20-21-22 sono state elaborate con gli strumenti messi a disposizione del portale *GeoCommons*®, le visualizzazioni di cui alle tav.23 e 23\_bis con gli strumenti di *Google*® e, infine, la tav. 24 è stata realizzata con *ArcGis online*®. Gli strumenti utilizzati e i dettagli delle visualizzazioni statiche e dinamiche elaborate saranno oggetto di descrizione e approfondimento al par. 8.2.1.

Per riassumere, invece, le risultanze relative alle occorrenze rinvenute nel testo, riportiamo la tabella *pivot* di Excel(di Microsoft® Office) che riproduce le prime località rinvenute nel testo, ordinate in modo decrescente per numero di occorrenze, il numero delle occorrenze registrate e la percentuale rispetto al totale complessivo delle occorrenze<sup>87</sup>:

<b>Localita'</b>	<b>O ccorrenze</b>	<b>% del Totale</b>
<i>Sumatra O ceano Indiano</i>	<b>56</b>	<b>5,07</b>
<i>Ulan Bator Mongolia Asia</i>	<b>53</b>	<b>4,80</b>
<i>Malabar Melibar India meridionale</i>	<b>43</b>	<b>3,89</b>
<i>Teheran Iran</i>	<b>42</b>	<b>3,80</b>
<i>Hang Tcheu Q inghai</i>	<b>36</b>	<b>3,26</b>
<i>Wenzhou Cina meridionale</i>	<b>34</b>	<b>3,08</b>
<i>Mumbai India</i>	<b>29</b>	<b>2,62</b>
<i>Pechino Repubblica Popolare Cina</i>	<b>23</b>	<b>2,08</b>
<i>Pegu Birmania</i>	<b>19</b>	<b>1,72</b>
<i>Altre Località</i>	<b>770</b>	<b>69,68</b>
<b>Totale C omplessivo</b>	<b>1105</b>	<b>100%</b>

Figura 14: Tabella Pivot

<sup>87</sup> Estratto da Foglio elettronico Il Milione di Marco Polo.xls presentato in allegato a questa relazione

Infine, il diagramma a torta (Pie Chart – Excel di Microsoft®Office), più intuitivo, per denotare le proporzioni in percentuale delle occorrenze rispetto al totale complessivo:

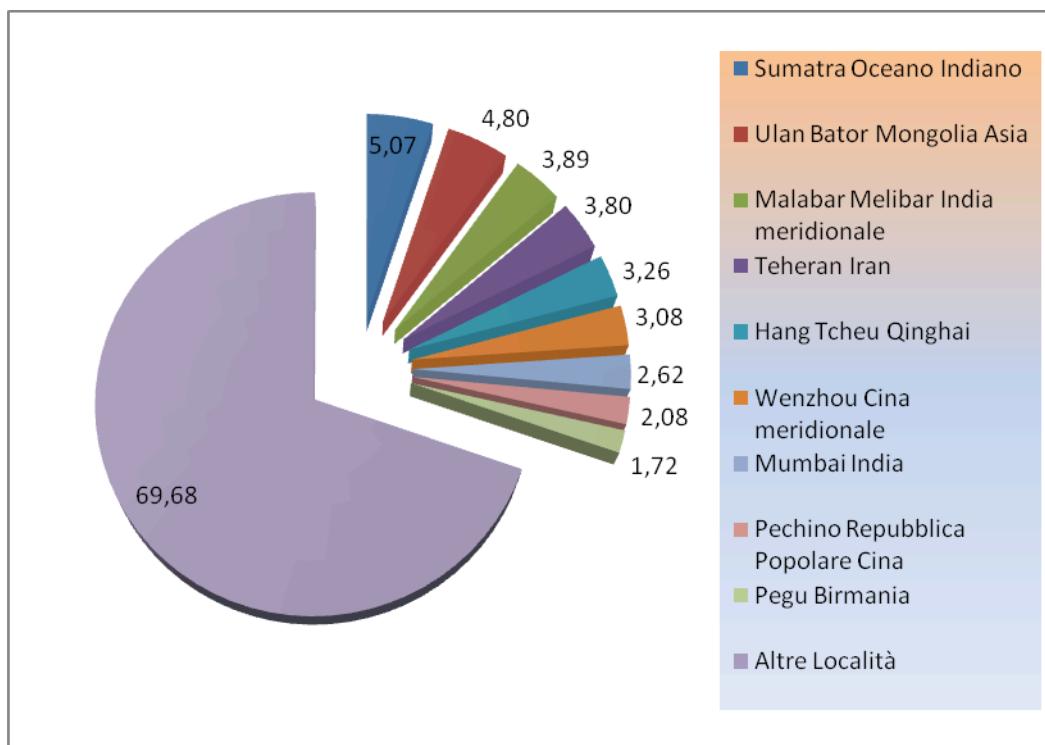


Figura 15: Pie Chart (modalità visuale) delle risultanze precedenti<sup>88</sup>

<sup>88</sup> Estratto da Foglio elettronico Il Milione di Marco Polo.xls presentato in allegato alla relazione

#### 6.2.4 Lo scenario culturale

Il calcolo formale eseguito sul corpus de *Il Milione* e derivante dall'applicazione degli algoritmi complessi produce un raggio della percezione di *3.172 km* e una distanza euclidea di *2709 km*. Il centro matematico è stato localizzato nel punto georeferenziato con  $33,8836^\circ$  di latitudine e  $80,2528^\circ$  di longitudine. L'ellisse della percezione che è stata definita è formata dall'asse maggiore **a** di *2.892 km* e quella minore indicata con **b** di *1.809 km*. L'angolo  $\Phi'$  in ragione del quale, rispetto all'asse polare, risulta posizionata l'ellisse della percezione è pari  $-0,100456^\circ$ . Con questa configurazione geometrica e topografica si definisce lo scenario cognitivo e mentale espresso da Marco Polo ne *Il Milione*. Ci troviamo di fronte a una risultanza che può essere, senza dubbio, definita inusuale nell'epoca medievale.

Con un'analisi approfondita del contesto socio culturale, nel quale l'opera di Marco Polo si è sviluppata, è possibile confermare e comprovare la fondatezza del risultato numerico ottenuto dall'impiego degli algoritmi formali complessi impiegati. Si introducono, quindi, di seguito, i *fattori-chiave* che sono implicati nella definizione di un orizzonte cognitivo così ampio. Si nota dalle risultanze che solo marginalmente esse risultano determinate, nella loro ampiezza, dal genere letterario di *letteratura di viaggio*, del quale, l'opera di Marco Polo appartiene. Abbiamo infatti asserito al par.6.2.1, alla luce dei parametri numerici calcolati, come sia possibile ritenere «parziale e riduttiva» la qualificazione culturale data a *Il Milione* di «resoconto di viaggio», in quanto, esso, come dimostrano le teorie fisiche applicate nell'analisi del testo, fu redatto in un'ottica enciclopedica manifestando un interesse di Marco Polo per ogni luogo «altro».

Come si è appurato dall'esame dei testi medievali, la cui analisi è stata riportata al par. 6.2.2, e come abbiamo confermato dall'analisi dei due casi di studio, la sostanziale *immobilità* del mondo medievale fu smentita dalla presenza di numerosi viaggiatori, spinti da esigenze e da obiettivi diversi: *clerici vagantes*, girovaghi, giullari, pellegrini (i viaggiatori medievali per eccellenza), mercanti.

Accanto ai viaggiatori *reali*, il Medioevo, ne conosce altri, puramente letterari, come i cavalieri dei romanzi cortese-cavalleresco impegnati in un cammino individuale di perfezionamento e di affermazione del proprio valore. In tutti i casi, *il «viaggio», reale o immaginario*, si accompagna al pericolo e assume un valore iniziatico di crescita, di scoperta di sé attraverso la scoperta dell'altro e dell'altrove. In questo contesto storico è, soprattutto, il mercante dell'età comunale a introdurre nel chiuso e statico mondo feudale, una spinta dinamica, destabilizzante e che rappresenta l'espressione di una rivalutazione dell'individuo, di una ritrovata fiducia nelle capacità personali. I mercanti-viaggiatori di Boccaccio ne sono una testimonianza letteraria esemplare; alcuni di loro lasciando interessanti autobiografie, come nel caso del mercante fiorentino Giovanni di Paolo Morelli con *Ricordi*<sup>89</sup>.

Non è una casualità, con questi presupposti, che si debba proprio a un mercante medievale come Marco Polo, il primo grande libro di viaggio, *Il Milione*. Dall'analisi del suo testo, egli risulta, sicuramente, spinto dalla necessità di esplorare nuove vie commerciali, ma anche dalla *pura curiosità* e dall'esigenza di ampliare la sua conoscenza diretta o immaginata che lo dirige verso l'ignoto, verso un qualsiasi luogo «altro». In altre parole, Marco Polo interpreta, ne *Il Milione*, l'intraprendenza della nuova classe in ascesa<sup>90</sup>.

Per tutti queste ragioni è chiaro comprendere, in questa fase, il motivo della *considerevole ampiezza* dell'ellisse della percezione che si è delineata dall'analisi del testo di Marco Polo e che denota uno scenario cognitivo che oltrepassa largamente i territori compresi nel suo viaggio in Cina.

Nelle tav. dalla 20 alla 24 che riportano le visualizzazioni elaborate, in modalità statica e dinamica, relative agli oggetti geografici rinvenuti ne *Il Milione*, si evidenzia chiaramente l'interesse enciclopedico per il *mondo intero* del viaggiatore Marco Polo.

---

<sup>89</sup> Il fiorentino Giovanni di Paolo Morelli (1371-1444), mercante iscritto all'Arte della Lana, compose i suoi *Ricordi* dal 1393 al 1411

<sup>90</sup> [http://www.letteratour.it/altro/A01\\_letteratura\\_e\\_viaggio.asp](http://www.letteratour.it/altro/A01_letteratura_e_viaggio.asp)

Riportiamo, in conclusione, altri aspetti, che pensiamo possa essere attinenti e altrettanto interessanti, che gravitano intorno alla stesura dell'opera di Marco Polo [6] [7] e che contribuiscono a chiarire le ragioni della grande diffusione del libro.

Il Milione all'origine fu scritto in lingua d'oïl, ma già prima della fine del XV secolo fu tradotta in lingua latino-francese, nei dialetti toscani e veneziani e, successivamente, in moltissime lingue europee. La scelta del francese, in particolare, fu quasi certamente dettata dalla volontà dell'autore di diffondere il testo a una platea internazionale. Da numerosi analisi filologiche è stato dimostrato che il manoscritto franco-italiano edito da Marco Polo agli inizi del XIV secolo, è *molto vicino* all'originale andato perduto e presenta numerose corrispondenze, dal punto di vista lessicografico, con il testo del romanzo di Rustichello da Pisa Meliadus<sup>91</sup>. Il manoscritto, nella sua prima redazione, porta, infatti, tutti i segni della collaborazione tra Marco Polo e Rustichello da Pisa e dimostra anche che il ruolo di quest'ultimo è stato qualcosa di diverso da quello che avrebbe caratterizzato un semplice scrivano.

Meliadus, l'opera di Rustichello, inizia, infatti, quasi con le stesse parole riportate in apertura de Il Milione: «Signori imperatori, re e duci e tutte altre genti».

Ma l'aspetto che, più di ogni altro, interessa l'analisi delle motivazioni della grande diffusione dell'opera, coincide con il fatto che Rustichello, in realtà, contribuisce in maniera determinante a indirizzare il romanzo verso un certo tipo di pubblico, fornendo a Marco Polo l'accesso alla lingua della letteratura cortese<sup>92</sup> con cui venivano scritti i romanzi che gli imperatori, i re e i principi ascoltavano leggere e che parlavano di *aventures* dei cavalieri erranti. Anche se, in effetti, la rappresentazione di Marco Polo nell'opera non corrisponde precisamente a quella del cavaliere errante, l'utilizzo della lingua d'oïl<sup>93</sup> mista a termini veneziani e italiani tipica del romanzo cortese, e a prima vista *inadatta* a un resoconto di viaggio, in realtà apre la strada a Marco Polo verso il vasto pubblico che si intratteneva prevalentemente con la letteratura cavalleresca. Ma non solo, Il Milione, per i suoi caratteri, si inseriva perfettamente e fin dalle origini nel genere della letteratura scientifica per documentare un ambito sempre più ampio del sapere e della volontà di dominio del tempo.

---

<sup>91</sup> [http://it.wikipedia.org/wiki/Rustichello\\_da\\_Pisa](http://it.wikipedia.org/wiki/Rustichello_da_Pisa)

<sup>92</sup> [http://it.wikipedia.org/wiki/Letteratura\\_medievale\\_francese](http://it.wikipedia.org/wiki/Letteratura_medievale_francese)

<sup>93</sup> [http://it.wikipedia.org/wiki/Lingua\\_d%27oïl](http://it.wikipedia.org/wiki/Lingua_d%27oïl)

Per questa ragione si può affermare che Rustichello da Pisa abbia contribuito fornendo in modo determinante anche il suo apporto di letterato. Egli fu, colui che, ascoltando il racconto per primo, ne poté intuire le potenzialità e le mise a frutto nel modo migliore; ne è la prova il successo riscosso da *Il Milione* fin dall'epoca della sua stesura e le numerose lingue nelle quali venne trascritto. In altre parole, come ascoltatore, Rustichello da Pisa prestò a Marco Polo la propria sapiente voce affinché egli potesse diffondere il suo grandioso sapere e grazie a questi fattori, Marco Polo poté diventare colui che conosceva l'Asia più di ogni altro essere umano del suo tempo.

Concludiamo con la frase di apertura de *Il Milione*, perché riteniamo possa riassumere in modo esaustivo, le motivazioni che hanno spinto Marco Polo a intraprendere il suo lavoro e che ne ha decretato il successo:

«(...) Ma voglio che voi sappiate che poi che Iddio fece Adam nostro primo padre insino al dí d'oggi, né cristiano né pagano, saracino o tartero, né niuno uomo di niuna generazione non vide né cercò tante maravigliose cose del mondo come fece messer Marco Polo (...)».

### **6.2.5 Problematiche operative**

Nell'ambito dell'analisi del secondo caso di studio: *Il Milione* di Marco Polo alcune delle problematiche che sono state affrontate coincidono con quelle che abbiamo dettagliato e analizzato parlando, al par. 5.2.4, a proposito del *Decameron*; altre sono state riscontrate e affrontate solamente per il testo in esame:

- un prima difficoltà ha riguardato la presenza di riferimenti a *luoghi geograficamente molto estesi*. Per la sua risoluzione, abbiamo definito, come dice la metodologia fisica, un *baricentro* al quale abbiamo attribuito le rispettive georeferenze che sono state estese diventando rappresentative di tutto l'oggetto geografico. In altre parole, per le nazioni menzionate nel corpus, sono state considerate le coordinate delle capitali e, nel caso delle regioni, quelle proprie dei capoluoghi.

Per i corsi d'acqua e i fiumi, invece, è stato calcolato il baricentro, considerando le coordinate geografiche della zona di origine e quelle della foce;

- per interpretare correttamente i toponimi degli oggetti geografici, espressi con riferimenti *impliciti*, sono state consultate plurime versioni dell'opera e le relative note integrative. In questo modo è stato possibile avere una quantità di informazioni considerevole che si è dimostrata determinante per la risoluzione di questa problematica;
- la corretta interpretazione dei nomi degli oggetti geografici identificati con toponimi antichi e desueti e di difficile riconoscimento soprattutto nel caso de Il Milione, ha causato un ritardo notevole nella conclusione dell'analisi e ha comportato necessariamente molteplici fasi di controllo delle risultanze evidenziate nelle analisi precedenti. La questione si è dimostrata di difficile soluzione in quanto la maggior parte dei toponimi menzionati hanno un'origine cinese, mongola o asiatica (ecc.) e, per questo motivo, la loro identificazione è stata possibile solo in seguito a un lavoro di ricerca e di verifica molto accurato. E' stato necessario consultare le note aggiornate delle differenti fonti utilizzate; esse contengono le informazioni più recenti che derivano dalle aggiornate interpretazioni filologiche dei testi e degli annali cinesi. Inoltre, ciò che nel Decameron era presente in maniera irrilevante e che, al contrario, nel resoconto di Marco Polo, rappresenta una componente frequente, è la presenza di luoghi con toponimi rappresentati da *nome di fantasia* oppure da nomi di località che le note relative al testo definiscono come «*non identificate*». Gli oggetti, in questo caso, non sono stati conteggiati perché non supportati da alcuna interpretazione utile alla loro identificazione.



Riassumendo, ne Il Milione, la frequente presenza di oggetti espressi da *toponimi simili fra loro* nella sintassi e di difficile interpretazione è stata superata grazie alla continua consultazione di differenti traduzioni dell'opera in parallelo. Esse riportano nelle note le teorie filologiche che sono alla base dell'interpretazione dei luoghi de Il Milione. Inoltre, tale analisi quantitativa ha reso necessario un riscontro puntuale e attento di diverse fonti accreditate disponibili aggiornate con le interpretazioni recenti degli annali cinesi, al fine di identificare i toponimi moderni che corrispondono ai termini desueti inseriti nel testo di Marco Polo e poter incorrere, nella misura minore possibile, nella possibilità di errori materiali.

Le fonti alle quali abbiamo fatto riferimento sono riportate nella bibliografia ai riferimenti: [3][4][5][6][7].

Nel foglio elettronico allegato alla relazione (Il Milione di Marco Polo.xls) è possibile notare che nella colonna A sono stati riportati i nomi antichi o di fantasia attribuiti da Marco Polo ai luoghi geografici della narrazione; nella colonna E sono riportati i toponimi *moderni corrispondenti* frutto dell'analisi filologica e della consultazione delle differenti fonti utilizzate;

- infine, un'ulteriore problematica sorta nell'analisi del testo di Marco Polo è rappresentata dall'assetto amministrativo nel quale, fin dall'antichità, è suddivisa la Cina. Attualmente, la costituzione della Repubblica Popolare Cinese stabilisce, *de iure* tre livelli di suddivisione, ma *de facto* nella Cina Continentale, vi sono cinque livelli di governo locale: la provincia, la prefettura, la contea, il comune e il villaggio. In seguito alla localizzazione geografica dell'oggetto rinvenuto nel testo, si è provveduto a collocarlo nella provincia amministrativa *moderna* definita *geograficamente* tra quelle più prossime rispetto a quella menzionata nel testo.

## 6.3 Conclusioni

Ne Il Milione, che rappresenta *la parte originale* di questo studio, sono stati rinvenuti n. 169 oggetti geografici per un totale di n. 1.105 occorrenze. Il calcolo formale del raggio di percezione del resoconto del viaggio di Marco Polo produce un risultato molto ampio (km 3.172), com'è facile intuire nel caso di un viaggio che copre una distanza notevole come quella da Venezia in Cina e una distanza euclidea altrettanto rilevante di 2.709 km.

Lo spazio cognitivo, risultante dall'analisi dell'opera, assume la forma *evoluta* (rispetto al cerchio della percezione delineato nel Decameron ) di un'ellisse della percezione, incentrata sul centro matematico del Tibet identificato con le coordinate 33,8836° e 80,2528°, e definita dai quadrati dei seguenti assi ortogonali a e b e dall'angolo  $\Phi$  che definisce la posizione della stessa ellisse rispetto all'asse polare:

$$\mathbf{a} = \mathbf{km} \ 2.892$$

$$\mathbf{b} = \mathbf{km} \ 1.809$$

$$\Phi = -0,100456^\circ$$

Sulla base del centro matematico si sono effettuati i calcoli dei parametri avanzati al fine di determinare l'ellisse della percezione. L'ellisse della percezione è formata dagli assi principali  $\mathbf{a} = \mathbf{km} \ 2.892$  e  $\mathbf{b} = \mathbf{km} \ 1.809$  ed è posizionata in ragione dell'angolo  $\Phi = -0,100456^\circ$  calcolato rispetto all'asse polare. Seguendo questa metodologia, è stata definita l'area cognitiva nella sua ampiezza (e soprattutto *larghezza*) che deriva dall'analisi quantitativa de Il Milione.

A conferma della fondatezza delle teorie applicate, le risultanze numeriche, che da esse si sono ricavate, sono state inserite nel contesto storico sociale e culturale di origine, al fine di proporre una logica serie di motivazioni agli interessanti connotati di uno scenario cognitivo così ampio e variegato e costituito da localizzazioni cognitive anche molto distanti rispetto alla traiettoria del viaggio narrato da Marco Polo nel suo resoconto «*enciclopedico*» di viaggio.

## 7 Il sito web: *Lo scenario Cognitivo in Letteratura*

### 7.1 Struttura e funzionalità

Il sito web *Lo scenario Cognitivo in Letteratura*, non pubblicato sul web, è stato realizzato esclusivamente al solo scopo *dimostrativo e illustrativo* delle tematiche e delle relative visualizzazioni che costituiscono l'oggetto della tesi di laurea specialistica dal titolo *Scenari Cognitivi e Culturali dei testi letterari: 2 casi di studio* presentata in questa relazione, nelle tavole e nei file.xls allegati alla stessa.

Il sito web si compone delle seguenti pagine:

➤ *index.html*:

Home Page: in questa pagina web vengono presentati i caratteri generali della tesi, il titolo, il candidato e il relatore. Viene inoltre riportato l'*abstract* del lavoro per riassumere le tematiche trattate nella relazione;

➤ *tesi.html* :

Tesi di laurea: vengono qui elencati i punti svolti che costituiscono la struttura della tesi;

➤ *decameron.html* :

il *Decameron*: in questa pagina è stata realizzata una galleria di immagini con JQuery nella quale, cliccando su ciascuna miniatura, appare una lente di ingrandimento che consente di visionare l'immagine corrente a dimensione reale relativa alla visualizzazione che ha per oggetto i fenomeni evidenziati nell'analisi quantitativa alle occorrenze e agli oggetti geografici dell'opera del Decameron. Viene, infatti, proposta la galleria fotografica delle visualizzazioni realizzate per il primo caso di studio e che sono state riunite dalla tav. 1 alla tav. 12;

➤ *milione.html* :

Il Milione : in questa pagina è stata realizzata una galleria di immagini con JQuery nella quale, cliccando su ciascuna miniatura, appare una lente di ingrandimento che consente di visionare l'immagine corrente a dimensione reale relativa alla visualizzazione che ha per oggetto i fenomeni evidenziati nell'analisi quantitativa alle occorrenze e agli oggetti geografici dell'opera de Il Milione. Viene, infatti, proposta la galleria fotografica delle visualizzazioni realizzate per il secondo caso di studio e che sono state riunite dalla tav. 13 alla tav. 24;

➤ *risultati.html*:

gli obiettivi raggiunti: vengono proposti i risultati raggiunti (v.par. 9.2 della relazione);

➤ *credits.html*:

vengono elencate le generalità della candidata alla laurea specialistica;

➤ *validazioni.html*:

in questa pagina finale vengono dettagliate le tipologie di validazione: W3CXHTML, W3CCSS e W3CWAI.

## 8 Le visualizzazioni

### 8.1 Visualizzazioni per i testi

Si è introdotta al par. 4.1 la linguistica computazionale facendo riferimento, in particolare, alle potenzialità insite nella disciplina per la comprensione delle proprietà dei linguaggi e per l'analisi della struttura e del contenuto di un corpus.

I principi delle visualizzazioni, sia statiche che dinamiche, applicate al testo offrono la possibilità di amplificare le potenzialità e l'efficacia dell'analisi sia qualitativa che quantitativa.

In relazione, ai nostri casi di studio, le metodologie della visualizzazione offrono un valido strumento per soddisfare task definiti in termini *quantitativi* e *qualitativi* comprendendo sia le rappresentazioni delle occorrenze rinvenute all'interno di ciascuna opera (analisi qualitativa delle occorrenze rinvenute), sia la rappresentazione grafica delle georeferenziazioni registrate nello stesso documento (analisi dei dati derivanti dalla georeferenziazione degli oggetti).

Queste ultime rappresentazioni, in modalità visuale e intuitiva, mostrano, in maniera assai chiara, la singolare ampiezza degli orizzonti cognitivi evidenziati nel Decameron e ne Il Milione.

Elaboreremo, in particolare, due categorie di visualizzazioni per focalizzare sia *l'aspetto delle occorrenze* (visualizzazioni relative alle occorrenze) evidenziate nel testo e quindi la loro percentuale di incidenza rispetto al totale rilevato, che *l'aspetto geografico* (visualizzazioni geografiche) strettamente legato alla distribuzione degli oggetti geografici nel globo terrestre.

## 8.2 Le visualizzazioni della geografia cognitiva

### 8.2.1 Le visualizzazioni relative alle occorrenze

Questa tipologia di visualizzazioni, nelle modalità sia statiche che dinamiche, applicate ai casi di studio, ha per oggetto la rappresentazione, in una modalità intuitiva, delle risultanze relative alla frequenza di comparizione di ciascun toponimo rinvenuto nel testo mostrandone, in particolare, il *peso* rispetto alla frequenza totale. Per mostrare le *risultanze relative alle occorrenze* useremo i seguenti strumenti:

- *Microsoft Office*® :
  - *Pie Chart*<sup>94</sup> o diagramma a torta in modalità statica (v.fig. 10 e fig.15):

Il diagramma circolare (comunemente indicato con il termine di *grafico a torta* o *diagramma circolare*) è un metodo utilizzato in statistica descrittiva per rappresentare graficamente le variabili quantitative *misurate* su classi di categorie (valori nominali). Un diagramma circolare viene costruito dividendo un cerchio in spicchi le cui ampiezze angolari sono proporzionali alle classi di frequenza. Nel caso di frequenze relative, la somma delle aree è 1 (ovvero 100%). Possiamo affermare che il grafico a torta è il modo forse più familiare e diffuso per rappresentare visivamente le percentuali.

- *Many Eyes*® di IBM:

è uno strumento<sup>95</sup> molto performante che fornisce uno spazio dinamico dando l'opportunità di utilizzare strumenti grafici di modellazione dei contenuti grafici e dinamici sia in 2D che in 3D. Con l'avvento dei siti Web e dei Social network, le visualizzazioni come *manyeyes.com* di IBM contribuiscono a rendere queste metodologie accessibili a un pubblico dei *non addetti ai lavori* e a decretarne una crescente popolarità [8].

- *Pie Chart* o diagramma a torta in modalità dinamica in 2D e 3D (v.tav. 2 e tav.14):

---

<sup>94</sup> [http://it.wikipedia.org/wiki/Diagramma\\_circolare](http://it.wikipedia.org/wiki/Diagramma_circolare)

<sup>95</sup> <http://www-958.ibm.com/software/data/cognos/manyeyes/>

- *Bubble Chart* o diagramma a bolle<sup>96</sup> (in versione dinamica v.tav.1 e tav.13 ):

Si tratta di una tipologia di grafico in cui ogni entità rappresentata è definita nel termine di tre parametri numerici distinti; in questa tipologia di grafico viene visualizzato un insieme di valori numerici usando una simbologia a cerchi di differenti dimensioni in base alla proporzione di incidenza dell'oggetto sul totale. Il bubble chart è particolarmente utile per insiemi di dati con decine a centinaia di valori, o con i valori che differiscono di diversi ordini di grandezza. Selezionando ogni bubble è possibile scorgere informazioni ulteriori. Muovendo il cursore sulle bolle vengono aperte finestre di dialogo che forniscono informazioni aggiuntive;

- *Tag cloud* (e *wordle*) (versione dinamica v.tav.4, tav.5, tav.6 , tav.16, tav.17,tav.18):

E' una visualizzazione delle frequenze con cui compaiono le parole (nel nostro studio viene utilizzata per rappresentare le frequenze di comparizione dei toponimi che indicano gli oggetti geografici). Essa permette di visualizzare la frequenza con cui parole compaiono in un dato testo e vedere la relazione tra una colonna degli oggetti geografici e la colonna indicante le rispettive occorrenze. La dimensione della parola visualizzata corrisponde alla *quantità* delle occorrenze associata a quello stesso termine. Per chiarire forniamo un esempio attinente per un testo: se il set di dati è un riassunto della trama di *Il Padrino*, sarà probabile vedere frequenti, che parole come «Corleone» e «mafia» redatte con un font di dimensione maggiore rispetto a quello utilizzato per indicare tag con parole meno frequenti come «aperto» o «ristorante».

In altre parole, il Tag Cloud consiste nella rappresentazione, in un formato molto gradevole, delle parole più frequenti (per i nostro casi dei riferimenti geografici associati alle rispettive occorrenze all'interno del corpus).

Per rendere, ancora maggiormente intuitiva tale metodologia e aumentarne le potenzialità, vengono utilizzati dei font di dimensioni proporzionali alla frequenza di comparizione nel testo (Viegas et al, 2007);

---

<sup>96</sup> [http://it.wikipedia.org/wiki/Diagramma\\_a\\_bolle](http://it.wikipedia.org/wiki/Diagramma_a_bolle)

- Il *Word Cloud Generator* (versione dinamica v.tav.3 e tav.15) consente di visualizzare la frequenza con cui i nomi riferiti agli oggetti geografici compaiono in un testo, o vedere la relazione tra una colonna indicante i nomi degli oggetti geografici e quella che rappresenta le relative occorrenze. Muovendo il cursore sull'indicazione delle cifre che rappresentano le occorrenze viene presentata una finestra di dialogo contenente gli oggetti geografici con quel numero di occorrenze;
- *Tree Map* (mappa albero in versione dinamica v. tav. 7 e tav. 19) è una modalità di visualizzazione compatta di dati organizzati in forma gerarchica. Una mappa albero evidenzia strutture gerarchiche in modalità molto efficace soprattutto per mostrare gli attributi dei nodi-foglia utilizzando la codifica della dimensione e del colore (gradienti di colori in base all'intensità di manifestazione del fenomeno). Esse consentono agli utenti di confrontare i nodi e sotto-alberi anche a diversi livelli di profondità. Anche in questo caso muovendo il cursore si evidenziano ulteriori informazioni relative all'oggetto geografico selezionato.



## 8.2.2 Le visualizzazioni relative agli oggetti geografici

Per rappresentare in modalità visuale la distribuzione geografica sulle mappe raffiguranti il globo terracqueo si sono utilizzate le *visualizzazioni geografiche*. Esse hanno per oggetto, infatti, la raffigurazione, tramite segnaposto segnalati con simbologie a scelta, delle latitudini e delle longitudini di ogni singolo oggetto geografico rinvenuto nel testo. In altre parole, la visualizzazione geografica risponde all'esigenza di visualizzare la disposizione degli oggetti rinvenuti e definiti nelle loro coordinate geografiche, all'interno dello spazio della mappa geografica.

Gli strumenti che abbiamo adoperato per questo genere di visualizzazioni sono:

- *Google® Earth*<sup>97</sup> ;

è un software che genera vedute virtuali della Terra utilizzando immagini satellitari ottenute dal telerilevamento terrestre, fotografie aeree e dati topografici memorizzati in una piattaforma GIS<sup>98</sup>. SIT è l'acronimo italiano di *Sistema Informativo Territoriale*; l'acronimo inglese *Geographic(al) Information System*, GIS, appunto, viene spesso usata erroneamente come sinonimo di SIT. Un sistema informativo territoriale o SIT è infatti «il complesso di uomini, strumenti e procedure (spesso informali) che permettono l'acquisizione e la distribuzione dei dati nell'ambito dell'organizzazione e che li rendono disponibili, validandoli, nel momento in cui sono richiesti a chi ne ha la necessità per svolgere una qualsivoglia attività» (Mogorovich, 1988). Si denomina GIS un sistema informativo computerizzato che permette l'acquisizione, la registrazione, l'analisi, la visualizzazione e la restituzione di informazioni derivanti da dati geografici (geo-riferiti). Secondo la definizione di Burrough (1986) «il GIS è composto da una serie di strumenti software per acquisire, memorizzare, estrarre, trasformare e visualizzare dati spaziali dal mondo reale».

---

<sup>97</sup> [http://it.wikipedia.org/wiki/Google\\_Earth](http://it.wikipedia.org/wiki/Google_Earth)

<sup>98</sup> [http://it.wikipedia.org/wiki/Sistema\\_informativo\\_geografico](http://it.wikipedia.org/wiki/Sistema_informativo_geografico)

Si tratta, quindi, di un sistema informatico in grado di produrre, gestire e analizzare dati spaziali associando a ciascun elemento geografico una o più descrizioni alfanumeriche.

- *Google® Maps*<sup>99</sup> (nome precedente *Google Local*):

è un servizio accessibile dal relativo sito web<sup>100</sup> e che consente la ricerca e la visualizzazione di mappe geografiche di buona parte della Terra. Il programma è distribuito gratuitamente dalla società Google. Inizialmente il programma si chiamava *Keyhole* ed era sviluppato dall'omonima società nel 2004 la società venne acquisita da Google e il software rinominato di conseguenza.

- *GeoCommons*<sup>101</sup> ®:

è uno strumento con ottime potenzialità inserito all'interno di una piattaforma libera di tipo collaborativo. Essa, infatti, offre la possibilità di condividere dati geografici e mappe, nonché effettuare operazioni che prevedono filtraggio dei dati, differenti modalità di visualizzazione, animazioni di vario genere, comprese le opzioni di condivisione. Il portale dà la possibilità di creare visualizzazioni interattive. E' possibile, infine, con GeoCommons esportare i dati, per esempio, nel formato geospaziale KML e utilizzarli su Google Earth e su applicazioni mobili.

- *Microsoft®Excel* (Scatter Plot in modalità statica) consente di creare diagrammi di dispersione rappresentando su un piano bidimensionale x, y rispettivamente la longitudine e la latitudine degli oggetti geografici rappresentandone visivamente la concentrazione (v.fig. 11 e fig.13);
- *ArcGis® (onLine)* – ESRI (v.tav. 12-24-27-28) è un sistema software completo per la creazione, diffusione e utilizzo dell'informazione geografica. ArcGIS fornisce un ambiente scalabile che consente l'implementazione di sistemi GIS per singoli utenti o più utenti in applicazioni desktop, server, sul Web ecc.

---

<sup>99</sup> [http://it.wikipedia.org/wiki/Google\\_Maps](http://it.wikipedia.org/wiki/Google_Maps)

<sup>100</sup> <http://maps.google.it/>

<sup>101</sup> <http://geocommons.com/>

ArcGIS è una famiglia integrata di prodotti software per la costruzione di un GIS completo ed è costituito da quattro ambienti operativi principali:

- **ArcGIS Desktop:** Una suite integrata di applicazioni professionali GIS, comprendente i prodotti software ArcView®, ArcEditor® e ArcInfo®.
- **Server GIS:** ArcIMS, ArcGIS Server e ArcGIS Image Server.
- **GIS mobile :** ArcPad® e ArcGIS Mobile per l'utilizzo sul campo.
- **Developer GIS:** Componenti software integrabili per sviluppatori che permettono di estendere i GIS desktop, costruire applicazioni GIS personalizzate, creare servizi GIS e applicazioni Web e sviluppare soluzioni *mobile*.

La versione di ARC GIS utilizzata per l'analisi geografica dei casi di studio è quella *online* libera ed è accessibile dal web; presenta funzionalità ridotte rispetto alla versione proprietaria.

I tool grafici sono strumenti software che permettono di utilizzare risorse di tipo collaborativo per visualizzare dati all'interno del web sottoforma di presentazioni ricche e altamente intuitive nonché molto efficaci.

## 9 Riepilogo e sviluppi futuri

### 9.1 Le tematiche trattate

L'intento di questo studio è stato quello di rappresentare le percezioni umane insite nei testi letterari, tramite l'utilizzo di *parametri quantitativi* impiegati, di solito, nell'analisi dei fenomeni legati alle *scienze esatte*. Si è cercato di cogliere, studiare ed esporre il legame esistente tra discipline filosoficamente relegate a contesti separati e antistanti.

Per consentire una migliore comprensione della materia trattata, sono stati dettagliati i *principi fisici* in base ai quali si sono fondati gli algoritmi formali utilizzati nell'analisi *quantitativa* dei due casi di studio. Nell'ambito del Decameron, primo caso di studio, infatti, si sono definiti lo *spazio* e l'*orizzonte* della *cognizione*, costituito dal *centro* e dal *raggio* della *percezione*. L'elemento *originale* è stato rappresentato dall'analisi quantitativa effettuata sul secondo studio: *Il Milione* di Marco Polo e dal calcolo, con l'impiego di algoritmi più complessi e avanzati, dell'ellisse della percezione. I risultati delle analisi quantitative di entrambi i testi sono stati oggetto, inoltre, di una valutazione nell'ottica dello scenario storico culturale dell'epoca di scrittura. In questo modo è stato possibile confermare l'efficacia delle teorie fisiche impiegate e la fondatezza dei parametri applicati nell'indagine. Per dare, quindi, completezza alla tematica trattata e sulla base delle risultanze ottenute, sono state presentate differenti tipologie di visualizzazione dei dati, nelle modalità statiche e dinamiche, che hanno come oggetto sia le occorrenze rintracciate che la distribuzione georeferenziata degli oggetti rinvenuti nei testi. Infine, i concetti esposti, nonché le relative visualizzazioni, sono stati inseriti in un sito web regolarmente *validato* dal titolo: *Scenari Cognitivi in Letteratura* creato, appositamente, allo scopo informativo e illustrativo.

## 9.2 Gli obiettivi raggiunti

Gli obiettivi, che ci siamo prefissati, di riuscire a proporre una metodologia nuova per eseguire l'analisi dei testi, basata essenzialmente sull'impiego di parametri quantitativi che appartengono alle scienze *dure*, crediamo siano stati raggiunti. Lo scopo era di trarre dai testi che hanno rappresentato i nostri due casi di studio, una serie di informazioni *aggiuntive* relativamente al testo, al loro autore e al periodo di scrittura dell'opera. Le risultanze sono state ottenute dall'applicazione di algoritmi formali di differente complessità e originalità per i due casi in esame. Per i due generi letterari, ai quali appartengono i testi che abbiamo studiato, sono state necessarie metodologie di approccio differente.

Nel caso del Decameron di Giovanni Boccaccio è stato interessante determinare il *centro di percezione* e il *raggio della cognizione* e scoprire come lo scenario cognitivo, dall'impiego di questi parametri, si è ottenuto fosse sorprendentemente *ampio* per il periodo di scrittura dell'opera.

Nel testo de Il Milione, l'analisi ci ha fornito l'occasione di indagare, non solamente sulla *lunghezza* della percezione insita nel resoconto di viaggio, dai risultati considerevoli in termini di chilometri, ma anche sulla *larghezza* della cognizione di Marco Polo. Abbiamo avuto, per Il Milione, la cui analisi ha fornito una connotazione *originale* al nostro studio, la necessità di operare, infatti, un'evoluzione dei concetti di centro e di raggio della percezione per introdurre, con l'applicazione di algoritmi più complessi, il calcolo dell'ellisse della percezione che esprime lo scenario cognitivo e culturale dell'opera. I risultati conseguiti, inoltre, sono stati oggetto di una serie di visualizzazioni diverse per evidenziare, in una modalità che risultasse particolarmente efficace, le risultanze *quantitative* e *geografiche* degli studi. Queste ultime hanno, infine, costituito la base sulla quale elaborare approfondite interpretazioni, in ambito socio culturale e storico, dei parametri ottenuti e verificare, in questo modo, la fondatezza delle teorie applicate.

### 9.3 Gli sviluppi futuri

Ipotizziamo che un'analisi che abbia per oggetto un confronto diacronico tra autori diversi ma affini per genere potrebbe aiutare all'individuazione degli elementi evolutivi (o involutivi) nella percezione spaziale o geografica, come l'effetto della trasformazione della sensibilità (che può essere segnalato come un fattore sociale che coinvolge una determinata collettività) nel corso del tempo.

Inoltre, sono molti gli esempi che potrebbero essere forniti e sui quali, un confronto diacronico tra opere dello stesso autore potrebbero essere molto interessanti per evidenziare un'evoluzione o anche un'involuzione sia culturale sia cognitiva dell'autore stesso nei diversi periodi della produzione. I risultati della comparazione tra testi forniscono, secondo i casi, parametri che possono portare a considerare un allargamento o, al contrario, a un restringimento dell'orizzonte cognitivo e culturale dell'autore.

Un aspetto del quale occorre essere assolutamente consapevoli è che la percezione della distanza, in ciascun individuo, è fortemente condizionata da fattori psicologici e sociali, che implicano un'anisotropia<sup>102</sup>, che, ricordiamo, è la caratteristica che lega un determinato fenomeno alla direzione, e la conseguente presenza di un'effettiva, quanto soggettiva, metrica dai caratteri non euclidei. Questa considerazione che deve stare alla base delle analisi quantitative sui testi, non deve distogliere la ricerca ad approfondire un ambito a nostro parere davvero interessante.

Un altro approccio davvero interessante, in questo ambito di studio, potrebbe essere rivolto all'identificazione delle *scale di distanza*, concretizzabile con lo studio di fonti contemporanee che presentino, al loro interno, elementi identificativi del concetto di distanza percepita; nel caso, per esempio, di resoconti di viaggio può essere costituito dalle specifiche relative ai tempi di percorrenza. A proposito dell'anisotropia, noi pensiamo, che la determinazione dell'ellisse della percezione possa dare un primo sostanziale contributo alla sua identificazione. In linea di principio, la precisione delle informazioni ricavate dai dati potrebbe essere importante per l'analisi degli elementi principali dell'ellisse, anche se occorre avere un campione di rilevamenti piuttosto numeroso.

---

<sup>102</sup> <http://it.wikipedia.org/wiki/Anisotropia>

Infine, un nuovo sviluppo dell'analisi quantitativa dei testi potrebbe avere per oggetto la determinazione dello spazio cognitivo all'interno del romanzo contemporaneo, compiendo anche comparazioni all'interno di stessi autori o diversi. Sarebbe interessante anche allargare le analisi dei testi in base alle nazionalità di provenienza degli autori, per generi differenti e anche tenendo conto di caratteri fra loro comparabili (connotati similari ecc.); infine, fornirebbe sicuramente molti spunti un'analisi di testi non specificatamente letterari che spaziassero, ad esempio, nell'ambito storico o scientifico, ecc.

# Sitografia e Bibliografia

- 
- <sup>i</sup> [http://it.wikipedia.org/wiki/Giovanni\\_Boccaccio](http://it.wikipedia.org/wiki/Giovanni_Boccaccio)  
[http://en.wikipedia.org/wiki/Giovanni\\_Boccaccio](http://en.wikipedia.org/wiki/Giovanni_Boccaccio)
- <sup>ii</sup> <http://it.wikipedia.org/wiki/Decameron>
- <sup>iii</sup> [http://en.wikipedia.org/wiki/The\\_Decameron](http://en.wikipedia.org/wiki/The_Decameron)  
[http://www.brown.edu/Departments/Italian\\_Studies/dweb/index.php](http://www.brown.edu/Departments/Italian_Studies/dweb/index.php)  
[http://it.wikipedia.org/wiki/Modello\\_matematico](http://it.wikipedia.org/wiki/Modello_matematico)  
<http://it.wikipedia.org/wiki/Interpolazione>  
<http://it.wikipedia.org/wiki/Estrapolazione>  
<http://it.wikipedia.org/wiki/Statistica>  
[http://it.wikipedia.org/wiki/Distribuzione\\_%28statistica%29](http://it.wikipedia.org/wiki/Distribuzione_%28statistica%29)  
[http://it.wikipedia.org/wiki/Statistica\\_descrittiva](http://it.wikipedia.org/wiki/Statistica_descrittiva)  
[http://it.wikipedia.org/wiki/Statistica\\_inferenziale](http://it.wikipedia.org/wiki/Statistica_inferenziale)  
<http://it.wiktionary.org/wiki/aleatorio>  
[http://it.wikipedia.org/wiki/Media\\_%28statistica%29](http://it.wikipedia.org/wiki/Media_%28statistica%29)  
<http://it.wikipedia.org/wiki/Varianza>  
[http://it.wikipedia.org/wiki/Mediana\\_%28statistica%29](http://it.wikipedia.org/wiki/Mediana_%28statistica%29)  
[http://it.wikipedia.org/wiki/Errore\\_sistematico](http://it.wikipedia.org/wiki/Errore_sistematico)  
[http://dizionari.corriere.it/dizionario\\_italiano/O/occorrenza\\_2.shtml](http://dizionari.corriere.it/dizionario_italiano/O/occorrenza_2.shtml)  
<http://www.wordreference.com/definizione/perifrasi>  
<http://progettomatematica.dm.unibo.it/Trigonometria/archi.htm>  
<http://en.wikipedia.org/wiki/Radian>  
[http://it.wikipedia.org/wiki/Raggio\\_terrestre](http://it.wikipedia.org/wiki/Raggio_terrestre)  
[http://it.wikipedia.org/wiki/Distanza\\_euclidea](http://it.wikipedia.org/wiki/Distanza_euclidea)  
[http://it.wikipedia.org/wiki/Seno\\_%28matematica%29](http://it.wikipedia.org/wiki/Seno_%28matematica%29)  
<http://it.wikipedia.org/wiki/Coseno>  
<http://edu.os3.it/html/manual/trigonometria.pdf>  
<http://it.wikipedia.org/wiki/Isotropia>  
<http://it.wikipedia.org/wiki/Varianza>  
<http://it.wikipedia.org/wiki/Ellisse>  
<http://it.wikipedia.org/wiki/Anisotropia>  
<http://www.wordreference.com/definizione/perifrasi>  
<http://progettomatematica.dm.unibo.it/Trigonometria/archi.htm>  
<http://en.wikipedia.org/wiki/Radian>  
[http://it.wikipedia.org/wiki/Raggio\\_terrestre](http://it.wikipedia.org/wiki/Raggio_terrestre)  
[http://it.wikipedia.org/wiki/Distanza\\_euclidea](http://it.wikipedia.org/wiki/Distanza_euclidea)



---

[http://it.wikipedia.org/wiki/Seno\\_%28matematica%29](http://it.wikipedia.org/wiki/Seno_%28matematica%29)  
<http://it.wikipedia.org/wiki/Coseno>  
<http://edu.os3.it/html/manual/trigonometria.pdf>  
<http://it.wikipedia.org/wiki/Isotropia>  
<http://it.wikipedia.org/wiki/Varianza>  
<http://it.wikipedia.org/wiki/Ellisse>  
<http://it.wikipedia.org/wiki/Anisotropia>  
<http://www.summagallicana.it/lessico/m/Milione.htm>  
[http://kidslink.bo.cnr.it/irrsaeer/marcopolo/il\\_milione/racconto/indicemilione.html](http://kidslink.bo.cnr.it/irrsaeer/marcopolo/il_milione/racconto/indicemilione.html)  
<http://www.df.unipi.it/~rossi/storia.html>

La data dell'ultimo accesso alle URL citate è: 7 novembre 2011

- [1] Rossi P., *Measuring Large Scale Space Perception in Literary Texts*, in «Physica A: Statistical Mechanics and its Applications», Volume 380, 1 July 2007, Pages 439-446
- [2] Boccaccio G., *Decameron*, a cura di V. Branca, Einaudi, Torino, 1980-1987 ISBN 88-06-1283-X
- [3] Polo M., *Il Milione di Messere Marco Polo Viniziano*, tomo II, G.Pagani, Firenze, MDCCCXXVII
- [4] Polo M., *Il Milione*, a cura di R. Alulli, Mondadori, 1954
- [5] Polo M., tradotto da M. Bellonci, Eri Roma, 1982
- [6] Münkler M., *Marco Polo: vita e leggenda*, Vita e Pensiero, Milano, 2001 EAN 9788834301241 ISBN 88-343-0474-8
- [7] Polo M., *Il Milione*, introduzione e note di M. Ciccuto, Rizzoli, Milano, 1998
- [8] Hearst Marti A., *Search User Interfaces*, Cambridge University Press.
- [9] Viegas FB, Wattenberg M, Kriss J, van Ham F., *Talk Before You Type: Coordination in Wikipedia*, Proceedings of the 40th Annual Hawaii International Conference on System Sciences, 2007

---

## Indice delle figure

<b>Figura 1:</b> Moda mediana media .....	13
<b>Figura 2:</b> Nella gaussiana la media, la moda e la mediana coincidono .....	14
<b>Figura 3:</b> Rappresentazione grafica della latitudine di Torino.....	26
<b>Figura 4:</b> Rappresentazione grafica della longitudine sempre di Torino .....	26
<b>Figura 5:</b> Un angolo misurato in radianti .....	28
<b>Figura 6:</b> Rappresentazione grafica del seno e del coseno.....	31
<b>Figura 7:</b> La rappresentazione grafica di un tensore.....	41
<b>Figura 8:</b> John William Waterhouse, The Decameron, 1916 (Liverpool, Lady Lever Art Gallery) .....	47
<b>Figura 9:</b> Elaborazione della tabella Pivot Microsoft® Excel per le occorrenze del Decameron .....	58
<b>Figura 10:</b> Pie Chart delle risultanze precedenti .....	59
<b>Figura 11:</b> Scatter Plot del Decameron .....	62
<b>Figura 12:</b> Il testo latino di Marco Polo( forse tradotto da Fra' Pipino da Bologna all'inizio del XIV secolo) .....	69
<b>Figura 12_bis:</b> Riproduzione di uno schizzo informale (non corrispondente nelle misure delle risultanze).....	80
<b>Figura 13:</b> Scatter Plot de Il Milione.....	81
<b>Figura 14:</b> Tabella Pivot.....	82
<b>Figura 15:</b> Pie Chart (modalità visuale) delle risultanze precedenti .....	83

---

# Indice delle tavole

- Tavola 1:** Bubble Chart Occorrenze Decameron (Many Eyes® visualizzazione dinamica);
- Tavola 2:** Pie Chart Occorrenze Decameron (Many Eyes® visualizzazione dinamica);
- Tavola 3:** Word Cloud Generator Occorrenze Decameron (Many Eyes® visualizzazione dinamica);
- Tavola 4:** Tag Cloud Color Occorrenze Decameron (Many Eyes® visualizzazione dinamica);
- Tavola 5:** Tag Cloud b/n Occorrenze Decameron (Many Eyes® visualizzazione dinamica);
- Tavola 6:** Tag Cloud n/b Occorrenze Decameron (Many Eyes® visualizzazione dinamica);
- Tavola 7:** Tree Map Color Occorrenze Decameron (Many Eyes® visualizzazione dinamica);
- Tavola 8:** Dati geografici Decameron GeoCommons® Grigio/Rosso;
- Tavola 9:** Dati geografici Decameron GeoCommons® Color;
- Tavola 10:** Dati geografici Decameron GeoCommons® Color\_2;
- Tavola 11:** Dati geografici Decameron Google®Earth;
- Tavola 11\_bis:** Dati geografici Decameron Google®Map;
- Tavola 12:** Dati geografici Decameron ArcGisOnLine® ;
- Tavola 13:** Bubble Chart Occorrenze Il Milione (Many Eyes® visualizzazione dinamica);
- Tavola 14:** Pie Chart Occorrenze Il Milione (Many Eyes® visualizzazione dinamica);
- Tavola 15:** Word Cloud Generator Occorrenze Il Milione (Many Eyes® visualizzazione dinamica);
- Tavola 16:** Tag Cloud Color Occorrenze Il Milione (Many Eyes® visualizzazione dinamica);
- Tavola 17:** Tag Cloud b/n Occorrenze Il Milione (Many Eyes® visualizzazione dinamica);
- Tavola 18:** Tag Cloud n/b Occorrenze Il Milione (Many Eyes® visualizzazione dinamica);
- Tavola 19:** Tree Map Color Occorrenze Decameron (Many Eyes® visualizzazione dinamica);
- Tavola 20:** Dati geografici Il Milione GeoCommons® Grigio/Rosso;
- Tavola 21:** Dati geografici Il Milione GeoCommons® Color;
- Tavola 22:** Dati geografici Il Milione GeoCommons® Color\_2
- Tavola 23:** Dati geografici Il Milione Google® Earth;
- Tavola 23\_bis:** Dati geografici Il Milione Google®Map;
- Tavola 24:** Dati geografici Il Milione ArcGisOnLine® ;
- Tavola 25:** L'itinerario del viaggio di Marco Polo;
- Tavola 26:** Screenshot della pagina index.html del sito web informativo;
- Tavola 27:** Screenshot della pagina decameron.html del sito web informativo;
- Tavola 28:** Screenshot della pagina milione.html del sito web informativo;

in allegato alla relazione:

- Tavole dalla n. 1 alla n. 28 specificate nell'indice delle tavole;
- Decameron di Giovanni Boccaccio.xls;
- Il Milione di Marco Polo.xls;