

# **IL SANGIOVESE ITALIANO NEL MONDO: NARRAZIONE E CONOSCENZA**



Costanza Nosi  
Gustavo Victor Rugnone  
Università LUMSA, Roma

# SOMMARIO

Indice delle tabelle.....	iv
Indice delle figure .....	v
Indice dei box.....	vii
<b>INTRODUZIONE .....</b>	<b>1</b>
<b>OBIETTIVI CONOSCITIVI E DISEGNO DI RICERCA.....</b>	<b>4</b>
Gli obiettivi della ricerca.....	4
Il disegno di ricerca .....	5
La procedura di raccolta dei dati.....	7
I criteri di selezione dei paesi.....	11
Elaborazioni e analisi condotte sui corpora .....	15
Impostazioni di analisi dei testi .....	16
Le elaborazioni realizzate .....	17
<b>I PAESI DI LINGUA INGLESE.....</b>	<b>19</b>
Analisi dei termini di ricerca .....	20
Analisi degli URL.....	20
Analisi tematica degli snippet.....	25

Analisi dei corpora.....	27
Australia .....	28
Canada .....	28
Regno Unito.....	28
USA.....	29
Associazioni di parole: Sangiovese .....	29
Analisi dei temi emergenti.....	34
Cluster semantica: Australia .....	36
Cluster semantica: Canada .....	41
Cluster semantica: Regno Unito .....	47
Cluster semantica: Stati Uniti d’America.....	54
<b>L’ITALIA.....</b>	<b>62</b>
Associazioni di parole: Sangiovese .....	63
Analisi dei temi emergenti.....	64
<b>CONSIDERAZIONI DI SINTESI .....</b>	<b>71</b>
<b>BIBLIOGRAFIA .....</b>	<b>75</b>

## **INDICE DELLE TABELLE**

Tabella 1: Caratteristiche corpus lingua inglese (Australia).....	28
Tabella 2: Caratteristiche corpus lingua inglese (Canada).....	28
Tabella 3: Caratteristiche corpus lingua inglese (Regno Unito) .	28
Tabella 4: Caratteristiche corpus lingua inglese (USA) .....	29
Tabella 5: Lemmi e associazioni per paese .....	33
Tabella 7: Caratteristiche corpus lingua italiana .....	62
Tabella 8: Lemmi e associazioni Sangiovese Italia.....	63

## **INDICE DELLE FIGURE**

Figura 1: Quote di mercato mondo dei motori di ricerca online .	8
Figura 2: Trend ricerche "Sangiovese" (01/2017-01/2018).....	8
Figura 3: Esempio di Google Autocomplete.....	9
Figura 4: Keyword e volume di ricerche online.....	10
Figura 5: Contribuzione keyword ai siti selezionati .....	11
Figura 6: Esportazioni di vini italiani .....	12
Figura 7: Esportazione di vini rossi toscani DOP .....	12
Figura 8: Concentrazione del volume medio di ricerche mensili	19
Figura 9: Concentrazione url - gruppo linguistico inglese.....	21
Figura 10: Codifica tematica degli snippet .....	22
Figura 11: Concentrazioni siti - gruppo linguistico inglese .....	23
Figura 12: Url comuni - gruppo linguistico inglese .....	24
Figura 13: Matrice degli url comuni .....	24
Figura 14: Collegamenti e distanza fra url e paesi .....	24
Figura 15: Relazioni fra le categorie degli snippet.....	26
Figura 16: Diagramma associazioni Sangiovese Australia .....	30
Figura 17: Diagramma associazioni Sangiovese Canada.....	31
Figura 18: Diagramma associazioni Sangiovese Regno Unito ...	31
Figura 19: Diagramma associazioni Sangiovese USA .....	32
Figura 20: Cluster Australia.....	36
Figura 21: Cluster e lemmi Australia.....	37
Figura 22: Wordcloud cluster Australia.....	38
Figura 23: Cluster Canada.....	42
Figura 24: Cluster e lemmi Canada .....	42
Figura 25: Wordcloud cluster Canada .....	44
Figura 26: Cluster Regno Unito 2D .....	47
Figura 27: Cluster Regno Unito 3D .....	48

Figura 28: Cluster e lemmi Regno Unito.....	48
Figura 29: Wordcloud cluster Regno Unito.....	49
Figura 30: Cluster USA .....	54
Figura 31: Cluster e lemmi USA.....	54
Figura 32: Wordcloud cluster Stati Uniti.....	55
Figura 33: Diagramma associazioni Sangiovese Italia .....	63
Figura 34: Cluster Italia.....	64
Figura 35: Cluster e lemmi Italia.....	65
Figura 36: Wordcloud cluster Italia.....	66

## **INDICE DEI BOX**

Box 1: Verbatim cluster 1_AU e Cluster 2_AU.....	38
Box 2: Verbatim Cluster 3_AU e Cluster 4_AU .....	40
Box 3: Verbatim Cluster_1_CAN e Cluster_4_CAN.....	43
Box 4: Verbatim Cluster_3_CAN .....	45
Box 5: Verbatim Cluster_2_CAN .....	46
Box 6: Verbatim Cluster_1_UK, Cluster_2_UK, Cluster_3_UK ..	50
Box 7: Verbatim Cluster_4_UK .....	52
Box 8: Verbatim Cluster_5_UK e Cluster_6_UK.....	53
Box 9: Verbatim Cluster_1_USA .....	56
Box 10: Verbatim Cluster_2_USA.....	58
Box 11: Verbatim Cluster_3_USA e Cluster_4_USA .....	59
Box 12: Verbatim Cluster_5_USA.....	60
Box 13: Verbatim Cluster_1_ITA.....	65
Box 14: Verbatim Cluster_2_ITA.....	67
Box 15: Verbatim Cluster_3_ITA – ampelografia.....	68
Box 16: Verbatim Cluster_3_ITA – etimologia .....	69
Box 17: Verbatim Cluster_4_ITA.....	70

# INTRODUZIONE

Il Sangiovese è uno dei vitigni più diffusi in Italia. Secondo l'ultimo censimento ISTAT, che risale tuttavia al 2010, il vitigno coprirebbe oltre l'11% (circa 71.500 ettari) della superficie vitata del paese. Secondo invece un più recente rapporto dell'OIV<sup>1</sup> del 2017, pur rimanendo la varietà di uva maggiormente coltivata in Italia, l'area dedicata al Sangiovese rappresenta ad oggi il 7,9% dell'intera superficie vitata, corrispondente a circa 54.000 ettari. È da evidenziare che il trend decrescente non riguarda solo questo vitigno, ma si inserisce in un contesto in cui, negli ultimi 15 anni, l'intera superficie vitata italiana, analogamente a quella di altri paesi europei tradizionali produttori di vino, si è ridotta approssimativamente del 20%.

È interessante accennare al fatto che il Sangiovese viene coltivato anche all'estero. In particolare, si trova diffusamente allevato in California, sia a Napa Valley che a Sonoma County, nel Sud America, in special modo in Argentina, in cui la massiccia immigrazione dall'Italia ha consentito che vi fossero portate barbatelle di vari vitigni dal nostro paese, tra cui il Sangiovese. Più recentemente, si è diffuso anche in Australia e in Nuova Zelanda, dove è stato sottoposto ad

---

<sup>1</sup>[https://www.federvini.it/images/pdf/Report\\_Oiv.pdf](https://www.federvini.it/images/pdf/Report_Oiv.pdf).

attenti processi di selezione clonale per adattarlo alle condizioni pedoclimatiche di questi paesi<sup>2</sup>.

A detta di Wikipedia, sito su cui con maggiore probabilità (si veda il prosieguo della ricerca) atterrano tutte le persone del mondo che cercano notizie su questo vitigno, le sue origini e la sua provenienza appaiono incerte. Le prime notizie relative al Sangiovese risalgono al XVI secolo, quando Giovan Vettorino Soderini parla di Sangiochetto o Sangiovetto in un suo scritto dell'epoca, dicendo che è un vitigno rimarchevole per la sua produttività regolare<sup>3</sup>. L'origine del nome è altresì incerta e le teorie alla sua base diverse: una possibile ipotesi è quella del "sangiovese", pensando al suo luogo d'origine in San Giovanni Valdarno, nell'aretino, in Toscana. Un'altra voce invece sembra far risalire il nome a forme dialettali ("san giovannina" uva primaticcia per il suo precoce germogliamento a fine giugno per la festa di San Giovanni Battista). C'è inoltre chi sostiene che derivi da "sangiovese", ovvero il "sangue di Giove", poiché il vitigno sembra provenire dal Monte Giove, vicino a Santarcangelo di Romagna.

Da recenti studi genetici sembra che il "Sangiovese", a dispetto della sua diffusa e storica presenza nell'area romagnola e tirrenica, possieda parentele con vitigni coltivati nel Sud Italia, soprattutto in Campania e Calabria (Palummina Mirabella e Calabrese Montenuovo). Dieci varietà ne costituiscono la "famiglia" e il Ciliegiole sembra sia un suo discendente diretto.

Indipendentemente dalle radici etimologiche del Sangiovese, il vitigno ha senza dubbio una grande rilevanza economica per il nostro paese venendo utilizzato, in purezza o in uvaggio, per la produzione di oltre 200 tra DOCG, DOC e IGT. Alcune di queste indicazioni

---

<sup>2</sup><https://cantine.pampanorama.it/magazine/i-vitigni-il-sangiovese-un-1164>.

<sup>3</sup><http://bressanini-lescienze.blogautore.espresso.repubblica.it/2014/10/06/il-mistero-del-sangiovese/>.

geografiche sono riservate a vini come il Chianti, il Chianti Classico e il Brunello di Montalcino, indiscussi ambasciatori dell'enologia Made in Italy nei mercati mondiali.

Nonostante questa rilevanza, non esiste alcuno studio sistematico sul Sangiovese con l'obiettivo di comprendere le dimensioni quali-quantitative della sua produzione in Italia e delle vendite sul mercato nazionale e all'estero, né le modalità con cui viene comunicato e promosso. Ecco che questo studio, che verrà in futuro completato da una ricerca di matrice agronomica sui dati strutturali e delle dinamiche di sviluppo delle superfici a Sangiovese, si prefigge di studiare lo storytelling del Sangiovese che viene veicolato online al pubblico, individuandone le basi del valore comunicato e i punti di distintività.

# **OBIETTIVI CONOSCITIVI E DISEGNO DI RICERCA**

## **GLI OBIETTIVI DELLA RICERCA**

Il principale obiettivo conoscitivo della ricerca consiste nel descrivere le caratteristiche dello storytelling del Sangiovese che viene veicolato online al pubblico, individuandone le basi del valore comunicato, i punti di distintività enfatizzati, verificarne le associazioni, la possibile personalità e gli eventuali legami con i territori di produzione.

Data la rilevanza che il vitigno riveste per l'economia italiana, essendo alla base di molteplici vini indiscussi ambasciatori del Made in Italy nel mondo, si è deciso di intraprendere una ricerca di tipo *cross-country*. Oltre all'analisi dello storytelling che viene fatto in italiano, l'indagine è stata condotta in lingua inglese e sono stati esaminati i siti consultabili da quattro differenti paesi anglofoni: Australia, Canada, Regno Unito, Stati Uniti d'America.

Come verrà meglio spiegato nella sezione in cui si illustra il disegno di ricerca, i contenuti testuali da sottoporre ad analisi sono stati individuati sulla base di un processo di selezione dei siti in cui erano contenuti che ha tenuto conto della probabilità che un ipotetico navigatore aveva di atterrare su quei siti nel caso in cui avesse

ricercato informazioni sul Sangiovese attraverso la digitazione di parole chiave su un motore di ricerca.

Implementando questa metodologia di indagine, di fatto, si sono raggiunti due sub-obiettivi conoscitivi. Innanzitutto si sono individuate le parole chiave più diffusamente utilizzate per ricercare informazioni sul Sangiovese online, poi si sono individuati i siti su cui, con maggiore probabilità, una persona alla ricerca di informazioni sul vitigno atterra. In altre parole, questi siti rappresentano le fonti informative che più diffusamente vengono consultate dagli individui e che hanno dunque maggior potere di influenzarne percezioni, conoscenze e atteggiamenti attraverso i contenuti e il tono della comunicazione veicolata.

## **IL DISEGNO DI RICERCA**

La metodologia adottata per l'indagine è di tipo quali-quantitativo e combina tecniche di content analysis qualitativa (Stemler, 2001) con il text mining (Lancia, 2012). In particolare, il text mining è una tecnica quantitativa di analisi che non si limita a "osservare" i testi online ma ne studia il linguaggio attraverso software dedicati capaci di produrre nuove informazioni da fonti testuali destrutturate quali categorie semantiche, frequenze e combinazioni di parole. L'utilizzo di software, inoltre, consente l'elaborazione e l'estrazione di informazioni anche da imponenti quantità di dati di testo. Lo studio ha previsto le seguenti fasi:

a. *Individuazione e selezione dei siti* – I siti sono stati individuati attraverso metriche tipicamente utilizzate nella ricerca online per l'identificazione delle pagine maggiormente consultate dagli utenti della Rete che effettuano ricerche, come rank Alexa, share of voice, ranking di visualizzazione sui motori di ricerca, ecc. (Antonopolus et al., 2017).

b. *Estrapolazione e sistematizzazione testi* – Una volta identificati i siti, sono stati estrapolati e raccolti i testi di interesse (corpora) e sono stati organizzati e sistematizzati per essere sottoposti all'analisi. Già dalla fase di sistematizzazione dei testi ci si è avvalsi del supporto del software T-Lab versione 9.1 (Lancia, 2014). Le operazioni propedeutiche all'analisi sono consistite nella pulitura dei testi (correzione di eventuali errori di battitura), disambiguazione e lemmatizzazione<sup>4</sup> (per un approfondimento, si rimanda al paragrafo Elaborazioni e analisi condotte sui corpora).

c. *Software-assisted content analysis qualitativa e text mining* – I corpora sono stati sottoposti a un'analisi qualitativa, content analysis, che consiste nella lettura critica da parte dei ricercatori membri del team per estrarne le informazioni rilevanti ai fini dell'obiettivo conoscitivo della ricerca. Nonostante la natura qualitativa dell'analisi posta in essere, ci si è comunque avvalsi del supporto del software per facilitare la sistematizzazione e la codifica delle informazioni.

Parallelamente, si è proceduto a effettuare una elaborazione attraverso T-Lab per esplorare, misurare e mappare i temi emergenti nei testi. Attraverso la metodologia del text mining, siamo stati in grado di estrarre ulteriori informazioni applicando differenti tecniche di elaborazione (relazioni di co-occorrenza tra le parole-chiave che vengono individuate all'interno dei testi, classificazione delle unità testuali attraverso un approccio bottom-up tramite l'analisi dei temi emergenti nei diversi corpora, elaborazione di mappe semantiche in

---

<sup>4</sup> "Nei dizionari linguistici che consultiamo, ogni voce corrisponde a un lemma che - generalmente - definisce un insieme di parole con la stessa radice lessicale (o lessema) e che appartengono alla stessa categoria grammaticale (verbo, aggettivo, etc.). Di norma, la lemmatizzazione comporta che le forme dei verbi vengono ricondotte all'infinito presente, quelle dei sostantivi e degli aggettivi al maschile singolare, quelle delle preposizioni articolate alla loro forma senza articolo, e così via. Ad esempio, le forme flesse "parliamo" e "parlato", risultanti dalla combinazione di un'unica radice (<parl->) con due diversi suffissi (<-iamo>, <-ato>), sono ricondotte allo stesso lemma "parlare"." (Lancia 2014, p. 214, <http://tlab.it/it/download.php>).

grado di rappresentare gli aspetti dinamici dei testi esaminati, individuazione di cluster semantici, ecc.).

### **La procedura di raccolta dei dati**

Essendo la presente ricerca basata su un'indagine di tipo *cross-country*, l'omogeneità dell'informazione deve essere garantita tra le diverse lingue e i paesi esaminati, e rappresentativa del fenomeno di studio. Nell'ambito del disegno di ricerca globale, l'informazione deve essere funzionale sia all'individuazione e selezione dei siti (fase a), sia all'estrapolazione e sistematizzazione dei testi (fase b).

La procedura di raccolta dei dati si è strutturata come segue:

1. *Informazioni generali*: la prima esplorazione è consistita nella consultazione di varie fonti riguardanti l'esportazione dei vini italiani in generale, e dei vini rossi toscani DOP in particolare. L'attenzione riposta su questi ultimi si giustifica per la loro significatività nella cornice di questa ricerca: la porzione di area toscana coltivata a Sangiovese è infatti pari al 62% della superficie iscritta all'albo viticolo, e il vitigno è utilizzato per la produzione di circa l'88% dei vini rossi della regione. In particolare, i dati considerati come riferimento sulle dinamiche di mercato sono quelli prodotti da *Wine Monitor* Gennaio-Dicembre 2016 di Nomisma<sup>5</sup>.

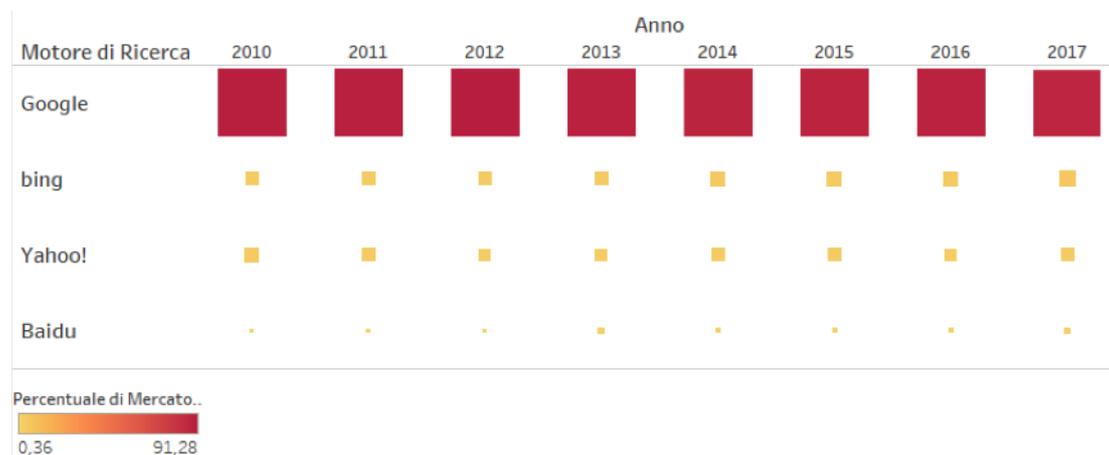
2. *Google Trends*: al fine di individuare il volume relativo di ricerca della parola "Sangiovese" e di altri termini di ricerca ritenuti di particolare importanza nella Rete è stato consultato *Google Trends*. La ricerca è stata effettuata considerando il periodo di tempo gennaio 2017-gennaio 2018, per avere un'immagine il più aderente possibile all'andamento delle ricerche. La scelta di utilizzare Google come motore di ricerca di riferimento si fonda sull'assoluta prevalenza di

---

<sup>5</sup><http://www.winemonitor.it/it/vino-numeri-2/dati/export-italia.html>.

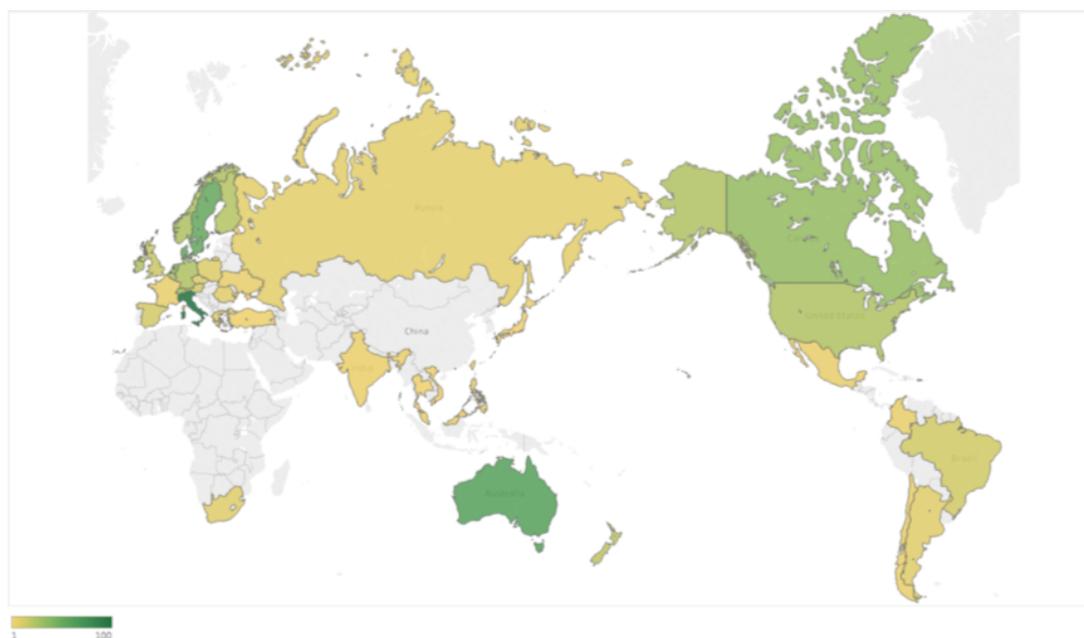
quest'ultimo come principale strumento di ricerca di informazioni online a livello mondiale.

Figura 1: Quote di mercato mondo dei motori di ricerca online



Fonte: elaborazione propria su dati StatCounter ([gs.statcounter.com](http://gs.statcounter.com))

Figura 2: Trend ricerche "Sangiovese" (01/2017-01/2018)

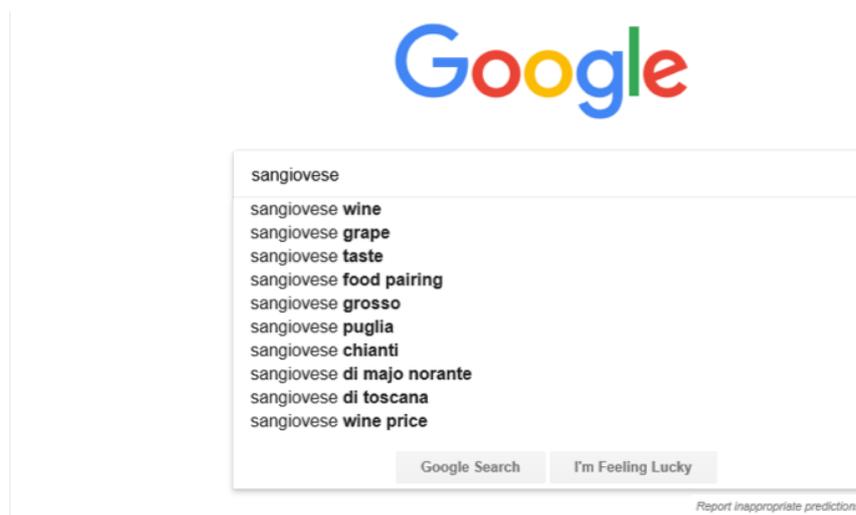


Fonte: elaborazione propria su dati Google Trends

Nel caso di *Google Trends* è necessario sottolineare che i risultati restituiti dallo strumento analitico sono sempre relativi al volume totale di ricerche effettuati in una certa zona nell'arco di tempo definito, e mai rappresentativi dal volume assoluto di ricerche

collegate a un termine. In particolare, i risultati di ricerca dello strumento sono proporzionali alla data e all'area geografica di una query. I dati rispecchiano le ricerche di una parola effettuate su Google. Ogni punto dati viene diviso per le ricerche totali dell'area geografica e l'intervallo di tempo che rappresenta per consentire di fare comparazioni sulla frequenza relativa. I risultati numerici ottenuti vengono scalati in un intervallo da 0 a 100 (rappresentato in Figura 2 dalla gamma di colore che va dal verde molto chiaro al verde intenso) in base alla proporzione dell'argomento su tutte le ricerche di tutti gli argomenti. Anche se per diverse aree geografiche viene indicato lo stesso interesse di ricerca per un termine, non significa che i relativi volumi di ricerca totali siano uguali<sup>6</sup>.

Figura 3: Esempio di Google Autocomplete



Fonte: Google.com

3. *Termini di ricerca*: la chiave di accesso all'informazione nei motori di ricerca sulla Rete sono i termini di ricerca o *keywords*. Per replicare nel modo più fedele possibile il processo di ricerca di un utente tipo su Google è stato utilizzato *Keyword Tool*, uno strumento che permette di avere accesso ai termini di ricerca associati e

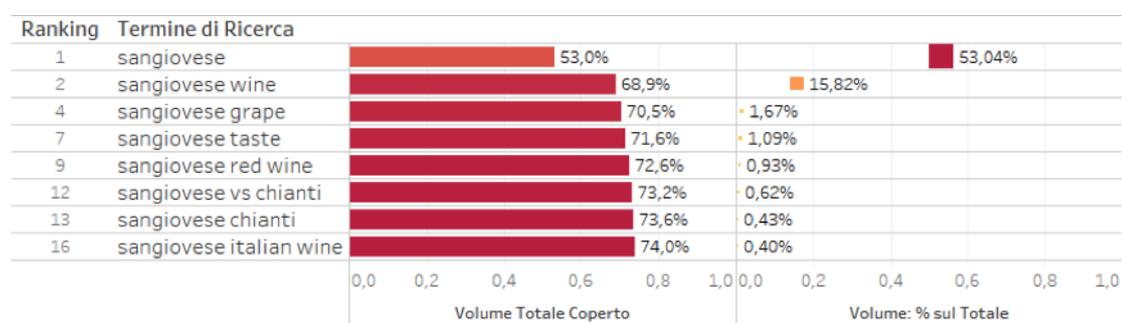
<sup>6</sup>[https://support.google.com/trends/answer/4365533?hl=it&ref\\_topic=6248052](https://support.google.com/trends/answer/4365533?hl=it&ref_topic=6248052).

automaticamente proposti da Google quando una persona inserisce un termine. Lo strumento rileva l'attività di *Google Autocomplete* per keyword che è stata digitata e restituisce i termini ad essa associati assieme al loro volume di ricerca per le lingue e/o le aree geografiche impostate nei filtri di ricerca di Keyword Tool.

I termini associati e suggeriti forniti da *Keyword Tool* sono stati organizzati in ordine di rilevanza decrescente tenendo conto del loro volume di ricerca.

Come mostrato in Figura 4, le ricerche relative alla parola Sangiovese equivalgono al 53% del volume di ricerche complessivo con tutte le keyword che includono Sangiovese realizzate in lingua inglese.

Figura 4: Keyword e volume di ricerche online



Fonte: elaborazione propria su dati Keyword Tool ([keywordtool.io](http://keywordtool.io))

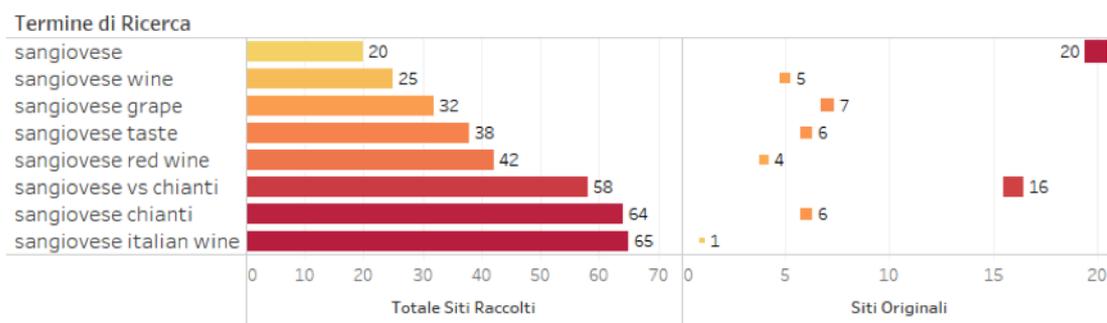
#### 4. Raccolta testi da Google:

- a) la raccolta dei testi è stata realizzata dalle pagine proposte da Google ipotizzando di effettuare la ricerca nei diversi paesi analizzati;
- b) partendo dal termine di ricerca con il maggior volume si è proceduto a selezionare e incorporare nella base dati dei primi 20 risultati restituiti da Google solo i testi pertinenti al

concetto "Sangiovese" inteso come narrazione attorno al vino oppure al vitigno<sup>7</sup>;

- c) nessun contenuto, indipendentemente da quante volte venisse riproposto nei risultati dell'esplorazione delle varie *keyword*, è stato inserito più di una volta nella base dati dei testi;
- d) la fine del processo di raccolta è stato determinato dal livello di saturazione ottenuto.

Figura 5: Contribuzione keyword ai siti selezionati



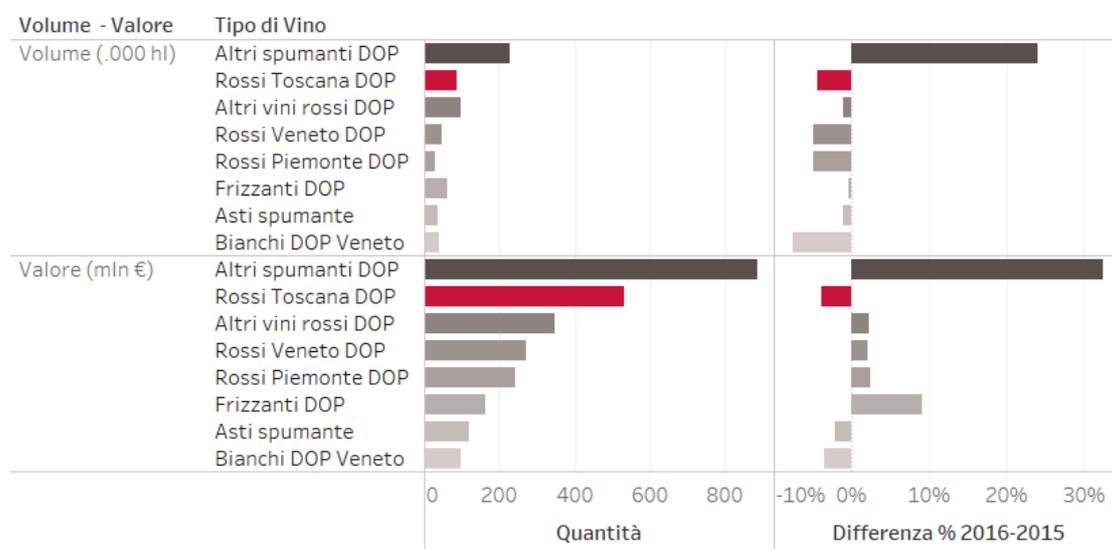
Fonte: elaborazione propria su dati Keyword Tool ([keywordtool.io](http://keywordtool.io))

## I criteri di selezione dei paesi

Le considerazioni fatte per selezionare i paesi da includere nell'analisi dello storytelling del Sangiovese online comprendono fattori differenti. Un primo elemento considerato riguarda innanzitutto i paesi di destinazione delle esportazioni di vino italiano in generale, oltre che dei vini rossi toscani DOP che sono in larghissima parte realizzati con questo vitigno.

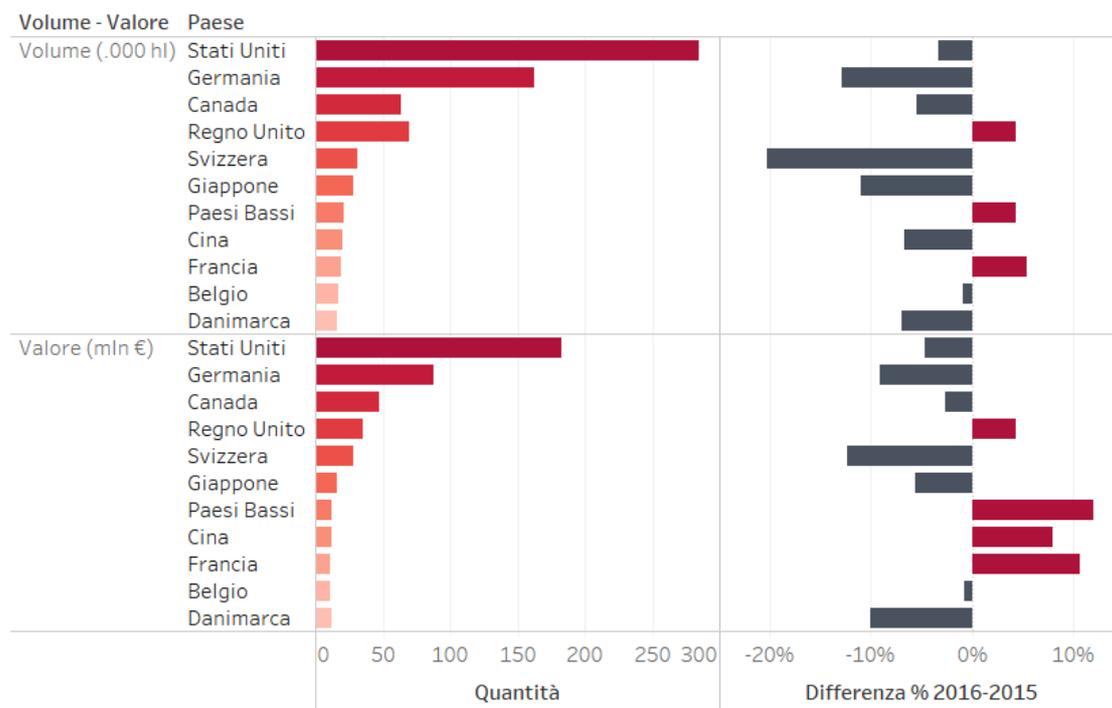
<sup>7</sup> Sono stati esclusi, per esempio, tutti i termini di ricerca relativi a e siti di vendita di vini, quelli focalizzati abbinamenti gastronomici, oppure sulla vendita di kit per la preparazione di Sangiovese a casa.

Figura 6: Esportazioni di vini italiani



Fonte: elaborazione propria su dati Wine Monitor Gennaio-Dicembre 2016

Figura 7: Esportazione di vini rossi toscani DOP



Fonte: elaborazione propria su dati Wine Monitor Gennaio-Dicembre 2016

Dai dati di mercato di *Wine Monitor* (2017) è evidente l'importanza dei vini rossi di Toscana DOP per l'esportazione italiana di vini, secondi fra le varie tipologie di vini per quanto riguarda il valore, e

terzi per volume esportato. Considerando i mercati d'esportazione di vini rossi toscani DOP, emerge chiaramente l'importanza dei paesi di lingua inglese, con tre paesi nelle cui lingue ufficiali rientra l'inglese nei top 5 importatori: Stati Uniti, Canada e Regno Unito.

Tra i paesi di lingua inglese si è deciso inoltre di inserire anche l'Australia per due ordini di motivi. Innanzitutto, come si evince da Figura 2, il volume di ricerche che viene fatto su Google nel paese è relativamente elevato rispetto a quanto avviene nel resto del mondo. In secondo luogo, chi opera nel business vitivinicolo sa che le attività di wine marketing e comunicazione che vengono fatte sia dalle imprese vitivinicole, sia a livello paese sono estremamente "evolute". Pur mostrando negli ultimi anni performance peggiorative in termini di export, l'Australia rappresenta un paese estremamente importante nel panorama del cosiddetto Nuovo Mondo ed è quindi sembrato opportuno esaminare come il vitigno Sangiovese viene raccontato online e a che tipo di informazioni accede un consumatore australiano che faccia ricerche sulla Rete.

È importante precisare che l'indagine è stata condotta tenendo in considerazione il paese all'interno del quale un ipotetico individuo effettua una ricerca online (e specificamente su Google) sul Sangiovese utilizzando delle parole chiave nella lingua ufficiale di quel paese. Attraverso un'impostazione avanzata del motore di ricerca è infatti possibile "simulare" la propria presenza in un certo paese e ricercare attraverso l'uso di keywords le informazioni su un certo argomento. Per chiarire il processo realizzato, se ne elencano di seguito gli step:

- *impostazione navigazione in "incognito"*: questa è una funzione utilizzabile su tutti i principali browser Web e consente di navigare senza lasciare tracce sul computer o su qualsiasi device si stia usando. In questo modo, gli indirizzi dei siti che vengono visitati non risultano elencati

nella cronologia e non vengono memorizzate sul dispositivo impiegato informazioni come i cookie, i suggerimenti per le ricerche o qualunque altro dato inserito nei moduli. La navigazione in incognito si è resa necessaria in quanto, altrimenti, i siti restituiti attraverso le ricerche sarebbero stati distorti, presentando un bias riconducibile alla presenza di cookie legati alle ricerche precedentemente fatte dai ricercatori sul dispositivo utilizzato. In altre parole, questo accorgimento fa sì che le ricerche condotte online per il presente studio riproducano quello che accadrebbe se l'ipotetico individuo che effettua le sue ricerche sul Sangiovese non avesse "una storia pregressa online", ovvero navigasse sempre in modalità incognita. Gli autori del presente studio riconoscono che di fatto anche questo è un "artificio", in quanto la stragrande maggioranza delle persone naviga in modalità "palese", alimenta progressivamente la propria cronologia, acconsente alla registrazione di cookies sul proprio PC, ecc., e dunque ha una sua propria storia sulla Rete che viene considerata dall'algoritmo di Google quando restituisce i risultati. Tuttavia, non potendo accedere a tutte queste informazioni ed essendo tecnicamente impossibile tenere conto delle cronologie dei navigatori di un paese, si è ritenuto questo l'unico modo possibile per raccogliere i dati;

- *impostazione lingua e area geografica*: successivamente, sempre attraverso le impostazioni avanzate di Google, si è proceduto a modificare la lingua di ricerca in base all'idioma di volta in volta ricercato e a modificare l'area geografica, in modo da far sì che il motore di ricerca non considerasse di default la località individuata tramite geolocalizzazione dell'IP che, evidentemente, era in Italia;

- *ricerca della parole chiave, individuazione url e raccolta corpora*: solo dopo aver compiuto i passi precedenti, si è proceduto alla digitazione delle parole chiave e all'individuazione dei siti e delle pagine, con il metodo della saturazione illustrato nella sezione precedente, da cui poi abbiamo raccolto i testi da sottoporre ad analisi.

Per concludere, il processo seguito predilige la prospettiva dell'individuo, potenziale consumatore di Sangiovese, e non quella della fonte delle informazioni. E questo evidentemente perché a noi interessava sapere quali fonti informative avevano la maggiore probabilità di essere consultate da un individuo qualunque localizzato in un certo paese, indipendentemente dalla localizzazione geografico-amministrativa della fonte in questione. In altre parole, un individuo localizzato in Australia che effettua ricerche sul Sangiovese utilizzando la lingua inglese, non necessariamente atterra su pagine con dominio .au (Australia), ma anche .uk, ecc., oppure effettuando ricerche in francese, pur simulando la localizzazione geografica in Francia, si possono consultare risultati che hanno dominio .fr (Francia), .ch (Svizzera), .be (Belgio), ecc.

Questo è del tutto coerente col fatto che, essendo Internet un network mondiale, la localizzazione geografica dei siti che possono essere consultabili non è rilevante ai fini delle ricerche. Quello che conta è di fatto l'ottimizzazione dei siti, il cosiddetto SEO, che se ne fa e che consente che vengano proposti per primi e nelle prime pagine di risultati dai motori di ricerca, in modo da incrementarne la probabilità di consultazione.

## **ELABORAZIONI E ANALISI CONDOTTE SUI CORPORA**

Per ogni lingua e ogni paese sono stati selezionati i corpora disponibili in ogni url, sono stati copiati e incollati in un file Word.

Successivamente, sono stati combinati in un unico file e trasformato in un file unico .txt, formato riconosciuto dal software per procedere al cleaning dei dati, alla preparazione del testo e all'elaborazione dello stesso. L'insieme dei corpora raccolti, così come indicato in ogni scheda paese, è stato analizzato come un unico corpus e sottoposto alle diverse elaborazioni. Prima di procedere alle impostazioni di analisi del corpus, questo è stato sottoposto a un processo di normalizzazione. Questa consiste nell'omogeneizzazione delle grafie utilizzate e viene realizzata per eliminare le possibili fonti di sdoppiamento del dato testuale, per esempio abbassando le maiuscole, uniformando la grafia dei nomi propri, di eventuali sigle, numeri, date, ecc. che potrebbero comportare una ingiustificata variabilità. Il software riconosce le parole presenti nel corpus come sequenze di caratteri dell'alfabeto comprese fra spazi vuoti o fra segni di punteggiatura che sono detti separatori.

### **Impostazioni di analisi dei testi**

Il cleaning e la preparazione dei testi da sottoporre all'analisi hanno richiesto una particolare attenzione e il ricorso a metodi manuali di aggiustamento dei dati per due principali motivi.

1. Innanzitutto per la molteplicità di multi-word contenute nei testi. Le parole multiple sono delle sequenze di due o più parole che costituiscono un'unica unità lessicale. Sono costituite da nomi composti, come "guerra mondiale" e locuzioni usate come modi di dire, come "in grado di". Se il dizionario del software ha consentito un processo automatico di individuazione delle multi-word per parte dei testi analizzati, una buona parte dei corpora è stato sottoposto a una procedura manuale di individuazione e codificazione. Alcuni esempi di multi-word che hanno richiesto una codifica manuale nel processo di preparazione dei dati sono: Sangiovese\_di\_Romagna, varietà\_clonali, Chianti\_Classico, ecc.

2. Il secondo passaggio di aggiustamento manuale ha riguardato una parziale personalizzazione del dizionario utilizzato per l'analisi. In particolare, sono stati realizzati dei processi di fusione tra parole nonostante avessero/non avessero la medesima radice lessicale, ma che presentavano la stessa specificità semantica. Un esempio è quello relativo alla variabilità cromatica del Sangiovese indicata con le parole "viola", "violetto" e "violaceo" che sono stati tutti combinati nel lemma "violaceo", oppure quelle parole che indicavano il "Sangiovese\_di\_Romagna" e "Sangiovese\_romagnolo" che sono stati tutti raggruppati nel lemma "Sangiovese\_di\_Romagna".

### **Le elaborazioni realizzate**

In questa sezione si descrivono sinteticamente le elaborazioni condotte sui corpora che hanno consentito di analizzare i tratti dello storytelling del Sangiovese fatto nelle diverse lingue e frutto delle ricerche realizzate nei diversi paesi.

- **Word association analysis:** in un primo momento è stata effettuata un'analisi delle associazioni di parole al fine di interpretare la keyword principale "Sangiovese" nel contesto semantico in cui è stata prodotta nell'ambito delle fonti testuali raccolte nei diversi paesi. Questo strumento di analisi misura l'intensità dell'associazione tra coppie di parole (unità lessicali) concomitanti (cioè presenti simultaneamente nello stesso frammento di testo, ovvero nei contesti elementari – frasi – in cui i testi sono stati segmentati dal software). La selezione delle parole associate è presentata sia in una tabella con gli indici coseno associativi che in un diagramma radiale, dove il lemma selezionato è posto al centro e le parole associate sono distribuite intorno ad esso, ciascuna proporzionalmente al grado di associazione (Lancia, 2004).

- ***Thematic analysis of elementary contexts***: attraverso una procedura di clustering semantico, finalizzata all'individuazione dei temi emergenti nei testi analizzati, sono stati esaminati i corpora. Il metodo di clustering utilizzato è di tipo non supervisionato (approccio bottom-up) e utilizza un algoritmo bisecting K-Means. Con questa procedura è possibile costruire una rappresentazione dei contenuti dei corpora attraverso un numero limitato di cluster tematici. Ogni cluster è costituito da un insieme di contesti elementari (frasi) caratterizzati dagli stessi pattern di parole chiave ed è descritto attraverso le unità lessicali che caratterizzano maggiormente i contesti elementari di cui è composto. In altre parole, i cluster costituiscono dei sottoinsiemi dei testi caratterizzati da gruppi di parole che condividono gli stessi contesti di riferimento, rappresentando dei macro-temi all'interno dei corpora.
- ***Realizzazione wordcloud***: identificando i testi appartenenti a ciascuno dei cluster individuati, siamo stati in grado di procedere a una content analysis dei contesti elementari inclusi in ogni singolo cluster e di dare una interpretazione qualitativa ai temi individuati. Per facilitare la visualizzazione dei risultati ottenuti, per ogni macro-tema individuato, è stato elaborato un wordcloud che mette in evidenza le parole che caratterizzano maggiormente i testi presi in considerazione. La dimensione delle parole nei wordcloud è proporzionale alla "tipicità" dell'unità lessicale all'interno del cluster, ossia tanto maggiore è la dimensione della parola tanto più quella parola è tipica e caratterizzante del cluster.

## I PAESI DI LINGUA INGLESE

La scelta d'iniziare la presentazione con i paesi di lingua inglese non è soltanto dettata dalla loro importanza per l'export di vini rossi toscani DOP, ma anche dalla possibilità che essi offrono come gruppo linguistico di approfondire alcune caratteristiche che riguardano la diffusione dell'informazione su Google.

Figura 8: Concentrazione del volume medio di ricerche mensili

Termine di Ricerca	Paese			
	US	CA	UK	AU
sangiovese	■ 56,9%	■ 42,8%	■ 38,6%	■ 52,6%
sangiovese wine	■ 16,9%	■ 10,5%	■ 17,5%	■ 12,9%
sangiovese grape	• 1,8%	• 1,6%	• 2,1%	• 0,7%
sangiovese pronunciation	• 2,8%	• 1,3%	• 0,9%	• 0,7%
sangiovese red wine	• 1,0%	• 0,7%	• 2,7%	• 1,1%
sangiovese rose	• 1,0%		• 0,6%	• 2,0%
coriole sangiovese				• 2,4%
sangiovese indianapolis	• 2,3%			
sangiovese merlot		• 1,3%	• 0,6%	
sangiovese ristorante	• 1,8%			
sangiovese taste	• 1,5%	• 0,3%		
sangiovese tesco			• 1,5%	
sangiovese food pairing	• 1,0%			• 0,4%
sangiovese lcbo		• 1,3%		
sangiovese saq		• 1,3%		
sangiovese australia				• 1,1%
sangiovese sassoregale	• 0,8%			
dal zotto sangiovese				• 0,7%
sangiovese chianti		• 0,7%		
sangiovese di romagna			• 0,6%	
sangiovese italian wine			• 0,6%	
	■ 86,8%	■ 61,8%	■ 65,6%	■ 74,6%

Fonte: elaborazione propria su dati Keyword Tool ([keywordtool.io](http://keywordtool.io))

In Figura 8, per ogni paese, viene rappresentata la percentuale relativa alle keyword più utilizzate sul volume totale di ricerche mensili. A titolo di esempio, si osserva che negli USA, il 57% di chi fa ricerche su Google per trovare informazioni sul Sangiovese utilizza esclusivamente la parola Sangiovese. Nel Regno Unito, questa percentuale si ferma a 38,6%, mentre il 17,5% ricerca "Sangiovese

wine". Questo ha evidentemente importanza per gli operatori allorquando vogliono generare traffico sui propri siti, inserendo nei testi le parole chiave maggiormente utilizzate dai navigatori online.

## **ANALISI DEI TERMINI DI RICERCA**

Il volume di ricerca relativo al concetto di "Sangiovese" tende a concentrarsi attorno a pochi termini, specialmente "Sangiovese wine", con code più o meno lunghe a seconda del paese. Tranne che per termini "locali", come Tesco nel Regno Unito o LCBO e SAQ in Canada legati alla distribuzione di prodotti nel territorio, c'è un'alta sovrapposizione tra le parole ricercate nei diversi paesi.

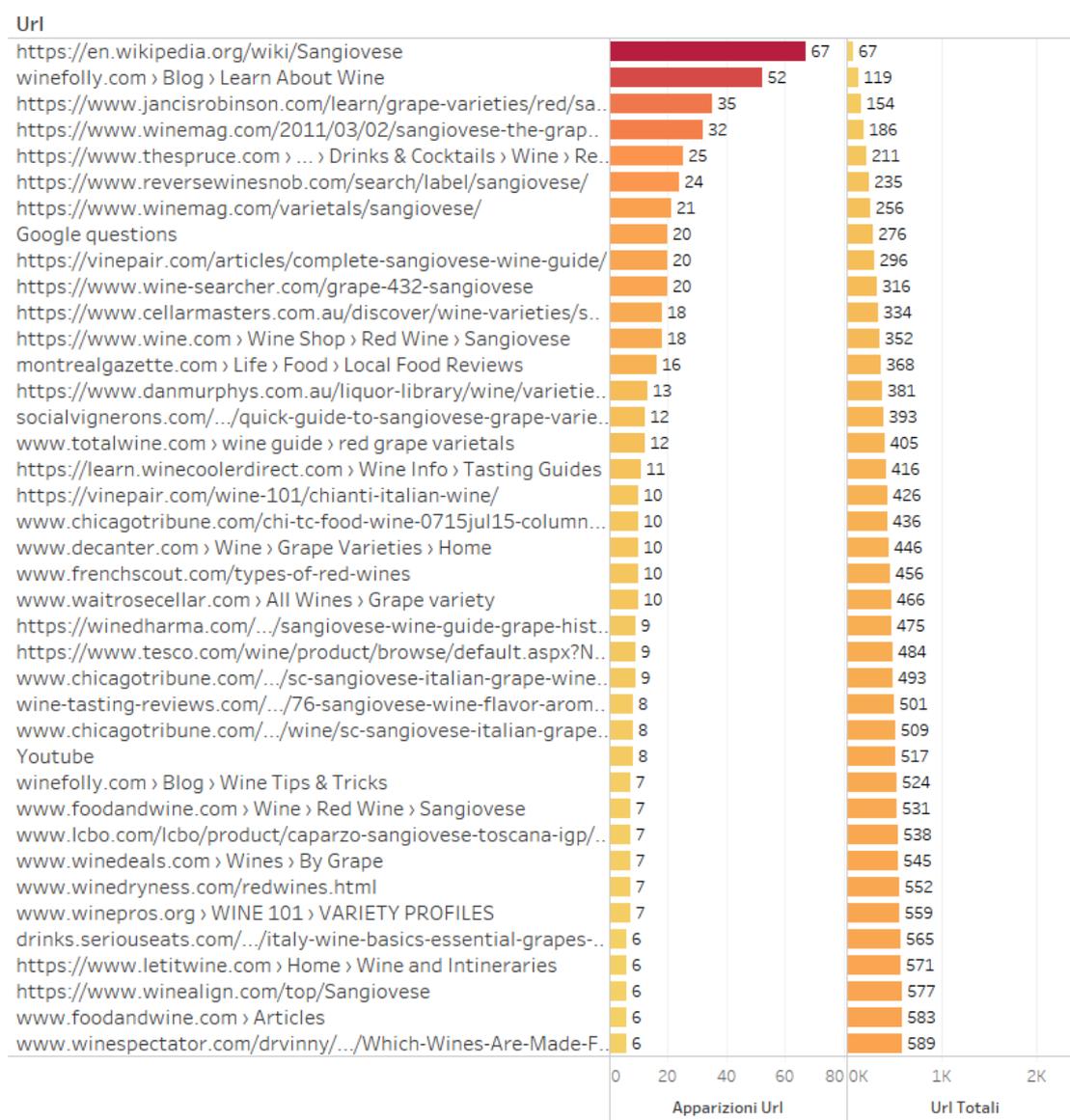
### **Analisi degli URL**

Prima di procedere alla descrizione delle ulteriori analisi effettuate, preme evidenziare che – come già accennato – nel processo di selezione degli url si è deciso di *non considerare* quelli che riguardavano esclusivamente proposte commerciali o siti istituzionali di aziende come per esempio quelle della ristorazione. È tuttavia opportuno segnalare che il volume di ricerca attraverso keyword specifiche finalizzate all'acquisto di bottiglie o all'individuazione di ristoranti non è del tutto trascurabile (si veda per esempio il 2,3% delle ricerche realizzate in USA è fatto con le parole chiave "Sangiovese Indianapolis" oppure l'1,8% delle ricerche in UK è fatto con le keyword "Sangiovese Tesco"). In questi casi, essendo l'algoritmo di Google estremamente selettivo ed efficiente, propone siti nei cui contenuti non c'è traccia di informazioni descrittive o di narrazione del Sangiovese, ma solo di natura prettamente commerciale.

Quando ognuno di noi effettua una ricerca online su un motore di ricerca, ottiene una lista di risultati composti da un titolo e da

descrizione sottostante. La descrizione è chiamata snippet. Lo snippet è rilevante per l'ottimizzazione sui motori di ricerca e ha una doppia funzione: la prima è quella di contenere parole chiave che contribuiscano a far sì che il sito sia visibile ai primi posti dei risultati proposti (SEO), la seconda consiste nell'essere sufficientemente descrittivo e accattivante per persuadere il navigatore a cliccare sul link.

Figura 9: Concentrazione url - gruppo linguistico inglese



Fonte: elaborazione propria su dati Google

Per comprendere se esistessero dei siti posizionati più efficacemente di altri, ovvero che fossero riproposti da Google nelle prime pagine nonostante l'utilizzo di diverse parole chiave che contenevano comunque la parola Sangiovese, si è proceduto nel seguente modo. Sulla base della maggior frequenza di utilizzo, sono state selezionate 117 keyword. Per ogni keyword è stata effettuata una ricerca su Google, sono stati selezionati i primi 20 link proposti e sono stati analizzati i relativi snippet. Complessivamente sono stati dunque esaminati 2.340 snippet. Dall'analisi è emerso che il 25% (corrispondenti a 589 snippet) dei risultati è riconducibile a 39 url (si veda Figura 9).

Figura 10: Codifica tematica degli snippet

Termine di Ricerca	Codifica Snippet		
	Uva & Vino	Vendita	Altro
sangiovese	■ 17	■ 2	■ 1
sangiovese wine	■ 17	■ 3	
sangiovese grape	■ 20		
sangiovese pronunciation	■ 5		■ 15
sangiovese red wine	■ 14	■ 6	
sangiovese rose	■ 1	■ 10	■ 7
coriole sangiovese	■ 4	■ 14	■ 2
sangiovese indianapolis			■ 20
sangiovese merlot	■ 9	■ 2	■ 8
sangiovese ristorante			■ 20
sangiovese taste	■ 19	■ 2	■ 1
sangiovese tesco	■ 5	■ 15	
sangiovese food pairing	■ 2		■ 19
sangiovese lcbo	■ 7	■ 13	
sangiovese saq		■ 20	
sangiovese australia	■ 10	■ 8	■ 2
sangiovese sassoregale	■ 5	■ 12	■ 2
dal zotto sangiovese	■ 6	■ 13	■ 1
sangiovese chianti	■ 16	■ 4	
sangiovese di romagna	■ 6	■ 10	■ 3
sangiovese italian wine	■ 17	■ 5	
	■ 180	■ 139	■ 101

Fonte: elaborazione propria

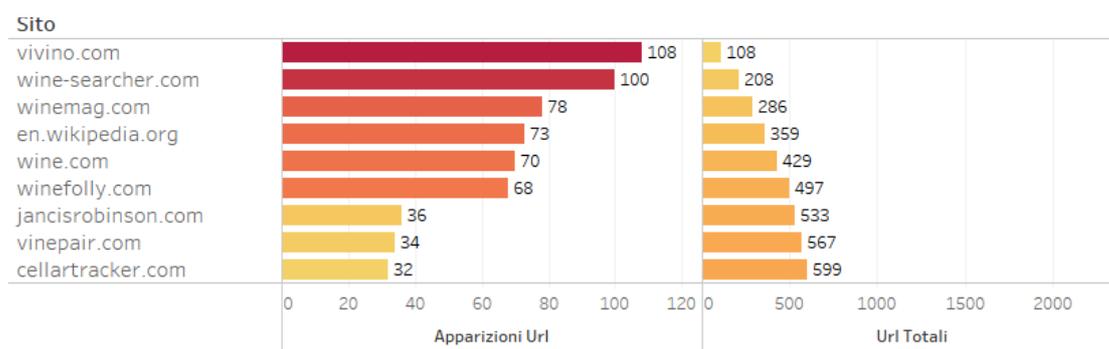
Questo significa che, indipendentemente dalle parole chiave utilizzate, esistono alcune fonti informative che ricorrono più spesso, sono dunque posizionate in modo particolarmente efficace e in grado di "presidiare" meglio di altri l'informazione.

Figura 10 illustra la codifica tematica degli snippet e, considerando le keyword selezionate e i relativi primi 20 snippet esaminati, mette

in evidenza la numerosità degli stessi che è riconducibile a diversi temi. Per esempio, utilizzando la keyword "Sangiovese", 17 snippet fanno riferimento a temi relativi a uva e vino, più ricchi di storytelling, mentre 2 a proposte commerciali e 1 ad altri argomenti. Mentre, nel caso in cui la parola chiave utilizzata sia "Sangiovese grape", i primi 20 snippet sono tutti riconducibili ai temi uva e vino.

Questa dinamica di "concentrazione" delle fonti informative è valida a maggior ragione se si considerano i siti. I 589 snippet/url, che corrispondono al 25% del totale di quelli consultati, è infatti riconducibile a soli 9 siti. E questo avviene perché un sito può avere più pagine ben posizionate sul motore di ricerca e incrementare di conseguenza la probabilità che il navigatore vi atterri.

Figura 11: Concentrazioni siti - gruppo linguistico inglese



Fonte: elaborazione propria su dati Google

Il risultato ottenuto è interessante anche considerando che le ricerche sono state fatte simulando la geolocalizzazione dell'utente in diversi paesi. 3 sono in cima alla classifica nel nostro database globale (Figura 9).

Figura 12: Url comuni - gruppo linguistico inglese

Url Comuni	Paese				
	USA	CA	UK	AU	
Url comuni ai 4 paesi	4	4	4	4	4
Url comuni a 3 paesi	6	8	8	2	
Url comuni a 2 paesi	3	7	5	1	
Url singole	7	1	3	13	
Totale	20	20	20	20	

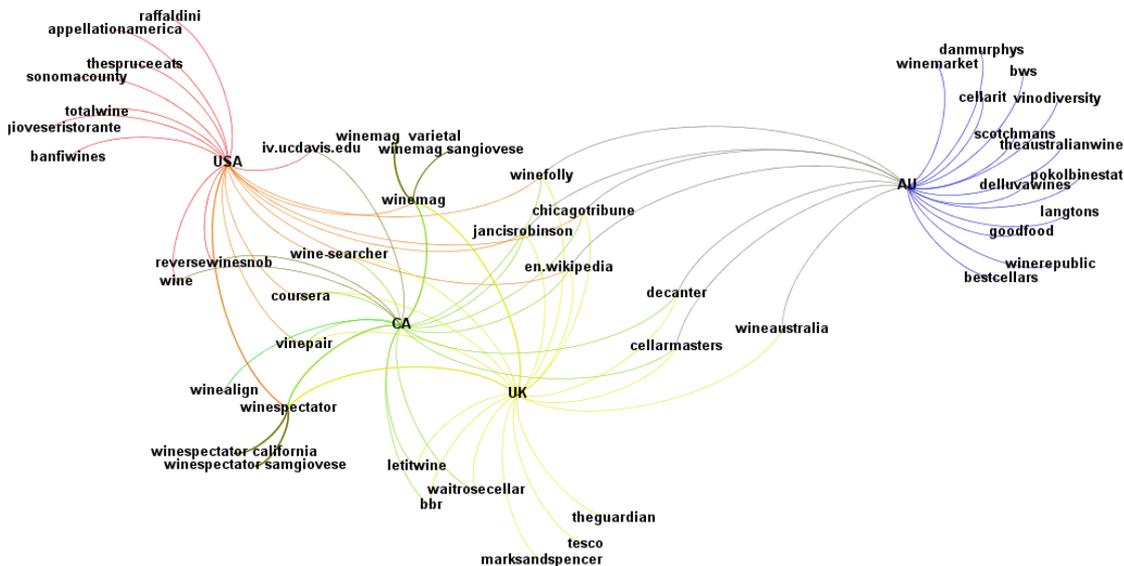
Fonte: elaborazione propria su dati Google

Figura 13: Matrice degli url comuni

Paese	USA	CA	UK	AU
USA	7	13	10	4
CA		1	16	6
UK			3	7
AU				13

Fonte: elaborazione propria su dati Google

Figura 14: Collegamenti e distanza fra url e paesi



Fonte: elaborazione propria su dati Google

Nel caso particolare della nostra ricerca, per esempio, il 20% degli url restituiti da Google per il termine Sangiovese con filtri specifici per paese e lingua sono comuni a tutti e 4 paesi considerati e sono quelli che risultano più ricorrenti tra gli url raccolti.

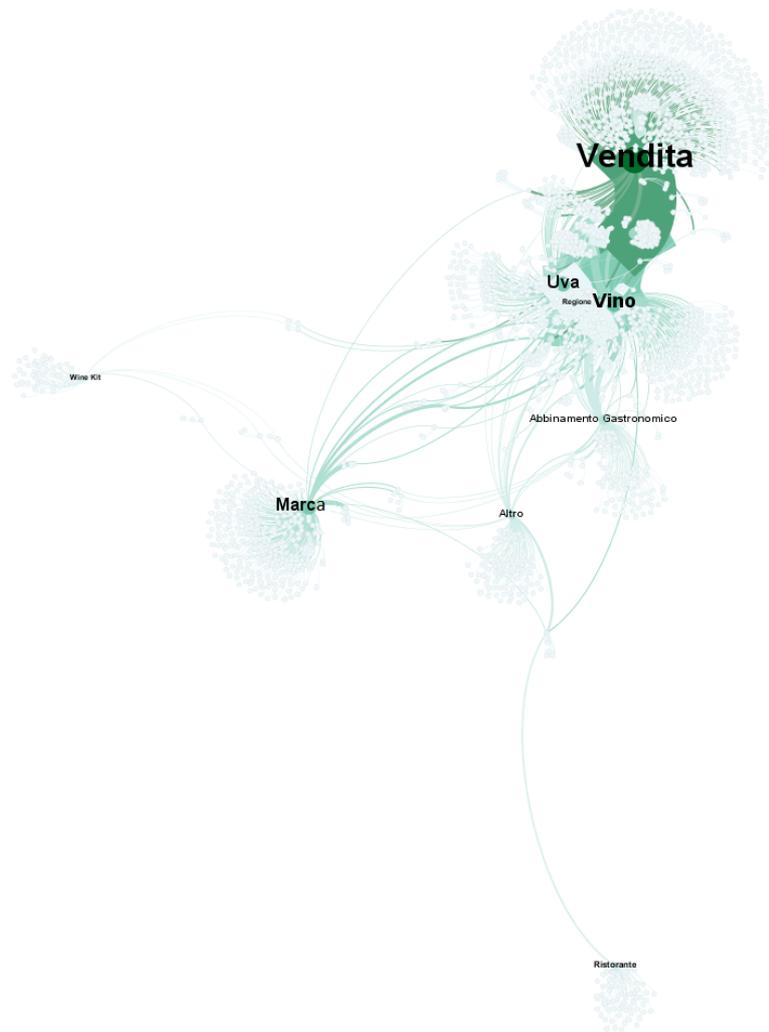
Al contempo, non è forse un caso che i paesi più diversi fra loro siano entrambi paesi produttori, gli Stati Uniti e l'Australia, e che quello con la maggiore proporzione di url singoli sia proprio l'Australia, non solo geograficamente dislocata rispetto agli altri, ma forse il paese più intenzionato, fuori dalla patria del Sangiovese, a produrre vini di qualità dal vitigno.

### **Analisi tematica degli snippet**

Sulla base dell'analisi tematica dei contenuti degli snippet sono stati categorizzati nel seguente modo:

- *Vendita*: attività commerciale relativa a vini a base di Sangiovese;
- *Vino*: contenuti in cui si parla di vini a base di Sangiovese;
- *Uva*: riferimenti al vitigno e alle sue caratteristiche;
- *Regione*: contenuti in cui il Sangiovese viene messo in relazione con un territorio;
- *Marca*: associazioni tra Sangiovese, brand e relativo posizionamento;
- *Abbinamento gastronomico*: suggerimenti su piatti e cibi che possono essere accompagnati da Sangiovese;
- *Ristorante*: riferimenti a specifiche aziende di ristorazione;
- *Wine kit*: promozione di kit (insieme di ingredienti e strumenti) per la produzione di vini "a base di Sangiovese" da realizzare a casa propria;
- *Altro*: categoria residuale in cui sono stati raggruppati altri contenuti meno ricorrenti.

Figura 15: Relazioni fra le categorie degli snippet



Fonte: elaborazione propria su dati Google

Dall'analisi degli snippet emerge che quando si aggiungono le keyword che sono nella coda della distribuzione del volume medio di ricerche mensili, i contenuti più comuni restituiti da Google sono in larga parte costituiti da brevi proposte di vendita. Per quanto riguarda lo storytelling del Sangiovese, è interessante notare la vicinanza fra i contenuti degli snippet relativi al vitigno, ai vini a base di Sangiovese, e alle regioni ad esso associate.

## ANALISI DEI CORPORA

Al fine di facilitare l'illustrazione dei risultati dell'indagine, si ritiene opportuno fornire alcune sintetiche informazioni relative all'analisi dei corpora raccolti simulando di essere geolocalizzati nei singoli paesi e di procedere poi con una disamina comparata dei risultati ottenuti.

Per ogni paese, viene proposta una tabella in cui si riportano le caratteristiche generali dei corpora esaminati:

- *Url selezionati e corpora*: è il risultato del processo di selezione delle fonti di testo da sottoporre a elaborazione e riporta il numero degli url e dei corrispondenti documenti di testo che sono stati estrapolati;
- *Occorrenze (token)*: numero di parole complessivo presenti nel testo;
- *Forme (type)*: numero di unità lessicali (intese come stringhe di caratteri riconosciuti dal software separati da spazi vuoti);
- *Hapax*: numero di parole che hanno occorrenza pari a 1, ovvero compaiono una sola volta nel testo esaminato;
- *Lemmi*: numero di label o tag frutto del processo di lemmatizzazione, ovvero di raggruppamento e classificazione delle unità lessicali;
- *Frequenza minima nel testo*: valore soglia di occorrenze nel testo al di sotto del quale i lemmi sono stati esclusi dall'analisi;
- *Lemmi sottoposti ad analisi*: numero di lemmi inclusi nell'analisi con occorrenza pari o superiore al valore soglia prescelto;
- *Contesti elementari*: numero di frammenti di testo derivanti dal processo di segmentazione del corpus che facilitano l'esplorazione dei documenti.

## Australia

Tabella 1: Caratteristiche corpus lingua inglese (Australia)

CARATTERISTICHE DEL CORPUS	NUMERO
Url selezionati e corpora	30
Occorrenze (token)	12.531
Forme (type)	2.842
Hapax (occ. = 1)	1.514
Lemmi	2.439
Frequenza minima nel testo	$\geq 4$
Lemmi sottoposti ad analisi	368
Contesti elementari	288

## Canada

Tabella 2: Caratteristiche corpus lingua inglese (Canada)

CARATTERISTICHE DEL CORPUS	NUMERO
Url selezionati e corpora	20
Occorrenze (token)	14.217
Forme (type)	3.141
Hapax (occ. = 1)	1.764
Lemmi	2.670
Frequenza minima nel testo	$\geq 4$
Lemmi sottoposti ad analisi	375
Contesti elementari	322

## Regno Unito

Tabella 3: Caratteristiche corpus lingua inglese (Regno Unito)

CARATTERISTICHE DEL CORPUS	NUMERO
Url selezionati e corpora	24
Occorrenze (token)	13.023
Forme (type)	3.097
Hapax (occ. = 1)	1.835
Lemmi	2.647
Frequenza minima nel testo	$\geq 4$
Lemmi sottoposti ad analisi	358
Contesti elementari	293

## USA

Tabella 4: Caratteristiche corpus lingua inglese (USA)

CARATTERISTICHE DEL CORPUS	NUMERO
Url selezionati e corpora	21
Occorrenze (token)	13.257
Forme (type)	3.080
Hapax (occ. = 1)	1.768
Lemmi	2.627
Frequenza minima nel testo	>=4
Lemmi sottoposti ad analisi	365
Contesti elementari	300

Fonte delle Tabelle: elaborazione propria

## ASSOCIAZIONI DI PAROLE: SANGIOVESE

Come detto, la prima analisi condotta consiste nel calcolo delle associazioni di parole. In particolare, si è deciso di esplorare le co-occorrenze all'interno dei corpora dei diversi paesi relative alla parola chiave principale e che compare il maggior numero di volte in tutti i testi, ossia Sangiovese.

Attraverso un indice, nel nostro caso quello del coseno, il software verifica le relazioni di co-occorrenza della parola selezionata all'interno dei contesti elementari. L'indice del coseno ha la seguente formula:

$$\frac{a}{\sqrt{(a+b) \times (a+c)}}$$

dove:

- $a$  è il numero di volte in cui i lemmi considerati (ipotizziamo lemma 1 e lemma 2) co-occorrono in un contesto elementare,

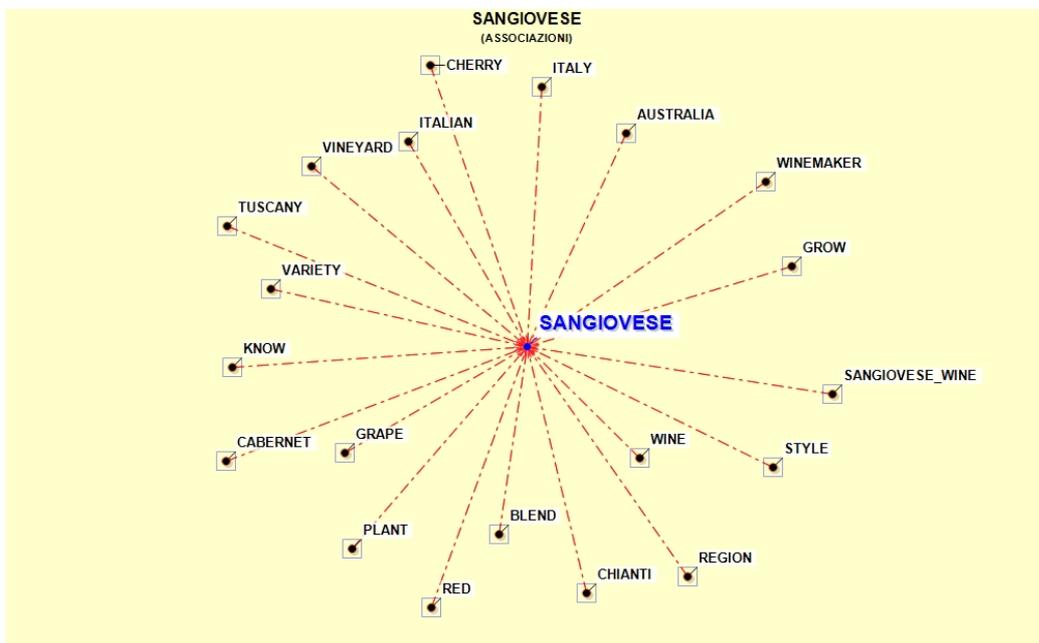
- $b$  è il numero di volte in cui è presente il lemma 1 e non il lemma 2,
- $c$  è il numero di volte in cui è presente il lemma 2 ma non il lemma 1.

Evidentemente, maggiore è l'indice, maggiore è il grado di associazione tra le due parole considerate.

In parole semplici, nel nostro caso, questo metodo verifica la compresenza della parola Sangiovese e altri lemmi presenti nel testo e riporta, in un diagramma radiale, le parole che co-occorrono con questa con maggiore frequenza. Nel diagramma, più vicini sono i lemmi alla parola centrale, maggiore è la frequenza di co-occorrenza.

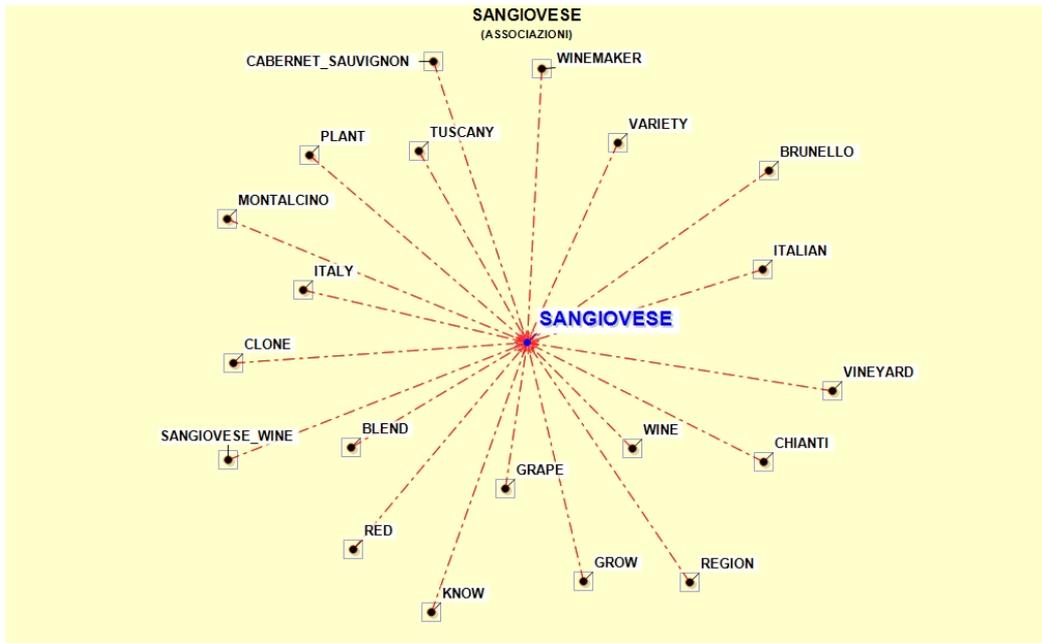
Di seguito, si riportano i diagrammi radiali relativi ai corpora dei diversi paesi.

Figura 16: Diagramma associazioni Sangiovese Australia



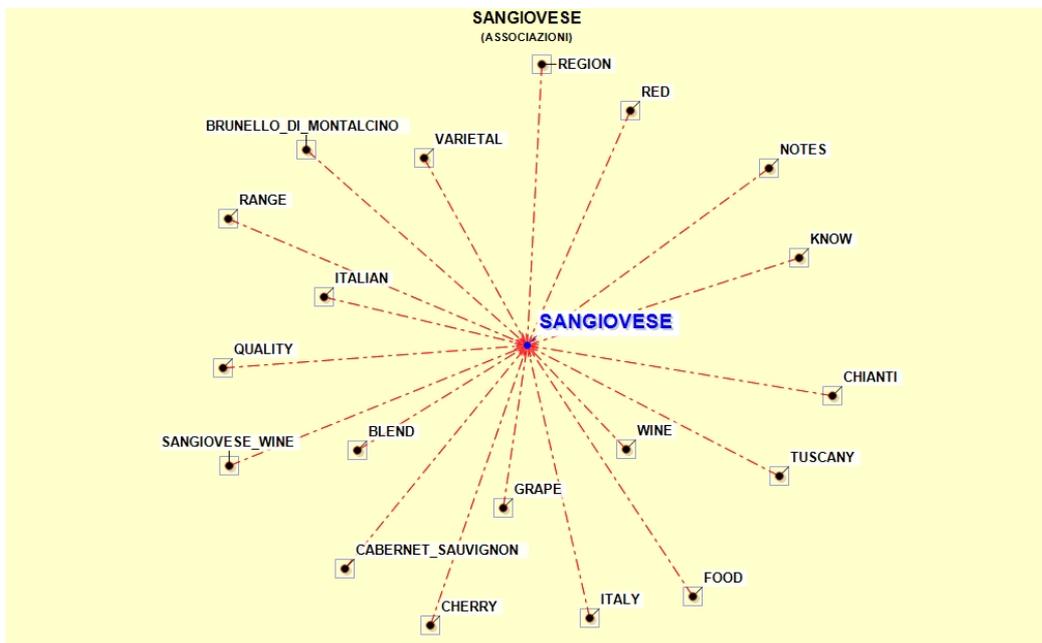
Fonte: elaborazione propria

Figura 17: Diagramma associazioni Sangiovese Canada



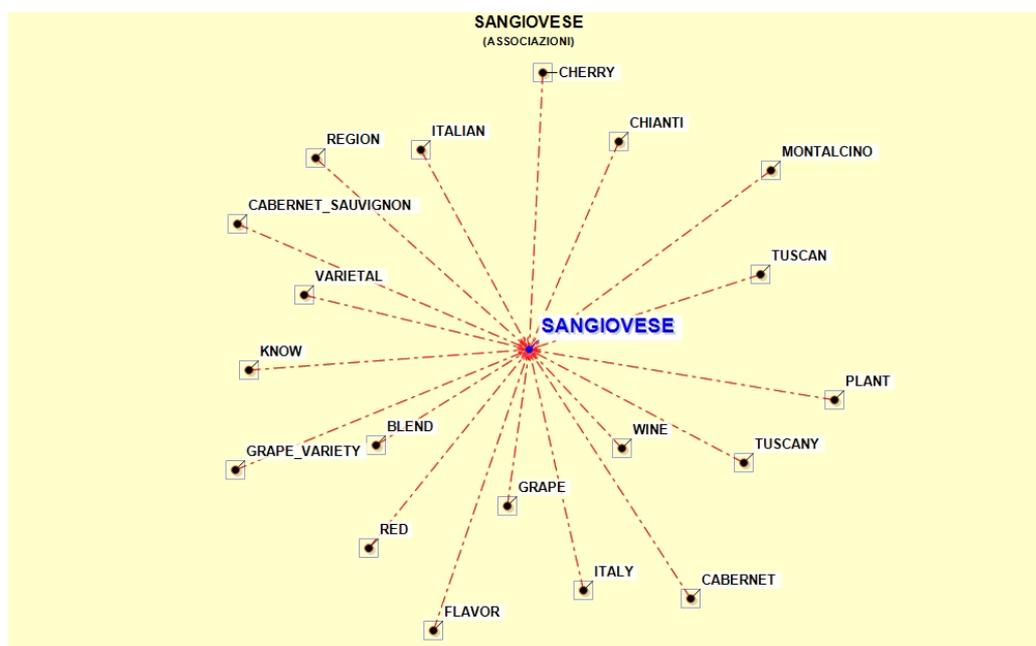
Fonte: elaborazione propria

Figura 18: Diagramma associazioni Sangiovese Regno Unito



Fonte: elaborazione propria

Figura 19: Diagramma associazioni Sangiovese USA



Fonte: elaborazione propria

Per fornire una lettura comparata, in Tabella 5, si riportano i dati dei primi 20 lemmi associati a Sangiovese con maggiore frequenza dei diversi paesi. In particolare:

- *Lemma A*: per ogni paese considerato, questo lemma è la parola Sangiovese;
- *Lemma B*: per ogni paese, nella prima colonna, viene riportato il lemma che co-occorre con Sangiovese nel testo con maggior frequenza in ordine decrescente;
- *C.E.(B)*: la seconda colonna indica il numero di contesti elementari in cui è presente il lemma B corrispondente. A titolo esemplificativo, il termine wine in Australia è presente complessivamente in 106 frammenti di testo;
- *C.E.(AB)*: la terza colonna indica il numero di contesti elementari in cui sono presenti sia il lemma selezionato (Sangiovese), sia la parola associata. Riprendendo l'esempio, le parole Sangiovese e wine in Australia sono presenti in 66 contesti elementari.

Tabella 5: Lemmi e associazioni per paese

AUSTRALIA			CANADA			REGNO UNITO			USA		
LEMMA B	C.E.(B)	C.E.(AB)	LEMMA B	C.E.(B)	C.E.(AB)	LEMMA B	C.E.(B)	C.E.(AB)	LEMMA B	C.E.(B)	C.E.(AB)
wine	106	66	wine	129	85	wine	123	79	wine	136	82
blend	38	31	grape	105	74	grape	100	60	grape	105	60
grape	70	42	blend	43	37	blend	37	32	blend	42	35
variety	36	25	Italy	54	38	Italy	44	34	varietal	46	29
Italian	28	22	Tuscany	47	34	varietal	46	32	Italian	31	23
Australia	21	19	variety	42	32	red	32	22	Chianti	53	30
grow	26	20	Italian	32	27	know	26	19	Tuscan	22	19
style	35	23	Chianti	50	31	Tuscany	42	24	Tuscany	52	29
Chianti	47	26	grow	28	23	Italy	46	24	Italy	48	26
plant	28	20	red	34	24	Cab_Sauv	18	15	red	35	22
know	26	19	clone	32	23	quality	29	19	know	23	17
vineyard	24	18	plant	31	22	Brunello	43	23	region	43	22
Italy	36	22	winemaker	24	19	region	43	23	cherry	23	16
winemaker	25	18	Brunello	49	27	notes	14	13	Montalcino	19	14
Sang_wine	35	21	vineyard	27	20	Chianti	44	23	plant	39	20
region	42	23	region	47	26	food	15	13	Cabernet	12	11
red	21	16	know	23	18	cherry	20	15	flavor	26	16
Cabernet	15	13	Sang_wine	43	24	Sang_wine	33	19	Cab_Sauv	15	12
Tuscany	36	20	Cab_Sauv	22	17	range	21	15	grape_var	15	12
cherry	16	13	Montalcino	22	17	winemaker	19	14	California	27	16

Fonte: elaborazione propria

Una prima considerazione riguarda la sostanziale omogeneità dei termini che co-occorrono con Sangiovese nei testi<sup>8</sup>. In particolare, le prime 3 parole, seppur con un ordine diverso in Australia, sono esattamente le medesime: wine, grape e blend.

Lasciati con sfondo bianco in Tabella 5, si nota come gran parte dei termini sia riferita agli aspetti che concernono la produzione, in particolare all'ambito agricolo (vineyard, plant, clone, region, grow).

È poi possibile individuare una categoria di termini (in azzurro) che fanno riferimento alle caratteristiche organolettiche (red, cherry, flavor, notes). E una relativa all'Italia (in verde) e in particolare alla Toscana (in giallo), relativamente alla quale vengono associati alla parola Sangiovese anche le denominazioni più famose legate al vitigno ossia Chianti e Brunello. A questo proposito, preme evidenziare fin da subito che non risultano tra le prime 20 associazioni con Sangiovese alcuni termini che hanno a che fare con la Romagna che vedremo essere invece molto più menzionata nei testi in italiano.

<sup>8</sup> Si rammenta, a questo proposito, che come detto nei paragrafi precedenti, indipendentemente dal paese in cui si è geolocalizzata la ricerca, spesso le fonti informative coincidono. Non sorprende dunque che, a causa della parziale sovrapposizione dei testi esaminati, questi risultino in parte omogenei.

Un'ultima osservazione riguarda l'Australia e gli USA (California) in cui il Sangiovese viene con più frequenza associato alle produzioni locali (in arancione). Vedremo che questo aspetto è particolarmente significativo nei testi australiani in cui una maggior parte dell'informazione è finalizzata a descrivere il Sangiovese impiantato nel paese e il processo di selezione clonale che ha condotto a un "ottimale" adattamento del vitigno alle condizioni pedoclimatiche locali, nonché allo "stile" australiano (blending e caratteristiche organolettiche dei vini a base di Sangiovese prodotti nel paese).

## **ANALISI DEI TEMI EMERGENTI**

Come descritto nella sezione metodologica (Le elaborazioni realizzate), i corpora raccolti sono stati sottoposti a una analisi tematica dei contesti elementari<sup>9</sup>. In sostanza, questa metodologia consente di costruire ed esplorare una rappresentazione dei contenuti dei testi attraverso pochi e significativi *cluster tematici*.

Ognuno di questi è costituito da un insieme di contesti elementari caratterizzati da analoghi pattern di parole chiave ed è descritto attraverso le unità lessicali che maggiormente caratterizzano i contesti elementari di cui è composto. È possibile affermare che il risultato di questa elaborazione propone una mappatura delle cosiddette *isotopie*, ovvero i temi, più o meno specifici, caratterizzati dalla co-occorrenza di tratti semantici. Ogni cluster ottenuto attraverso l'analisi si caratterizza per insiemi di unità lessicali che condividono gli stessi contesti di riferimento, consentendo di ricostruire discorsi specifici all'interno della struttura complessiva del corpus.

---

<sup>9</sup> La fonte della descrizione della metodologia riportata nel testo è il manuale del software T-Lab (Lancia, 2014).

Sempre come detto, il clustering realizzato in questo caso è di tipo *non supervisionato*. Questo significa che non si sono identificate delle categorie semantiche a priori, ma si è adottato una procedura che fa sì che i temi “emergano” dal testo. Solo successivamente, esaminando i cluster e i contesti elementari in essi contenuti attraverso la content analysis qualitativa, sono stati codificati dei macro-temi contenuti nei corpora.

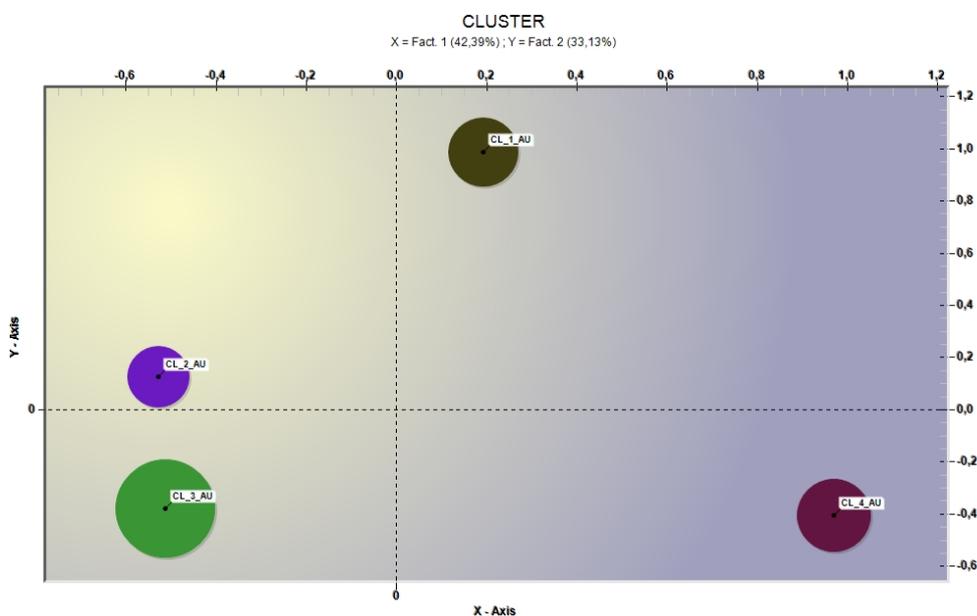
Per ogni paese, vengono riportati i risultati ottenuti attraverso l’elaborazione dei testi con il software e i wordcloud da noi generati utilizzando le unità lessicali contenute nei diversi cluster.

I wordcloud sono costruiti includendo nella “nuvola” le parole maggiormente significative che appartengono al cluster. Come detto, la dimensione delle parole nei wordcloud è proporzionale alla “tipicità” dell’unità lessicale all’interno del cluster, ossia tanto maggiore è la dimensione della parola tanto più quella parola è tipica e caratterizzante del cluster. Ogni cluster riporta inoltre una percentuale che rappresenta la porzione di contesti elementari dell’intero corpus che sono riconducibili a quello specifico cluster. Questo significa che tanto maggiore è la percentuale, tanto più preponderante è quel macro-tema rispetto agli altri.

Si premette fin d’ora che pur essendo alcuni temi trasversali e trattati in maniera omogenea dei testi raccolti nei diversi paesi, nella maggior parte dei casi è possibile individuare delle specificità che riguardano sia i temi stessi, sia il “tone of voice” con cui vengono illustrati. Preme evidenziare altresì che, in generale, l’approccio con cui vengono affrontati i temi è decisamente “scientifico”. Ovvero, le argomentazioni proposte fanno in generale riferimento all’agronomia, all’ampelografia, alla vitivinicoltura e all’enologia, alle caratteristiche fisico-chimiche e organolettiche dei vini ottenuti dal Sangiovese. Molto meno di frequente si fanno riferimenti alla storia, alle tradizioni,

a contenuti emozionali che potrebbero essere invece associati a una connotazione "mitologica" del Sangue di Giove.

Figura 20: Cluster Australia



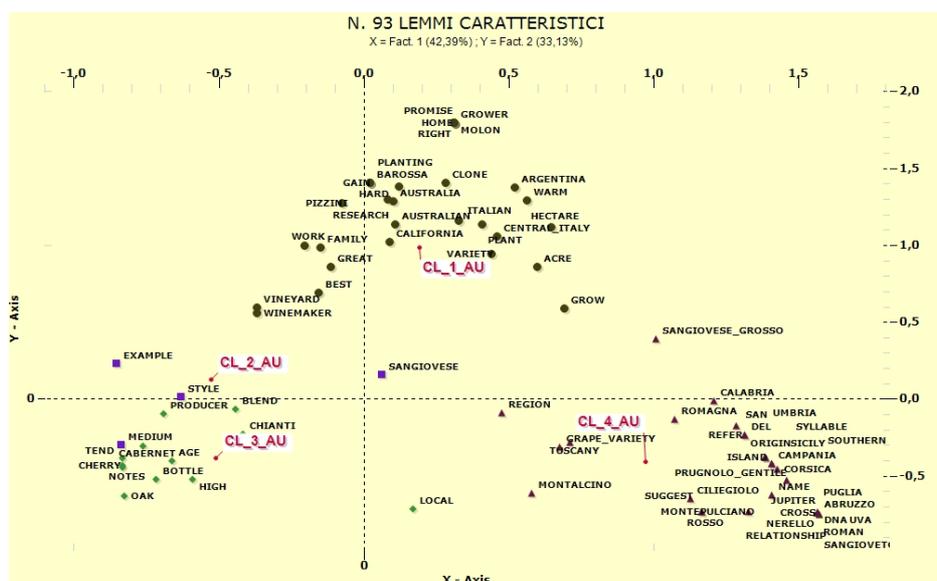
Fonte: elaborazione propria

### **Cluster semantica: Australia**

La procedura di analisi dei temi emergenti simulando la geolocalizzazione del navigatore in Australia ha fornito una soluzione a 4 cluster. Questo significa che i contesti elementari contenuti nel corpus esaminato sono riconducibili a 4 macro-temi principali.

Figura 20 mostra il risultato dell'analisi e la posizione dei cluster considerando i due assi fattoriali x e y (dimensioni) che ne determinano la variabilità. La visualizzazione del risultato può includere anche i lemmi che appartengono a ciascun cluster come mostrato in Figura 21.

Figura 21: Cluster e lemmi Australia

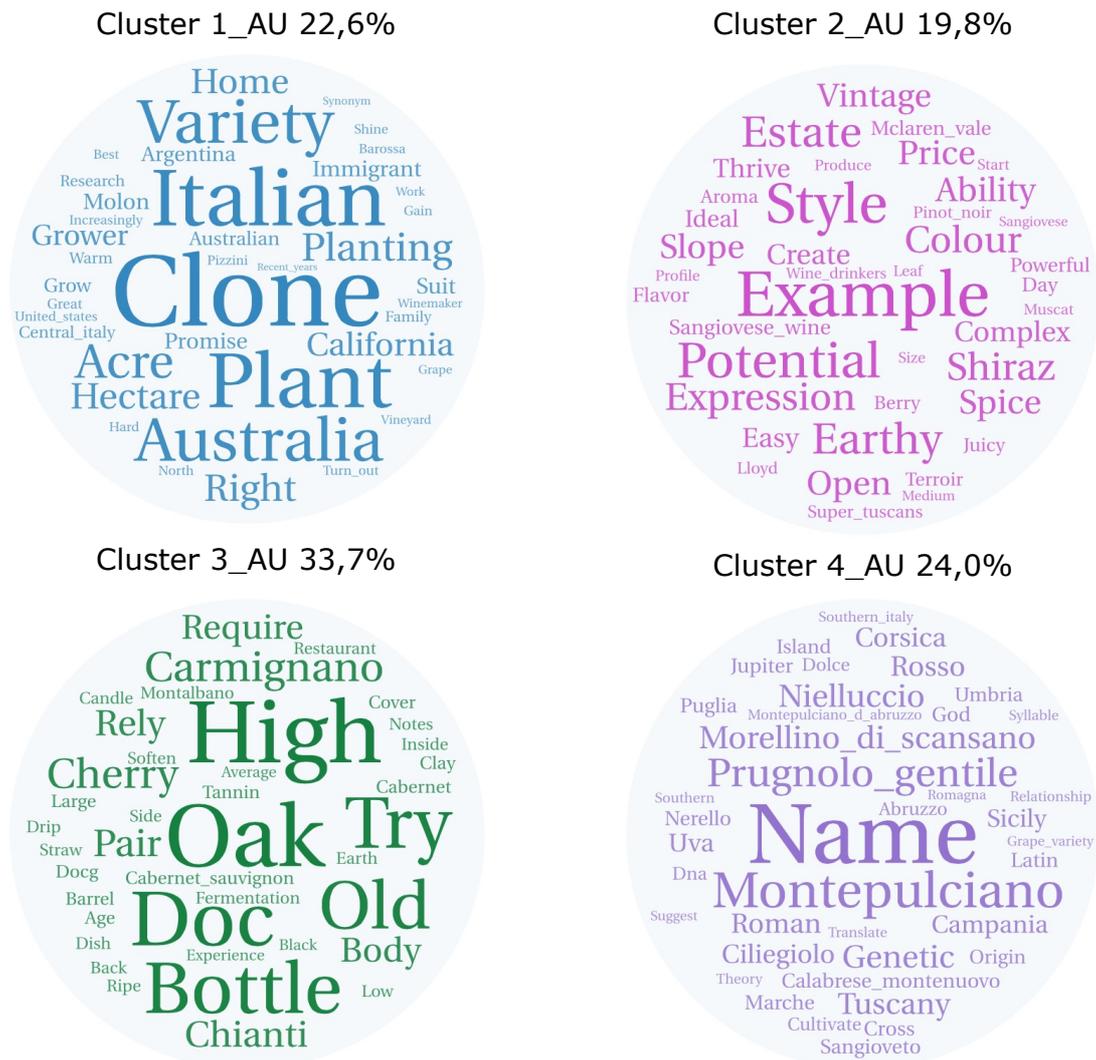


Fonte: elaborazione propria

L’Australia è probabilmente il paese che si caratterizza più degli altri nel proporre una narrazione del Sangiovese sì, ma di quello coltivato all’interno dei propri confini. Oltre il 40% dei contesti elementari classificati (Cluster 1\_AU – 22,6% e Cluster 2\_AU – 19,8%) riguarda infatti il Sangiovese coltivato in Australia. Il Cluster 1\_AU, in particolare, racconta dell’abilità degli imprenditori australiani di selezionare cloni (individuati grazie all’attività di ricerca condotta dall’Università di UC Davis, California, USA) capaci di adattarsi alle condizioni pedoclimatiche del territorio. Pur ammettendo alcune prime sperimentazioni fallimentari, si evidenzia come il vitigno si sia successivamente ben adattato e sia stato in grado di fornire la materia prima adatta per la realizzazione di vini di qualità elevata. In questo caso, il taglio con cui si racconta il Sangiovese è di tipo “agronomico”.

Contrariamente a quanto avviene in tutti gli altri paesi, in Australia viene fatto ricorrente riferimento a nomi di imprenditori e di aziende produttrici che sono stati capaci di valorizzare il vitigno a livello locale come Pizzini, Penfolds, Indigo, la famiglia Lloyd, ecc.

Figura 22: Wordcloud cluster Australia



Fonte: elaborazione propria

Box 1: Verbatim cluster 1\_AU e Cluster 2\_AU

**Cluster 1\_AU**

\*\*\*\* \*00001  
SCORE ( 92.349 )  
Italian immigrants introduced the Sangiovese vine to Argentina in the late 19th and early 20th century. Early site and clonal\_selection was less\_than ideal and, like California and Australia, recent endeavors have focused on finding the best clones to use and the right vineyard locations. The grape is not widely planted in Argentina and the focus is mostly on the export market.

\*\*\*\* \*00001  
SCORE ( 46.024 )  
Australian Sangiovese time to shine It has taken a\_great\_deal longer than anybody thought it would, but slowly, Sangiovese is beginning to shine. Not only have the best clones been identified, but winemaker and growers are also

working\_out the **best** regions for the **variety** as\_well. In the **Barossa**, there's Penfolds' Cellar Reserve Sangiovese.

### Cluster 2\_AU

\*\*\*\* \*00001

SCORE ( 33.539 )

In the early **days**, **Lloyd** added Cabernet and **Shiraz** for **colour** and depth, but felt it didn't show the true **expression** of the grape. Three decades of experience in the vineyard and winery has given him the **ability** to **produce** a 100% **Sangiovese** of immense quality and **style** that stands completely on its own, and Coriole continues to top the **Sangiovese** charts today.

\*\*\*\* \*00001

SCORE ( 25.994 )

Brokenwood sources fruit from the Indigo Vineyard, which also makes its own **fine Sangiovese**. Castagna is a biodynamic producer, **creating** truly dynamic wines **É** the La Chiave is a rich and **powerful expression** of the **style**. In Heathcote, outfits like Tar & Roses, Condie **Estate** and Greenstone Vineyards have\_been working hard with their **Sangiovese** with exceptional results.

*Fonte: elaborazione propria*

Box 1 riporta alcuni dei verbatim maggiormente significativi dei Cluster 1\_AU e 2\_AU con il relativo score. Questo punteggio rappresenta il "peso" che il singolo contesto elementare ha nell'ambito del cluster. Il software riporta la lista dei verbatim in ordine di score decrescente, ovvero dal contesto elementare che più caratterizza il cluster in oggetto a quello che lo caratterizza meno. Le parole in rosso rappresentano le unità lessicali che sono maggiormente tipiche del cluster.

Come detto, anche il Cluster 2\_AU riguarda il Sangiovese in Australia ma da un punto di vista "enologico". Questo macro-tema infatti riguarda la capacità dei winemaker australiani di "interpretare" il vitigno in modo da realizzare vini di qualità. In questo caso, l'accento è posto sullo stile dei vini, sui possibili blend realizzati con varietà locali e sulla capacità di riuscire a riprodurre vini che siano simili a quelli italiani, in particolare toscani, e non una mera versione australiana ritenuta troppo somigliante ai prodotti del luogo, come per esempio i vini a base di Shiraz. Anche in questo caso, gran parte della narrazione fa riferimento a una lunga e faticosa attività di sperimentazione, in questo caso riferita alla vinificazione e

all'invecchiamento, che solo dopo una molteplicità di tentativi di dubbia riuscita conduce alla realizzazione di vini di qualità "straordinaria".

Gli altri due cluster (Cluster 3\_AU - 33,7% e Cluster 4\_AU - 24,0%) riguardano invece il Sangiovese italiano. In particolare, il Cluster 3\_AU tratta dello "stile" dei vini rossi toscani a base di Sangiovese. È interessante evidenziare fin d'ora che, analogamente a quanto avviene in tutti gli altri paesi di lingua anglofona esaminati, il Sangiovese di Romagna o comunque riferimenti alla produzione romagnola a base di Sangiovese sono assolutamente marginali. Le potenzialità del vitigno di dare origine a vini di elevata qualità riguardano nella stragrande maggioranza dei casi gli ettari impiantati nelle diverse aree della Toscana e che fanno parte delle denominazioni più famose come il Chianti, il Chianti Classico, il Brunello di Montalcino e il Carmignano. In Australia, in particolare nel Cluster 3\_AU, sono numerosissimi i riferimenti a queste indicazioni geografiche messe in relazione con le caratteristiche fisico-chimiche e organolettiche dei vini.

*Box 2: Verbatim Cluster 3\_AU e Cluster 4\_AU*

#### **Cluster 3\_AU**

\*\*\*\* \*00001

SCORE ( 14.052 )

**CHIANTI DOCG Chianti** is the **largest** and best-known Sangiovese-dominant **wine\_region**, but its styles are diverse. Though it **covers** about one-third of Tuscany, overlapping several other famous Sangiovese-based regions, most of **Chianti** is **located** toward the northern **side** of Tuscany. as\_a\_result, it **tends** to exhibit more of Sangiovese's characteristic **acidity** and less of its **body**.

\*\*\*\* \*00001

SCORE ( 13.015 )

**Carmignano** is generally made from Sangiovese\_planted on the east **side** of the **Montalbano** hills, which has less sandstone and produces a more intense **version** of the grape. Therefore, while **Carmignano tends** to be drinkable like **Chianti**, it also has darker **fruit** and more structure. **Carmignano wines** must be **aged** for approximately 18 months, with eight months\_in\_oak or chestnut.

#### Cluster 4\_AU

\*\*\*\* \*00001

SCORE ( 30.790 )

In 2004, DNA profiling done by researchers at San Michele all'Adige revealed the grape to be the product of a crossing between Ciliegiolo and Calabrese\_Montenuovo. While Ciliegiolo has a long history tied to the Tuscan region, Calabrese\_Montenuovo ( which is not related to the grape commonly known as Calabrese, or Nero d'Avola ) has its origins in southern\_Italy

\*\*\*\* \*00001

SCORE ( 20.690 )

Evidence for this latter theory is the proliferation of seedless mutations of Sangiovese, known under various synonyms, throughout various regions of southern\_Italy including Campania, Corinto nero which is grown on the island of Lipari just north of Sicily and Tuccanese from the Puglia region in the heel of the Italian boot.

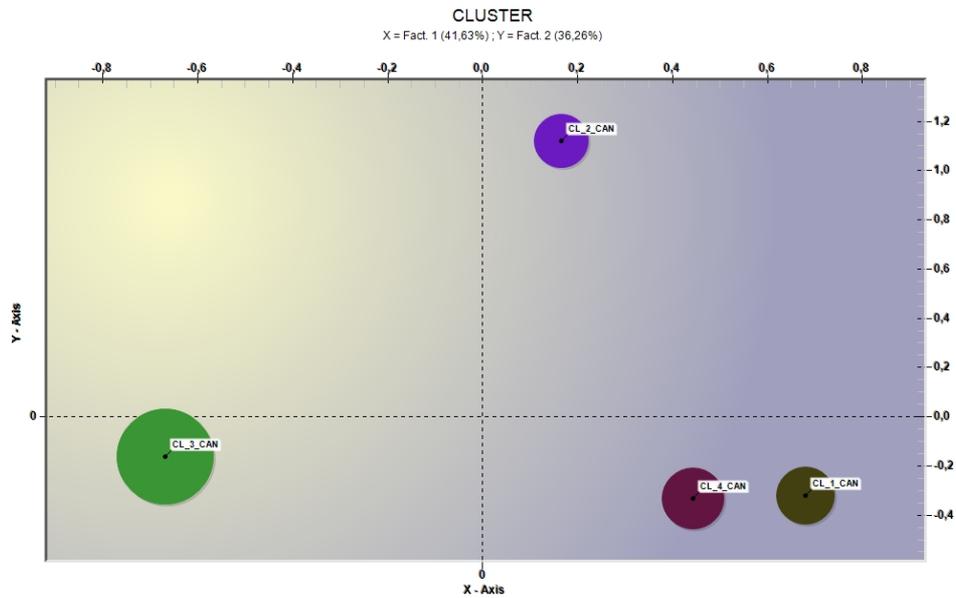
*Fonte: elaborazione propria*

L'ultimo cluster, il 4\_AU, riguarda le origini del Sangiovese, il suo patrimonio genetico e i diversi nomi che questo assume nelle diverse aree in cui è coltivato. Anche in questo caso, le argomentazioni sono molto tecniche e fanno spesso riferimenti a ricerche scientifiche condotte sul DNA del vitigno e alle teorie sulla sua origine e l'evoluzione nel tempo. Si segnala fin d'ora che questo macro-tema si ritrova in tutti i corpora dei paesi di lingua inglese esaminati ed ha quindi caratteristiche assai simili ai corrispondenti cluster che illustreremo relativamente a Canada, Regno Unito e USA. Un tratto che è forse maggiormente enfatizzato in Australia concerne il legame che il Sangiovese potrebbe avere con vitigni del sud Italia, avvalorato da riferimenti a ricerche che dimostrano come il suo patrimonio genetico potrebbe avere tratti in comune con il Calabrese Montenuovo (Calabria, Campania), il Tuccanese (Puglia), il Susumaniello (Salento), il Corinto nero (Lipari), ecc.

#### **Cluster semantica: Canada**

Il medesimo processo di elaborazione è stato realizzato sul corpus raccolto in Canada e, anche in questo caso, ha condotto all'individuazione di 4 cluster semantici (Figura 23).

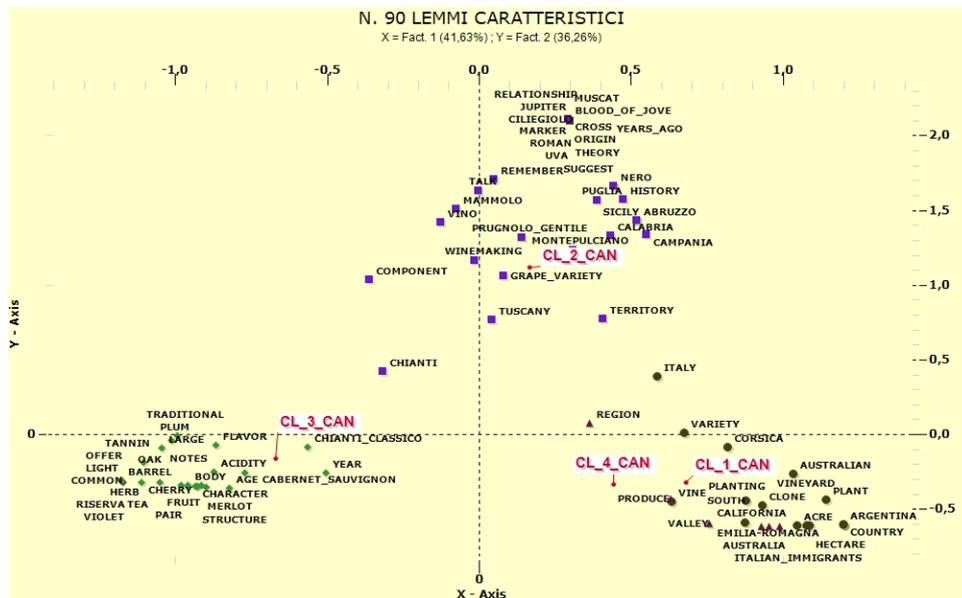
Figura 23: Cluster Canada



Fonte: elaborazione propria

I cluster con i rispettivi lemmi caratteristici sono rappresentati in Figura 24.

Figura 24: Cluster e lemmi Canada



Fonte: elaborazione propria

Anche simulando la geolocalizzazione della ricerca online in Canada si identificano due cluster (Cluster\_1\_CAN - 20,9% e Cluster\_4\_CAN - 22,7%) che riguardano temi legati al vitigno del

Sangiovese, alle condizioni pedoclimatiche che ne favoriscono la crescita e aumentano il potenziale per la realizzazione di vini di qualità. La differenza tra il primo e il quarto cluster risiede nel fatto che parte della narrazione del Cluster\_4\_CAN riguarda il Sangiovese in Italia. Quasi il 45% dei contesti elementari è dunque riconducibile a discorsi che parlano del vitigno adottando un'ottica agronomica. Anche in questo caso, molta della narrazione riguarda le capacità del vitigno di adattarsi a contesti ambientali al di fuori dei confini italiani. A differenza dei testi australiani, tuttavia, i riferimenti sono fatti più in generale al cosiddetto Nuovo Mondo del vino, come Stati Uniti (Stati di Washington, California), all'Argentina, al Cile, alla Nuova Zelanda, allo stesso Canada (Ontario, British Columbia, ecc.). Potremmo dunque dire che il lettore di questi testi viene informato in maniera più trasversale e di fatto più *oggettiva* sulla coltivazione del Sangiovese all'estero e delle aree in cui ha dato i risultati migliori (Box 3).

Box 3: Verbatim Cluster\_1\_CAN e Cluster\_4\_CAN

#### Cluster\_1\_CAN

\*\*\*\* \*00001

SCORE ( 130.093 )

[ 7 ] **Argentina** was next with 6, 928 **acres** ( 2, 804 ha ), followed by Romania with 1, 700 **hectares** ( 4, 200 **acres** ), the **Corsica** region in France with 1, 663 **hectares** ( 4, 110 **acres** ), **California** with 1, 371 **hectares** ( 3, 390 **acres** ) and **Australia** with 440 **hectares** ( 1, 100 **acres** ). [ 10 ] **Italy** In **Italy**, Sangiovese is the most widely planted\_red\_grape **variety**.

\*\*\*\* \*00001

SCORE ( 36.072 )

Canberra and Young in New **South** Wales; Stanthorpe in Queensland and the western edge of the **Great** Dividing Range in Victoria. [ 10 ] In New Zealand, the first varietal version of Sangiovese was released in 1998 and today there are 6 **hectares** ( 15 **acres** ) of the **grape planted**, mostly on the **North Island** around Auckland.

#### Cluster\_4\_CAN

\*\*\*\* \*00001

SCORE ( 9.676 )

Ha of **Sangiovese\_planted**, mostly in Ontario where some **producers** in Niagara-on-the-Lake are **experimenting** with ice wine versions of the grape. A small amount of the grape can also be found in British Columbia. [ 7 ] Other New

World regions A "Super Tuscan-style" Sangiovese\_blend from Washington\_State.

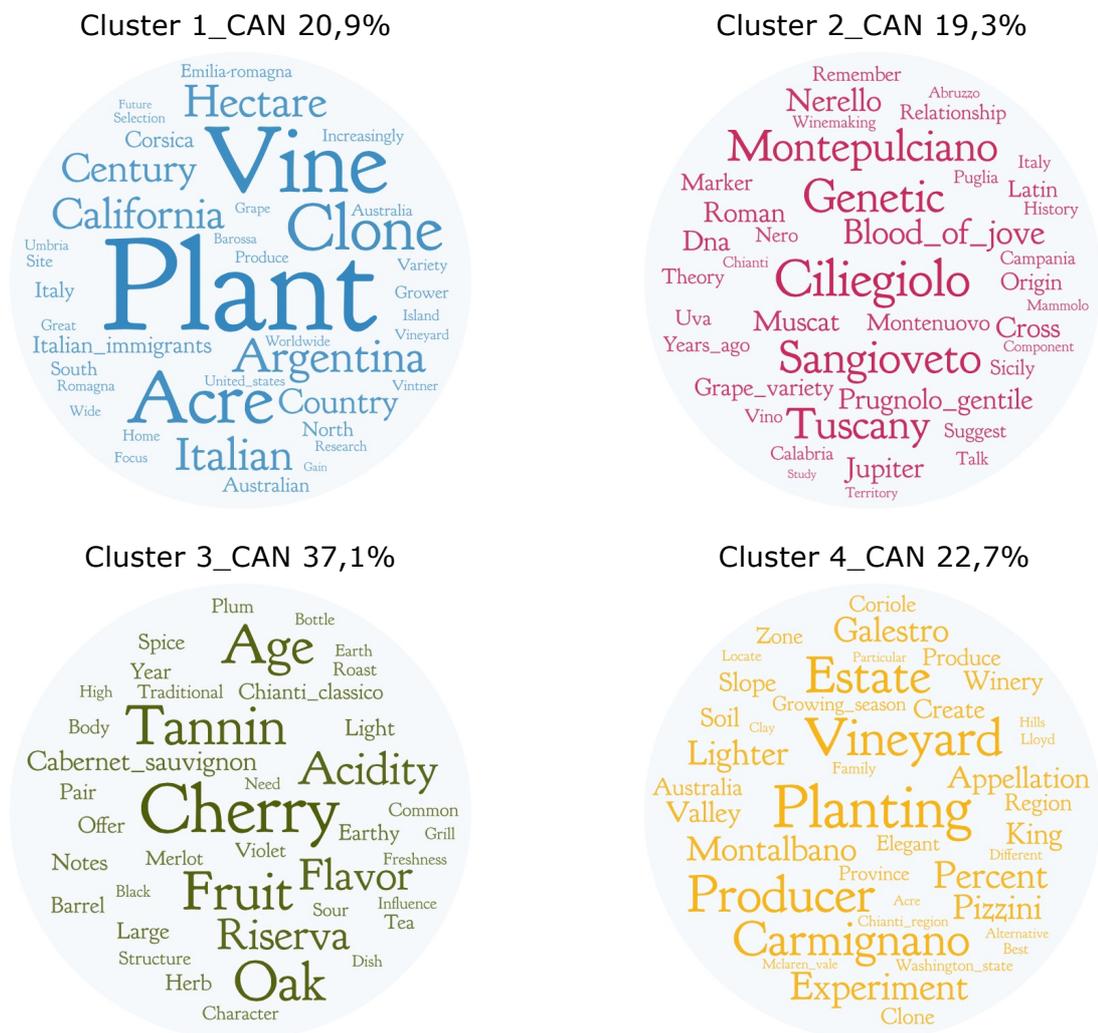
\*\*\*\* \*00001

SCORE ( 9.151 )

By 2003, that number rose to nearly 3, 000 acres ( 1, 200 ha ) with plantings across the state, most notably in Napa Valley, Sonoma county, San Luis Obispo, Santa Barbara and the Sierra Foothills. [ 4 ] However, in recent years plantings of the variety have declined to 1, 950 acres ( 790 ha ) by 2010.

Fonte: elaborazione propria

Figura 25: Wordcloud cluster Canada



Fonte: elaborazione propria

Il Cluster\_3\_CAN ha dimensioni quasi doppie rispetto agli altri, includendo oltre il 37% dei contesti elementari classificati.

#### Box 4: Verbatim Cluster\_3\_CAN

\*\*\*\* \*00001

SCORE ( 67.910 )

Four **years** for normale and five **years** for **riserva**, with two **years** in **oak** for both categories. Brunello **wines** still have a great variation in style, particularly when it comes to **oak aging**. **Modern versions aged** in new French\_oak can be rich, toasty and concentrated, while **traditional versions aged** in **large**, old **oak** often have more of an emphasis on non-fruit **flavors** and **earth**.

\*\*\*\* \*00001

SCORE ( 33.282 )

**Morellino** is more plush and juicy, the **acidity** softened by **dark**, drinkable **cherry fruit**. Despite being the most southerly Tuscan **DOCG** based on Sangiovese, **Morellino wines** have complex tasting **notes** beyond rich **fruit**. They maintain **freshness** and are suitable for youthful consumption, as evidenced by regulations that allow **Morellino** to be **released** a mere six months after **harvest**.

Fonte: elaborazione propria

Come risulta in maniera evidente dal wordcloud, il macro-tema in questo caso sono le caratteristiche fisico-chimiche e organolettiche dei vini a base di Sangiovese. In particolare, considerando le diverse denominazioni dei vini realizzati col vitigno (Brunello, Morellino, Chianti, Super Tuscan inclusi) e i diversi processi di invecchiamento e affinamento cui sono sottoposti, così come eventuali altri uvaggi utilizzati per la loro produzione (Cabernet Sauvignon, Merlot, Syrah, ecc.), si evidenzia la versatilità del Sangiovese e le differenze che caratterizzano i prodotti finiti.

Riferimenti frequenti sono fatti alla presenza di tannini, all'acidità, alla freschezza, così come alle note erbacee e speziate, di terra, di ciliegia, al colore violaceo. Solo saltuariamente, si fa riferimento ai possibili abbinamenti cibo-vino che vedremo essere invece addirittura un macro-tema che risulta dalla procedura di clustering realizzata per il Regno Unito.

Si è deciso di illustrare il Cluster\_2\_CAN per ultimo perché, considerando tutti i macro-temi emersi nella procedura di clustering dei paesi anglofoni, è probabilmente quello che più enfatizza gli aspetti "mitologici" del Sangiovese, facendo riferimento a storie e

tradizioni, anche utilizzando un tono maggiormente diretto e colloquiale. In altre parole è il cluster che più degli altri utilizza una narrazione maggiormente emozionale e si distacca, almeno in parte, dal tono "scientifico" di natura agronomica ed enologica diffusamente utilizzato in altri contesti. Come si evince dai verbatim nel Box 5, diversi contesti elementari parlano dell'etimologia del termine Sangiovese e fanno riferimento agli antichi Romani e agli Etruschi. Qui si fa riferimento alla *vitis vinifera*, specie coltivata da millenni, già dagli Assiri, i Babilonesi, i Sumeri e gli Egizi, all'etimologia della parola che si rifà a Zeus e alla mitologia greca, non senza qualche banalizzazione come "Giove, il dio numero uno dei Romani". Seppure più "ordinario", il tono con cui si veicolano le informazioni appare senza dubbio più fruibile da parte del grande pubblico, meno edotto da un punto di vista scientifico, e probabilmente più disponibile a leggere testi meno impegnativi e complessi.

*Box 5: Verbatim Cluster\_2\_CAN*

\*\*\*\* \*00001

SCORE ( 46.761 )

Early theories on the origin of Sangiovese dated the grape to the time of Roman winemaking. [ 4 ] It was even postulated that the grape was first cultivated in Tuscany by the Etruscans from wild *Vitis vinifera* vines. The literal translation of the grape's name, the "blood\_of\_Jove", refers to the Roman god Jupiter.

\*\*\*\* \*00001

SCORE ( 26.699 )

Over 40 years on, the variety is finally beginning to shine, and adventurous wine\_drinkers are discovering a whole new world of Sangiovese's exotic, evocative flavour and flair... Australian style. From the pretty hills of Tuscany Sangiovese is the famous grape of Tuscany. The name is derived from the Latin Sanguis Jovis blood\_of\_Jove after the mythical Roman God of Jupiter.

\*\*\*\* \*00001

SCORE ( 17.382 )

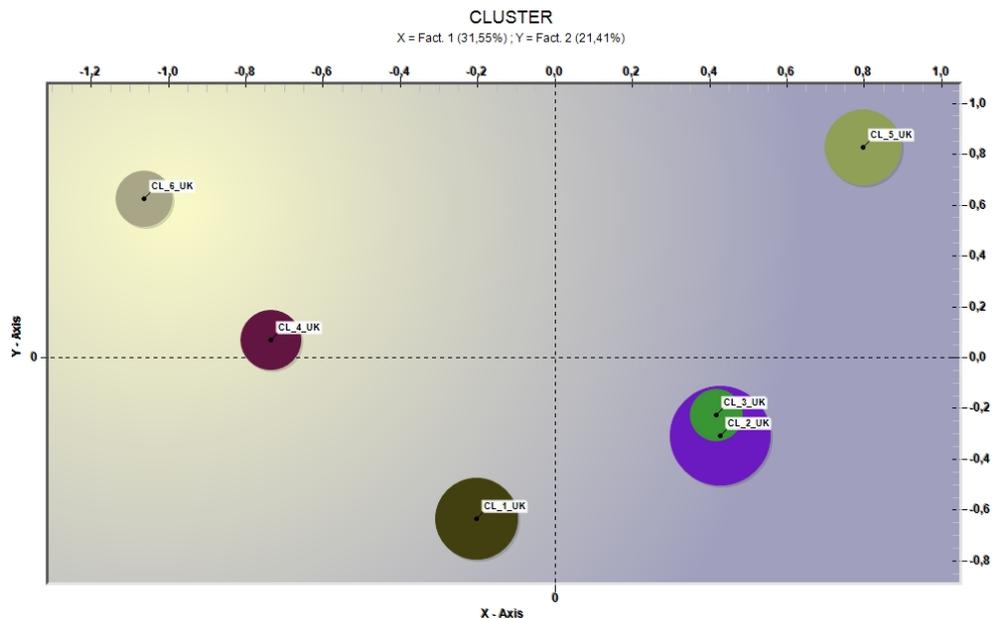
You may remember him as Jupiter, the Romans' numero uno god ( pardon me, my Latin is rusty ) and their answer to the Greeks' Zeus. You may also remember sangiovese as the wheelhouse, principal grape\_variety of Chianti, that once-forgettable red\_wine.

*Fonte: elaborazione propria*

## Cluster semantica: Regno Unito

Il processo di clustering del corpus relativo al Regno Unito ha dato origine a una soluzione a 6 cluster, evidenziando una maggiore variabilità dei macro-temi presenti nei testi.

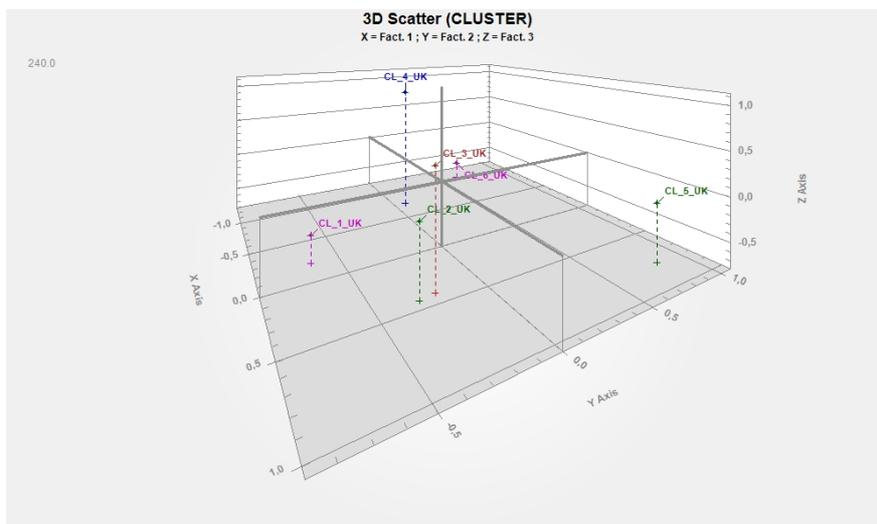
Figura 26: Cluster Regno Unito 2D



Fonte: elaborazione propria

Per meglio comprendere le differenze che esistono tra i cluster, poiché nella rappresentazione bidimensionale (Figura 26) i Cluster\_2\_UK e Cluster\_3\_UK appaiono sovrapposti, in Figura 27, mostriamo la rappresentazione tridimensionale.

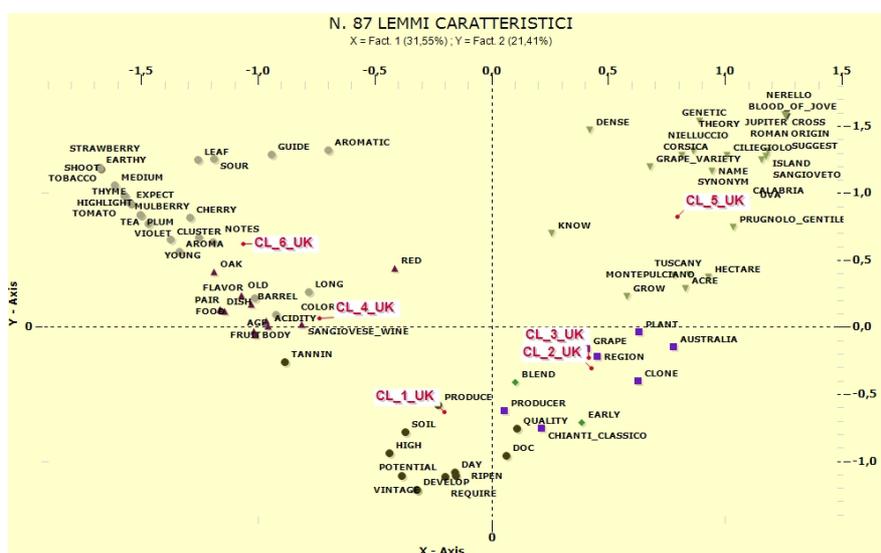
Figura 27: Cluster Regno Unito 3D



Fonte: elaborazione propria

In Figura 28, viene mostrata la rappresentazione dei cluster e dei lemmi relativi al Regno Unito.

Figura 28: Cluster e lemmi Regno Unito



Fonte: elaborazione propria



Tre cluster su sei, specificamente il Cluster\_1\_UK – 19,8%, il Cluster\_2\_UK – 24,2% e il Cluster\_3\_UK – 11,3%, propongono una narrazione del Sangiovese secondo l'ormai nota prospettiva agronomica e vitivinicola. Complessivamente dunque oltre il 55% dei contesti elementari del corpus Regno Unito è accomunato da un taglio prettamente tecnico.

*Box 6: Verbatim Cluster\_1\_UK, Cluster\_2\_UK, Cluster\_3\_UK*

#### **Cluster\_1\_UK**

\*\*\*\* \*00001

SCORE ( 42.874 )

The grape seems to **produce** the **highest quality wine** in the sandstone and clay rich hills south of the Via Emilia near the Apennines which is covered by much of the Sangiovese\_di\_Romagna **DOC** zone. The **higher** summer **time** temperatures of this **area** gives more opportunity for Sangiovese to sufficiently **ripen**.

\*\*\*\* \*00001

SCORE ( 37.186 )

Sangiovese has shown itself to be adaptable to many different types of vineyard **soils** but seems to thrive in **soils** with a **high** concentration of limestone, having the **potential** to **produce** elegant **wines** with forceful aromas. In the Chianti\_Classico region, Sangiovese thrives on the highly friable shale-clay **soil** known as **galestro**.

#### **Cluster\_2\_UK**

\*\*\*\* \*00001

SCORE ( 22.062 )

The former is well-known as an ancient **variety** in Tuscany, the latter is an almost-extinct relic from the Calabria, the toe of Italy. [ 5 ] At least fourteen Sangiovese **clones** exist, of which Brunello is one of the best regarded. An attempt to classify the **clones** into **Sangiovese\_grosso** ( including Brunello ) and **Sangiovese\_Piccolo** families has **gained** little evidential support.

\*\*\*\* \*00001

SCORE ( 13.072 )

Owing\_to Sangiovese's ability to take\_on characteristics of **region**, climate and those imparted by the **winemaker**, wines made from this **grape** vary widely in taste. CSIRO **experimented** with the Sangiovese **varietal** in **Australia** in the 1960s. Since then there has\_been an ongoing search for most suitable climates and **regions** to get the best out\_of the **grape**.

#### **Cluster\_3\_UK**

\*\*\*\* \*00001

SCORE ( 49.893 )

While the first DOC to be **permitted** to **blend Cabernet\_Sauvignon** with Sangiovese was approved for **Carmignano** in 1975, most of Tuscany's premier

wine\_regions were not permitted to blend Cabernet\_Sauvignon with Sangiovese till the late 20th century.

\*\*\*\* \*00001

SCORE ( 38.807 )

Newer Italian imports, yet to be evaluated under California conditions, include Sangiovese FPS 07 ( VCR 6 ), 08 ( VCR 19 ), 09 ( VCR 30 ), 10 ( VCR 23 ), and 13 ( VCR 102 ). Typically, production is 5 to 10 tons per acre.

*Fonte: elaborazione propria*

In particolare, il Cluster\_1\_UK è dominato da riferimenti al territorio, alle caratteristiche del suolo che favoriscono lo sviluppo del potenziale del vitigno. Si descrive in maniera dettagliata l'ambiente di coltivazione in Italia, facendo riferimento sia alla Toscana che all'Emilia Romagna, le tecniche di coltivazione, la distanza consigliata tra i filari, le rese consentite o raccomandate e, in parte, si descrivono alcuni aspetti relativi ai processi di vinificazione.

Il Cluster\_2\_UK, seppure in parte analogo al primo, riguarda in particolar modo i processi di selezione clonale che hanno permesso la coltivazione del Sangiovese nel Nuovo Mondo, con specifici e frequenti riferimenti all'Australia. Questo non stupisce visto il legame storico che esiste tra il continente australiano e la Gran Bretagna, l'affinità linguistica, e considerato che da quando l'Australia si è affacciata sul mondo vitivinicolo globale esportando i propri vini ha considerato il Regno Unito sia un mercato di sbocco, sia una "porta" di accesso agli altri paesi europei e non.

Il Cluster\_3\_UK propone descrizioni approfondite dei processi di selezione vinificazione, in particolare attraverso le caratteristiche dei vini a base di Sangiovese ma che risultano da blend con altri vitigni tipicamente francesi, come il Cabernet Franc e Sauvignon, il Pinot, il Merlot. Oltre che alle denominazioni toscane a base di Sangiovese, riferimenti frequenti in questo cluster vengono fatti al fenomeno dei cosiddetti Super Tuscan e dei "Cal-Ital", ovvero il movimento nato alla fine degli anni '80 in California, portato avanti dagli enologi

americani che erano alla ricerca di alternative ai vini rossi realizzati con varietal francesi.

Il Cluster\_4\_UK (13,6%) è assolutamente tipico e peculiare del Regno Unito e il macro-tema che lo caratterizza non si rintraccia, in modo così definito, in alcuno degli altri paesi anglofoni esaminati.

*Box 7: Verbatim Cluster\_4\_UK*

\*\*\*\* \*00001

SCORE ( 112.332 )

Sangiovese **pairs** with a wide range of **foods** because\_of its medium weighted **body** and **savory character**. Use **Sangiovese's savory** as a congruent **flavor** with **herbs** and tomatos. This **technique** will actually bring\_out more **fruity flavors** in the wine. A Sangiovese with high **tannins** will **work perfectly** with **rich roasted** meat, cured sausages and **hard** cheeses.

\*\*\*\* \*00001

SCORE ( 73.191 )

If you\_re serving up a **pizza**, Sangiovese will **work** an absolute treat Sangiovese **Food Pairings**: chicken **dishes** with **herbs**, **hard** acidic cheeses, **tomato-based dishes** especially ragu, **pizza** and pasta.

\*\*\*\* \*00001

SCORE ( 64.245 )

Varietal Sangiovese or those with a smaller proportion of the **powerful**, full-bodied Cabernet blended in, can accentuate the **flavours** of relatively bland **dishes** like meatloaf and **roast** chicken. **Herb** seasoning such\_as basil, thyme and sage play\_off the herbal notes of the grapes. Sangiovese that has\_been subject\_to more **aggressive oak treatment pairs** well with grilled and smoked **food**.

*Fonte: elaborazione propria*

Si tratta del tema del cosiddetto *food pairing*, ovvero dell'abbinamento cibo-vino, caratterizzato da suggerimenti su quali piatti (pizza, pasta, ribollita, ragù, arrostiti, ecc.) accompagnare con vini a base di Sangiovese. Il dettaglio con cui si parla degli abbinamenti è sorprendentemente elevato, differenziando i piatti consigliati in funzione degli altri uvaggi con cui viene vinificato il vitigno, del periodo di invecchiamento e della tipologia e dimensione dei contenitori (botti nuove o usate, barrique, tonneaux, capacità in ettolitri, ecc.).

Analogamente ad Australia e Canada, anche il processo di clustering del Regno Unito evidenzia due macro-temi: il primo

riconducibile all'etimologia del nome e ai diversi modi in cui viene chiamato il Sangiovese in funzione dell'area geografica in cui è coltivato in Italia (Cluster\_5\_UK – 18,1%) e uno alle caratteristiche fisico-chimiche e organolettiche dei vini a base di Sangiovese (Cluster\_6\_UK – 12,6%).

*Box 8: Verbatim Cluster\_5\_UK e Cluster\_6\_UK*

#### Cluster\_5\_UK

\*\*\*\* \*00001

SCORE ( 30.553 )

Brunelletto ( in the Grosseto region\_of\_Tuscany ), Brunello, Brunello\_di\_Montalcino, Cacchiano ( in Tuscany ), Calabrese ( in Tuscany ), Cardisco, Chiantino ( in Tuscany ), Cordisio, Corinto nero ( on the island of Lipari in Sicily ), Dolcetto Precoce, Guarnacciola ( in the Benevento region of Campania ), Ingannacane, Lambrusco Mendoza ( in Tuscany ).

\*\*\*\* \*00001

SCORE ( 16.642 )

Sangiovese has a long history on the island of Corsica where it is known as Nielluccio. The grape was likely brought to the island sometime between 14th and 18th century when it was ruled by the Republic of Genoa.

#### Cluster\_6\_UK

\*\*\*\* \*00001

SCORE ( 55.803 )

Young Sangiovese has fresh fruity flavours of strawberry and a little spiciness, but it readily takes on oaky, even tarry, flavours when aged in barrels. While not as aromatic as other red\_wine varieties such\_as Pinot\_noir, Cabernet\_Sauvignon, and Syrah, Sangiovese often has a flavour profile of sour red\_cherries with earthy aromas and tea leaf notes.

\*\*\*\* \*00001

SCORE ( 31.723 )

There are some hints that can help you, first of all the refreshing complexity of the bouquet, which ranges from the fragrance of violet, cherry, rose, sour cherry, peony, mulberry especially in Romagna's Sangiovese plum, carob, rhubarb, tomato jam, licorice and anise, to the aromatic waves of tea, thyme, marjoram.

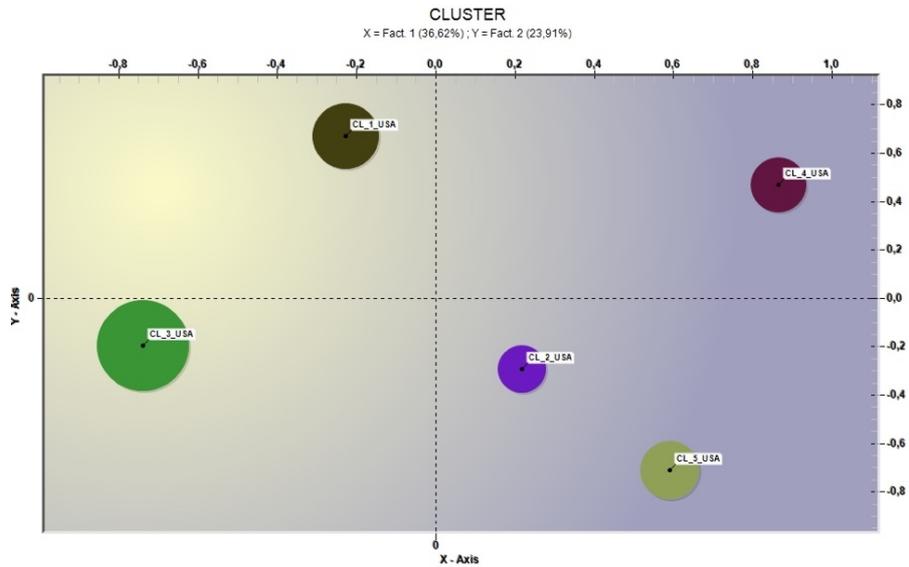
*Fonte: elaborazione propria*

Relativamente a questi due cluster, non si evidenziano particolari specificità rispetto a quanto già evidenziato nell'illustrazione degli altri paesi.

## Cluster semantica: Stati Uniti d'America

L'ultimo paese di lingua inglese che è stato incluso nell'analisi è rappresentato dagli USA.

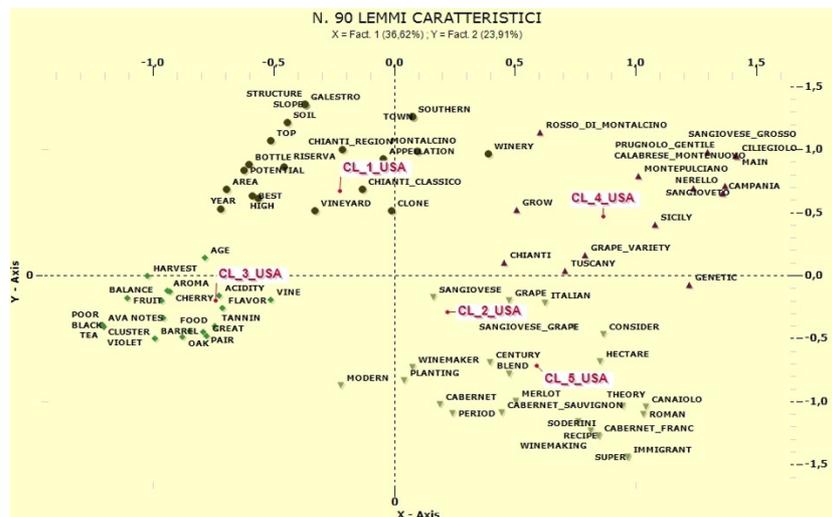
Figura 30: Cluster USA



Fonte: elaborazione propria

In questo caso, il processo di clustering ha dato luogo all'individuazione di 5 diversi macro-temi (si vedano Figura 30 e Figura 31).

Figura 31: Cluster e lemmi USA



Fonte: elaborazione propria



Analogamente agli altri paesi già trattati il primo cluster tematico statunitense (Cluster\_1\_USA), che include quasi il 21% dei contesti elementari classificati, racconta il Sangiovese attraverso la descrizione del suolo e delle condizioni climatiche che ne favoriscono la crescita e una maturazione ottimale. Rispetto agli altri paesi, tuttavia, è possibile rintracciare delle specificità che differenziano e sono propri dei corpora USA. Innanzitutto si coglie una maggiore precisione nelle descrizioni: del suolo si parla di argilla, galestro, pendenza, esposizione, altitudine, ecc. Non si parla solo in generale di zone, come quelle ampie coperte dai confini amministrativi delle indicazioni geografiche, ma si fa specifico riferimento a sotto-zone, in grado di differenziare la qualità dell'uva grazie al microclima che le caratterizza. Si parla in modo più ampio inoltre di territori che in altri contesti vengono solo sporadicamente nominati, come la Maremma, vedremo nel secondo cluster l'Umbria, della zona di Montepulciano e così via.

*Box 9: Verbatim Cluster\_1\_USA*

\*\*\*\* \*00001

SCORE ( 21.510 )

Brunello tends to **need** many more **years** in **bottle** to develop than even the sternest **Chianti\_Classico** but the **DOC Rosso\_di\_Montalcino** identifies the earlier-maturing reds of the region. Just to the east of **Montalcino**, around the **town** of Montepulciano, there is a similar system for the **local** wines, known as **vino\_nobile\_di\_Montepulciano**.

\*\*\*\* \*00001

SCORE ( 13.198 )

Stylistic and terroir based differences also emerge among the various sub-zones of the **Chianti\_region**. The **ideal vineyard** locations are found on south and southwest-facing **slopes** at altitudes between 490-800 ft ( 150-550 m ). **in\_general**, Sangiovese has a more difficult time fully **ripening** in the **Chianti\_region** than it does in the **Montalcino** and **Maremma** regions to the south.

\*\*\*\* \*00001

SCORE ( 12.982 )

Some **top** wines made predominantly from Sangiovese are Flaccianello ( from Fontodi, **Chianti\_Classico** ), Le Pergole Torte ( Monteverdine, **Chianti\_Classico** ), Zerbina **Riserva** Pietramora ( Zerbina, Romagna ).

*Fonte: elaborazione propria*

Si parla prevalentemente di Toscana, facendo riferimento esplicito ad aziende e a vini specifici segnalandone la qualità elevata (Banfi, Biondi Santi, Monteverdine con Le Pergole Torte, il Flaccianello di Fontodi, sempre nel Chianti Classico, Zerbina Riserva Pietramora in Romagna, ecc.).

Come poi verrà confermato nell'analisi del Cluster\_5\_USA (di cui una delle parole maggiormente caratteristiche, come visibile dal wordcloud, è Super Tuscan), non stupisce che le descrizioni, in special modo della Toscana, presentino un elevato grado di dettaglio e che questo sia peculiare degli Stati Uniti. Sappiamo infatti che la regione è meta di imponenti flussi turistici incoming dagli USA, che addirittura negli anni '80 è stato coniato il termine "Chiantishire" che fa riferimento all'area del Chianti e che allude al numero di persone anglofone che si sono trasferite o che visitano questa regione geografica per finalità turistiche. Tra alcune delle persone famose che possiedono proprietà nel Chianti ci sono: Polly Toynbee, Sting, Bryan Ferry, Antonio Banderas, Richard Gere, ecc.

Ancora, si pensi all'industria cinematografica che ha contribuito a creare stereotipi associati al paesaggio toscano e al lifestyle dei suoi abitanti come i cipressi, le crete senesi, i vigneti, gli uliveti, ecc., al buon cibo e vino, alla convivialità e all'accoglienza. Si pensi a film come "Sotto il sole della Toscana", del 2003, ambientato tra Cortona e Montepulciano, oppure "Letters to Juliet", del 2010, girato nei paesaggi collinari intorno a Castelnuovo Berardenga. O ancora, più recentemente e di fatto avulso dai tradizionali stereotipi toscani, si pensi alla saga di Twilight (ultimo episodio nel novembre 2012), che parla dell'antica stirpe dei vampiri di Volterra. Nonostante le scene del film siano state completamente girate a Montepulciano, in

provincia di Siena, grazie al film i flussi turistici a Volterra nell'aretino sono aumentati di circa 50.000 presenze all'anno<sup>10</sup>.

Il secondo cluster (Cluster\_2\_USA – 14,3%) racconta dello stile dei vini a base di Sangiovese e della versatilità del vitigno rispetto alle aree geografiche in cui viene coltivato. In funzione di queste, infatti, i vini si connotano di caratteristiche fisico-chimiche e organolettiche differenti. Pur mantenendo il focus sui vini rossi toscani, il macro-tema del secondo cluster include anche riferimenti ad altre regioni italiane (Marche, Lazio, Umbria) e all'estero (Corsica, Mendoza in Argentina e California). È interessante notare che, nonostante gli Stati Uniti siano un paese grande importatore di vini dall'Australia, i riferimenti al Sangiovese coltivato in questo paese sono quasi assenti.

*Box 10: Verbatim Cluster\_2\_USA*

\*\*\*\* \*00001

SCORE ( 7.632 )

Outside Tuscany, it is found throughout central\_Italy where it places an important **role** in the Denominazione di Origine Controllata e Garantita ( **DOCG** ) wines of Montefalco Sagrantino secco and Torgiano Rosso Riserva in **Umbria**, Conero in **Marche** and the Denominazione di Origine Controllata ( **DOC** ) wines of Lazio and Rosso Piceno in **Marche**.

\*\*\*\* \*00001

SCORE ( 4.636 )

The formula for Brunello\_di\_Montalcino **DOCG** was created in the 1860s by Biondi-Santi, but it wasn't until the 1980s that the appellation started gaining widespread recognition. Generally **full-bodied**, tannic, and complex, Brunello\_di\_Montalcino is often likened in **style** to Barolo and is generally considered to be the key region of central\_Italy.

\*\*\*\* \*00001

SCORE ( 4.525 )

We know and love the Tuscan varietal from **full-bodied** red\_wines like Chianti and Brunello\_di\_Montalcino, as\_well\_as for being the catalyst of the Super\_Tuscan **style**. However, Sangiovese has not always been admired. Throughout its storied past, the grape has paved the way for Italian wines while simultaneously been at the **heart** of several viticultural controversies.

*Fonte: elaborazione propria*

<sup>10</sup> <http://guide.supereva.it/twilight/interventi/2010/04/news-su-volterra-e-laiuto-della-saga-al-turismo>.

Il terzo e il quarto cluster (Cluster\_3\_USA - 29,7% e Cluster\_4\_USA - 17%) rappresentano due macro-temi analoghi a quelli rintracciati negli altri paesi anglofoni, ovvero le caratteristiche fisico-chimiche dei vini a base di Sangiovese e i nomi che assume il vitigno nelle diverse zone in cui è coltivato. Nel Cluster\_3\_USA, si parla di aromi, sapori, colori, facendo riferimento ai tratti organolettici tipici dei vini rossi realizzati col vitigno (il viola, il granato, il pepe, il tabacco, il timo, la cannella, la fragola, la ciliegia, la maggiorana, ecc.) e ai tratti chimici (i tannini, l'acidità, la freschezza), evidenziando anche le differenze che caratterizzano i vini realizzati all'estero.

*Box 11: Verbatim Cluster\_3\_USA e Cluster\_4\_USA*

#### Cluster\_3\_USA

\*\*\*\* \*00001

SCORE ( 34.788 )

Typical **harvest** may not be until mid to late **October**. The long\_growing\_season gives the grape **time** to **develop** richness and **body**. Sangiovese usually displays characteristics of **red fruits** tart **cherries**, **red plum** and **strawberries**. It portrays secondary characteristics of roses, oregano, **thyme**, **leather**, clay, **tobacco**.

\*\*\*\* \*00001

SCORE ( 25.018 )

Californian examples tend to have more **bright, red fruit flavours** with some Zinfandel-like **spice** or darker **fruits** depending\_on the proportion of Cabernet blended in. Argentine examples showing a hybrid between the Tuscan and California Sangiovese with juicy **red fruit** wines that end on a bitter **cherry** note.

#### Cluster\_4\_USA

\*\*\*\* \*00001

SCORE ( 84.053 )

The two **main** sub-types are **Sangiovese\_grosso** and **Sangiovese\_Piccolo**. **Sangiovese\_grosso**, also **known** as **dolce** and **gentile**, is **synonymous** with Sangiovese\_di\_Lamole of the Greve-Firenze region of **Chianti**, the **Prugnolo\_Gentile** of **Montepulciano**, and the **Brunello** of Montalcino.

\*\*\*\* \*00001

SCORE ( 53.788 )

Other **DNA studies** have also shown an existing genetic\_relationship between the Sangiovese and other **grapes** such\_as Frappato, Gaglioppo, **Nerello** Mascalese, etc. There are many **families** and clones of the Sangiovese\_grape, but here I will only mention the two most **common families**, namely the **Sangiovese\_Piccolo** and **Sangiovese\_grosso** ( **Prugnolo\_Gentile** ).

*Fonte: elaborazione propria*

Nel Cluster\_4\_USA si racconta dei diversi cloni che danno origine a diverse tipologie di uve in differenti aree produttive italiane. Si fa riferimento alla Toscana, alla Romagna, alla Puglia, Sicilia, Campania e si avvalorano le affermazioni riportando le evidenze di ricerche sui cloni e sul DNA del vitigno. È interessante notare come, in maniera peculiare negli USA, si faccia frequentemente riferimento alle famiglie del Sangiovese, ovvero il Sangiovese grosso, con cui viene prodotto il Brunello di Montalcino, e il Sangiovese piccolo.

*Box 12: Verbatim Cluster\_5\_USA*

\*\*\*\* \*00001

SCORE ( 52.831 )

One of **Sangiovese**'s more **modern** incarnations is in the so-called "**Super\_Tuscans**", which are made under the Toscana IGT category. These wines allow **winemakers** more freedom to **blend** indigenous **Italian grapes** (principally **Sangiovese**) with **Cabernet\_Sauvignon**, **Cabernet\_Franc**, and **Merlot**

\*\*\*\* \*00001

SCORE ( 41.882 )

In 1883, the **Italian** writer **Giovanni** Cosimo Villifranchi echoed a similar description about the quality of Sangiovese being dependent on the **grapes** with which it was **blended**. The **winemaker** and politician, Bettino Ricasoli formulated one of the **early recipes** for Chianti when he **blended** his **Sangiovese** with a sizable amount of **Canaiolo**.

\*\*\*\* \*00001

SCORE ( 39.622 )

There was a vogue in the late 1970s and 1980s to minimize **Sangiovese**'s role and **blend** in very obvious proportions of Cabernet and Merlot. Marchese Piero Antinori set in train a huge fashion for **blending** in these **Bordeaux grapes** with the local **Sangiovese** when he **released** the then mould-breaking Tignanello in the **early** 1970s mimicking the **recipe** for **Carmignano** on the other side of Florence.

\*\*\*\* \*00001

SCORE ( 22.521 )

Yet as recently as the 1970s, **Sangiovese** from some of Tuscany's most celebrated regions wasn't necessarily much to **talk** about. Tighter restrictions on **blending** with other **grapes** have improved the reputation of Chianti\_Classico, while so-called **Super\_Tuscan** wines blended with Cabernet Sauvignon and **Merlot** command premium prices just like cult California **Cabernet**.

*Fonte: elaborazione propria*

Come accennato, il quinto cluster (Cluster\_5\_USA – 18,3%) non trova alcuna corrispondenza negli altri paesi analizzati. Racconta del Sangiovese vinificato con altri uvaggi, tipicamente quelli francesi

come il Merlot e il Cabernet, che ha dato origine ai vini che sono diventati famosi in tutto il mondo con il nome di *Super Tuscons*. Come è noto, la nascita di questo nome – coniato da Wine Spectator – risale agli anni settanta, quando fu modificata la procedura di preparazione del Chianti eliminando le uve bianche per sostituirle con alcune varietà di Bordeaux come Cabernet Sauvignon e Franc, e Merlot. Erano vini rossi che non rispettavano di proposito le regole dei disciplinari di produzione dei vini a denominazione della regione, come il Tignanello, il Solaia, il Sassicaia, ecc., che incontravano il cosiddetto “gusto internazionale”, e che riscossero un successo planetario, specialmente negli Stati Uniti. Un macro-tema è dunque quello che narra la storia di questi vini e dei protagonisti della vitivinicoltura toscana, come Bettino Ricasoli o la famiglia Antinori.

## L'ITALIA

L'Italia è il paese con il numero più alto di termini riferibili a "Sangiovese". Come si evince da Tabella 6, il processo di saturazione utilizzato per selezionare i corpora da sottoporre ad analisi in Italia porta a raccogliere un testo decisamente più "ricco" di parole rispetto a quello dei paesi anglofoni. Il corpus costituito dall'unione delle fonti documentali raccolte è infatti caratterizzato da un numero di occorrenze che è circa il doppio di quello degli altri paesi (oltre 30.000 token).

*Tabella 6: Caratteristiche corpus lingua italiana*

<b>CARATTERISTICHE DEL CORPUS</b>	<b>NUMERO</b>
Url selezionati e corpora	32
Occorrenze (token)	30.106
Forme (type)	6.074
Hapax (occ. = 1)	3.633
Lemmi	4.727
Frequenza minima nel testo	>=10
Lemmi sottoposti ad analisi	261
Contesti elementari	744

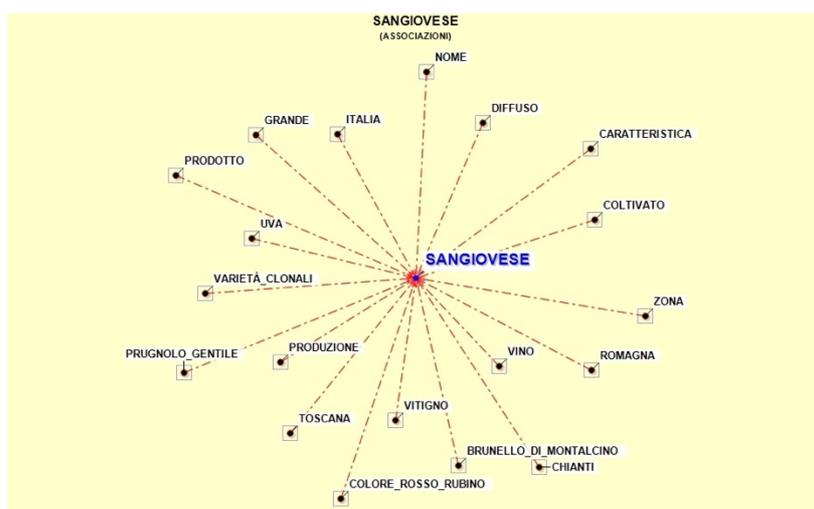
*Fonte: elaborazione propria*

Si comprende dunque come in Italia la narrazione del Sangiovese sia più cospicua e articolata rispetto a quella fatta all'estero. In questo caso, dunque, si è stabilito un limite minimo di occorrenza nel testo pari a 10 per includere i lemmi dei testi nelle elaborazioni successive.

## ASSOCIAZIONI DI PAROLE: SANGIOVESE

Figura 33 e Tabella 7 illustrano i risultati dell'analisi dei lemmi che co-occorrono più frequentemente con la parola Sangiovese nei contesti elementari dei testi. Si vede, come intuibile, che ai primi posti ci sono le parole vino e vitigno, al quarto uva, e parole che riguardano la produzione diffusa (produzione, diffuso, coltivato).

Figura 33: Diagramma associazioni Sangiovese Italia



Fonte: elaborazione propria

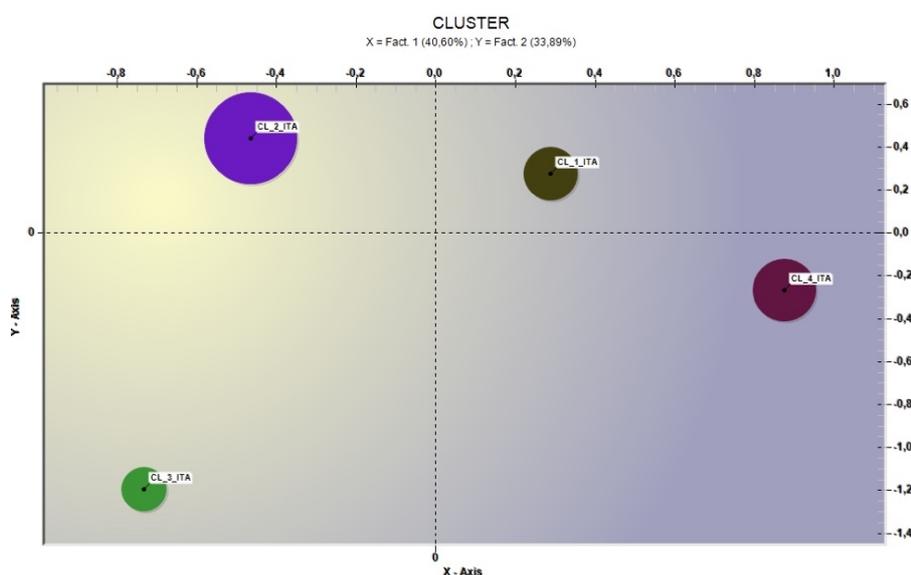
Tabella 7: Lemmi e associazioni Sangiovese Italia

LEMMA B	C.E.(B)	C.E.(AB)
vino	283	157
vitigno	154	95
produzione	134	84
uva	118	76
Italia	68	56
diffuso	63	52
coltivato	84	57
Romagna	88	55
Brunello_di_Montalcino	65	46
Toscana	88	53
varietà_clonali	51	40
grande	78	47
nome	82	48
caratteristica	79	47
zona	79	45
Chianti	53	36
colore_rosso_rubino	61	37
Prugnolo_Gentile	31	26
prodotto	61	35

Fonte: elaborazione propria

È interessante notare che, contrariamente a quanto avviene nei paesi anglofoni, tra le associazioni più ricorrenti con la parola Sangiovese c'è Romagna, addirittura prima della Toscana e delle più famose denominazioni di questa regione (Brunello di Montalcino e Chianti). Vedremo infatti che in Italia si parla molto diffusamente del Sangiovese di Romagna e che questo costituisce un macro-tema che include oltre il 20% dei contesti elementari classificati.

Figura 34: Cluster Italia



Fonte: elaborazione propria

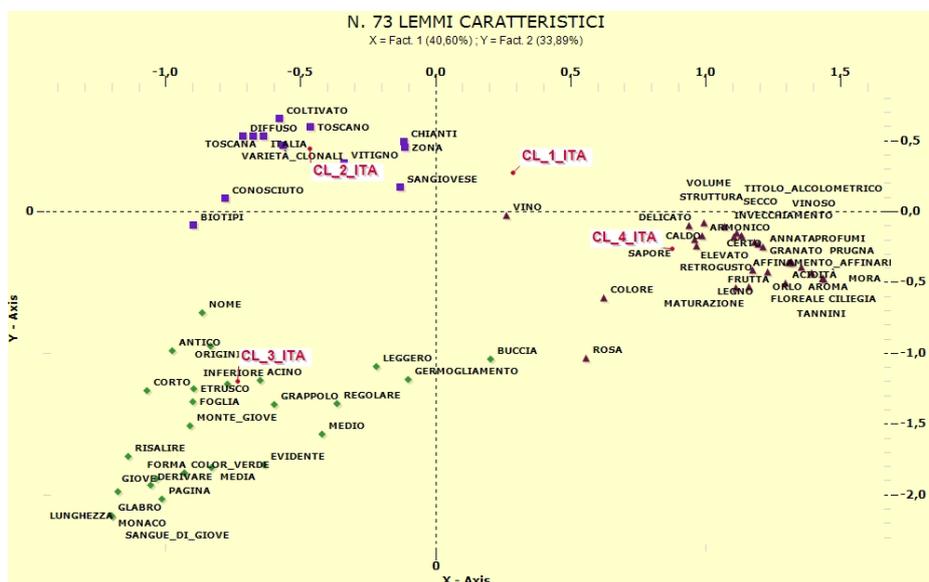
## ANALISI DEI TEMI EMERGENTI

Il processo di analisi semantica ha fornito una soluzione a quattro cluster per il corpus in lingua italiana.

Come già detto, un macro-tema chiaramente identificabile nelle fonti documentali italiane è riferibile al Sangiovese di Romagna e alla produzione di vini nella regione romagnola (Cluster\_1\_ITA – 20,6%). Nonostante la denominazione sia stata modificata in Romagna Sangiovese dal 2011, in questo cluster si parla sostanzialmente della DOC continuando a chiamarla con il nome originario. Si descrivono le zone di produzione, gli uvaggi consentiti nel processo di vinificazione,

le tipologie (Novello, Superiore e Riserva) e gli abbinamenti con il cibo. Il tono è descrittivo e piuttosto didascalico (Box 13).

Figura 35: Cluster e lemmi Italia



Fonte: elaborazione propria

Box 13: Verbatim Cluster\_1\_ITA

\*\*\*\* \*00001

SCORE ( 207.259 )

Il **Sangiovese\_di\_Romagna Doc** viene **prodotto** da **uve** raccolte in un'ampia zona collinare **compresa** nelle **province\_di\_Bologna, Forlì - Cesena, Rimini e Ravenna** ed è stato il primo ad essere riconosciuto tra i **vini Doc** nel comprensorio **romagnolo**.

La **Doc** prevede le **tipologie Novello, Superiore e Riserva**.

\*\*\*\* \*00001

SCORE ( 199.892 )

Gli **abbinamenti** del **Sangiovese\_di\_Romagna Superiore** e il **Sangiovese\_di\_Romagna Riserva** sono i tipici **abbinamenti** dei grandi vini\_rossi: **carni brasate, arrosti e cacciagione**.

La **denominazione Sangiovese\_di\_Romagna DOC** rappresenta una delle più importanti aree vitivinicole della regione.

\*\*\*\* \*00001

SCORE ( 97.867 )

Il **Sangiovese\_di\_Romagna** è coltivato invece nelle **province** di **Rimini, Forlì - Cesena, Ravenna e Bologna**. Questa **DOC** prevede l'utilizzo di **uve** Sangiovese non inferiore all'85%. Il rimanente 15% può essere composto da **uve** a bacca\_rossa a\_patto che siano coltivate in **terre romagnole**.

\*\*\*\* \*00001

SCORE ( 58.002 )

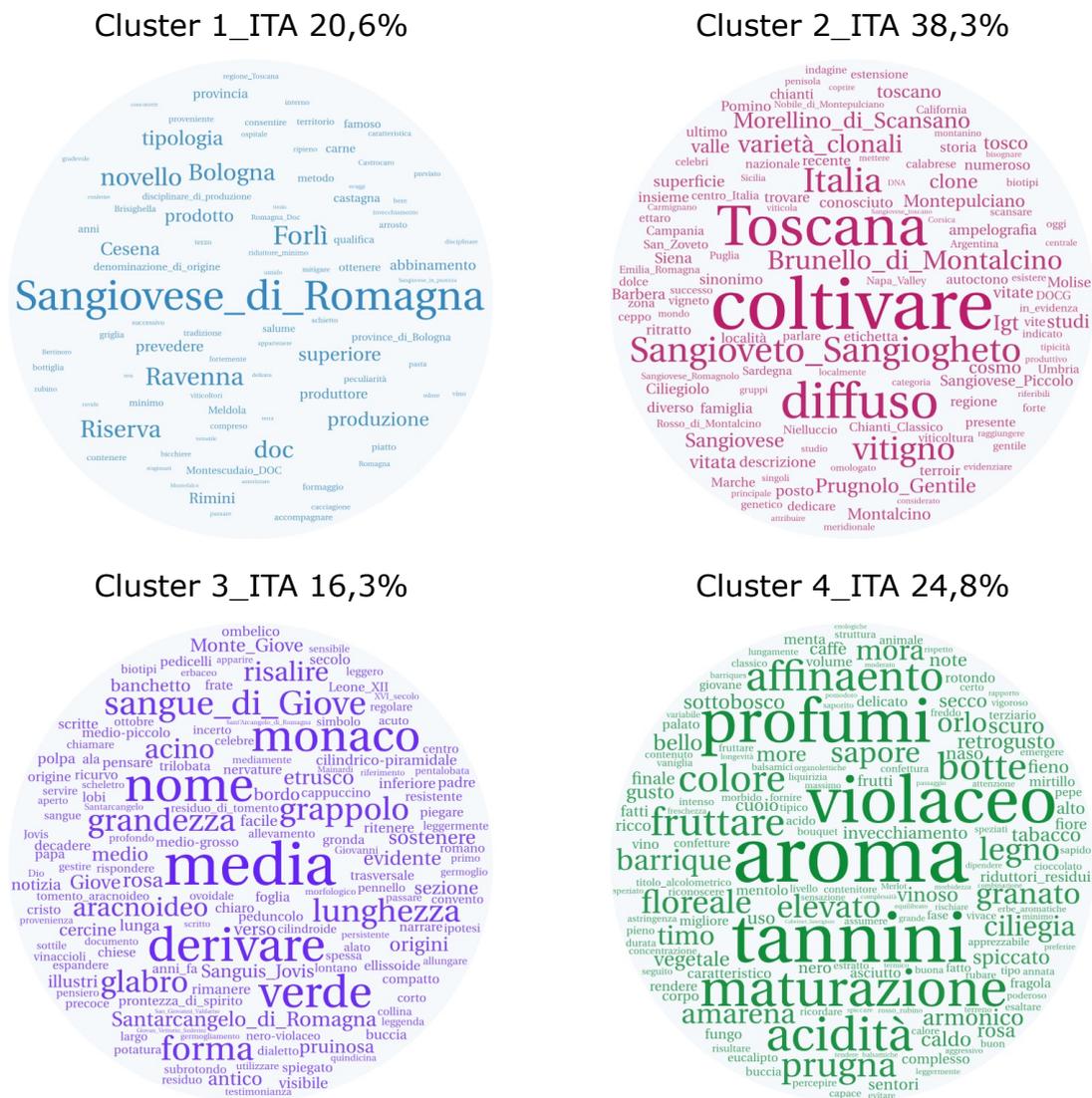
Questo **vino**, noto anche a livello\_internazionale ( viene esportato in **Australia, Canada, Giappone, USA e vari Paesi europei** ), viene **prodotto** in diversi comuni

della regione Emilia\_Romagna appartenenti alle province di Bologna, Forlì - Cesena, Ravenna e Rimini. Ogni anno vengono prodotte circa 6 milioni di bottiglie di Sangiovese di Romagna.

Fonte: elaborazione propria

Il secondo cluster (Cluster\_2\_ITA), di dimensioni decisamente più rilevanti rispetto agli altri includendo quasi il 40% dei contesti elementari classificati, è quello che narra il Sangiovese adottando la prospettiva agronomica ed enologica.

Figura 36: Wordcloud cluster Italia



Fonte: elaborazione propria

Si fa spesso riferimento all'ampelografia, si descrivono le caratteristiche del vitigno e le sue varietà, i cloni, i biotipi e si

evidenzia l'elevata adattabilità delle piante e le mutazioni che subisce il vitigno adattandosi a contesti ambientali anche molto diversi tra loro sia in Italia, sia all'estero. Si parla delle caratteristiche dei vini a base di Sangiovese, prevalentemente quelli toscani indicando le denominazioni o facendo riferimento ai Super Tuscan, e degli uvaggi che solitamente si vinificano insieme al Sangiovese, come il Trebbiano, il Canaiolo, la Malvasia, ecc. In questo caso, il linguaggio utilizzato è estremamente tecnico, specialistico, talvolta difficile da comprendere da parte di un ipotetico lettore non esperto in materia di vitivinicoltura.

*Box 14: Verbatim Cluster\_2\_ITA*

\*\*\*\* \*00001

SCORE ( 140.830 )

Tra le più diffuse e note sono da citare il Sangiovese grosso, coltivato in Toscana, che comprende tra gli altri i biotipi brunello e Prugnolo\_gentile, e il Sangiovese\_piccolo, tipico prevalentemente di Emilia-Romagna e Toscana, dove è famoso nel biotipo morellino, nella zona di Scansano. Fuori dall'Italia viene coltivato soprattutto in California e Argentina.

\*\*\*\* \*00001

SCORE ( 51.806 )

La Commissione\_Ampelografica della Provincia\_di\_Siena ( 1875-76 ) indica tra i vitigni più diffusi nel Chianti il Sangiovese e un Calabrese (ampelograficamente diverso), a Montepulciano il Prugnolo e a Montalcino il Brunello. La stessa Commissione si pone il dubbio che Sangiovese, Prugnolo e Brunello, insieme al Sangiovese piccolo, abbiano in realtà un'unica identità.

\*\*\*\* \*00001

SCORE ( 45.416 )

Esistono comunque numerose varianti e cloni, soprattutto perché come il Pinot Noir il Sangiovese tende a mutare spontaneamente, come ad esempio in Sardegna dove è conosciuto con il nome di Nielluccio. Il Sangiovese è un vitigno molto diffuso in Italia, il più diffuso insieme al Barbera, con l'11 per cento delle superfici vitate di vitigni rossi.

\*\*\*\* \*00001

SCORE ( 34.154 )

Il Sangiovese è tra i vitigni di uva rossa più diffusi grazie alla sua capacità di crescere in presenza di clima temperato, pertanto si può incontrare la coltivazione di vitigno del Sangiovese in ogni zona d'Italia ad esclusione delle regioni del nord e dell'estremo sud.

*Fonte: elaborazione propria*

Il terzo cluster (Cluster\_3\_ITA) è invece quello di dimensioni minori, includendo poco più del 16% dei contesti elementari classificati. Anche in questo cluster parte della narrazione è di tipo molto tecnico ed è dedicata a una descrizione estremamente dettagliata delle caratteristiche ampelografiche e delle modalità di coltivazione del Sangiovese (foglia media, trilobata o pentalobata, seno peziolare a U, ecc.). Anche in questo caso, per un non esperto, la lettura e la comprensione di queste informazioni risulta assai difficile (Box 15).

*Box 15: Verbatim Cluster\_3\_ITA – ampelografia*

\*\*\*\* \*00001

SCORE ( 26.169 )

Il vitigno Sangiovese ha le seguenti caratteristiche varietali: **foglia media, trilobata o pentalobata**, di colore\_verde\_chiaro; **grappolo medio, compatto, cilindrico-piramidale, alato**; **acino** medio-grande, **ovoidale di forma regolare**; **buccia** di colore nero-violaceo e ricca di pruina, non molto **spessa**.

\*\*\*\* \*00001

SCORE ( 21.174 )

Caratteristiche ampelografiche **Foglia: media**, pentagonale, **pentalobata**; lembo piano con **lobi leggermente piegati a gronda**; seno\_peziolare ad U **aperto**; seni\_laterali\_superiori ad U con margini **leggermente** sovrapposti, seni\_laterali\_inferiori appena accennati, poco **aperti**; **pagina** superiore e **pagina inferiore glabre** o con **leggerissima** peluria.

*Fonte: elaborazione propria*

Parte invece della narrazione del medesimo cluster si riferisce all'etimologia della parola Sangiovese e si raccontano le storie e le leggende che formano un vero e proprio "alone di mistero" intorno all'origine del nome di questo vitigno.

In questo caso si tratta di una narrazione che non ha una mera finalità "informativa" e che invece abbiamo visto caratterizzare i contenuti della maggior parte dei documenti che sono stati analizzati sia in Italia, sia all'estero. Riferimenti sono fatti agli Etruschi, ai Romani, agli dei, a monaci e papi che sono i protagonisti delle storie che costituiscono il mito del Sangiovese. Possiamo quindi affermare che questo è il cluster che maggiormente si avvicina a un vero e

proprio storytelling del vitigno e che si rivolge maggiormente alla sfera emotiva del lettore e non meramente a quella cognitiva (Box 16).

*Box 16: Verbatim Cluster\_3\_ITA – etimologia*

\*\*\*\* \*00001

SCORE ( 49.191 )

Gli ampelografi **ritengono** che il luogo di **origine** del Sangiovese sia la Toscana, dove tuttora rappresenta la varietà di uva a bacca\_rossa\_più importante. Il mito **narra** che il **nome** derivi da **Sanguis\_Jovis**, **sangue\_di\_Giove**, **nome** attribuitogli da un **monaco cappuccino** del **convento** di **Sant'Arcangelo\_di\_Romagna**, nei pressi del **Monte\_Giove**.

\*\*\*\* \*00001

SCORE ( 37.196 )

Questi, gradì tale vino e ne **chiese** il **nome**, mettendo i **frati** in imbarazzo poiché non avevano mai **pensato** di dargli un **nome**. Uno dei **monaci**, però, con **prontezza\_di\_spirito** disse che il vino si **chiamava Sanguis\_Jovis** ( **sangue\_di\_Giove** ), **nome** che nei **secoli** si mutò in "**sangue\_di\_Giove**" e poi in "sangiovese".

\*\*\*\* \*00001

SCORE ( 34.350 )

Le **prime notizie** sul vitigno denominato Sangiovese **risalgono** al XVII **secolo** e le sue **origini** sono avvolte da un velo di leggendario mistero. Il suo curioso **nome**, che in **dialetto** romagnolo suona come sanzves, sembrerebbe, infatti, **derivare** dal **Monte\_Giove** o **Colli Jovis**, una **collina** che si trova nell'entroterra riminese.

\*\*\*\* \*00001

SCORE ( 23.145 )

Il suo **nome** potrebbe **derivare** da **sangue\_di\_Giove**, il suo **antico nome** derivato dalla tradizione **etrusca** del Settecento avanti **Cristo**, come ex\_voto di ringraziamento **verso** il **padre** degli dei che regalava all'uomo questo frutto prezioso.

\*\*\*\* \*00001

SCORE ( 21.338 )

Si **narra** che durante un **banchetto** tenuto nel Monastero\_dei\_Frati **Cappuccini** in **Santarcangelo\_di\_Romagna**, alla\_presenza di **Papa Leone\_XII** ed **illustri** ospiti, fu **servito** questo vino prodotto dagli stessi **monaci**. Il vino fu molto apprezzato e ne fu chiesto il **nome**. Un **monaco**, con **prontezza\_di\_spirito**, disse che il vino si **chiamava Sanguis Jovis** = **sangue\_di\_Giove** ( Sanjovese ).

*Fonte: elaborazione propria*

L'ultimo cluster (Cluster\_4\_ITA – 24,8%) è analogo al cluster individuato in ciascuno dei paesi anglofoni analizzati e si caratterizza per la descrizione delle caratteristiche fisico-chimiche e organolettiche dei vini a base di Sangiovese. Anche in questo caso, si

parla degli aromi tipici dei vini e di come cambino in funzione dei processi di invecchiamento sui sono sottoposti.

*Box 17: Verbatim Cluster\_4\_ITA*

\*\*\*\* \*00001

SCORE ( 85.932 )

Dal colore rosso\_rubino\_intenso fino al Granato, dopo lungo invecchiamento la sfumatura potrà assumere anche tonalità\_aranciate. Gli aromi del Sangiovese sono prevalentemente orientati ai frutti\_rossi e neri tra cui spiccano amarena, mora e prugna. Fra gli aromi floreali la violetta è il più caratteristico seguito dall'aroma di rosa.

\*\*\*\* \*00001

SCORE ( 75.586 )

I profumi classici del Sangiovese sono la prugna e la viola, ma poi il cuoio e gli aromi terziari classici dei grandi vini. Grandi aromi maturi annegano in colori scuri, con vini ricchi e variabili per le annate, dove quelle fredde accentuano i tannini.

\*\*\*\* \*00001

SCORE ( 59.759 )

Il Sangiovese\_di\_Romagna è un vino dal colore\_rosso\_rubino che talvolta può avere orli violacei, odore vinoso con un profumo che ricorda quello della viola; il sapore è secco, armonico, leggermente tannico, con un retrogusto\_amarognolo. Il titolo alcolometrico\_volumico\_totale minimo deve essere 12% vol.

\*\*\*\* \*00001

SCORE ( 31.117 )

Altri aromi di frutta che si possono percepire nel Sangiovese sono ribes, fragola, ciliegia, lampone e mirtillo. Il Sangiovese viene maturato prevalentemente in contenitori di\_legno - botti e barriques - pertanto saranno frequenti anche aromi di tipo speziato - vaniglia e liquirizia su tutti - così come aromi empireumatici come tostato, caffè e cioccolato.

*Fonte: elaborazione propria*

## **CONSIDERAZIONI DI SINTESI**

In base all'analisi effettuata, è possibile compiere alcune considerazioni. Innanzitutto, un dato che sembra caratterizzare trasversalmente l'Italia e i paesi esaminati di lingua inglese è il taglio diffusamente tecnico (agronomico, ampelografico, enologico) con cui si parla di Sangiovese. Questo non stupisce in quanto si tratta di un vitigno, e non di un vino, dunque della tipologia e delle qualità che caratterizzano la materia prima principale con cui si realizzano i vini, l'uva. In questo caso, la consapevolezza dei consumatori (quella che chiameremmo awareness nel caso in cui si trattasse di una marca vera e propria) relativa al vitigno è inferiore rispetto ai nomi delle famosissime denominazioni dei vini con questo realizzati, come il Brunello, il Chianti e così via. Si pensi, per esempio, che il Chianti è la categoria di vino italiano più nota al mondo (Penco, 2018), ma nessuno studio è stato realizzato per misurare la conoscenza che i consumatori italiani e/o stranieri hanno del vitigno con cui viene realizzato. Mentre consapevolezza sui vitigni esiste (almeno in parte) relativamente ai cosiddetti vitigni internazionali, come lo Chardonnay, il Sauvignon, il Pinot, nel caso del Sangiovese si potrebbe affermare che il nome del vitigno sembrerebbe "soffrire" a causa dell'elevata notorietà delle categorie di vino con questo prodotte. Ecco che chi parla online di Sangiovese ne parla in maniera tecnica, in una prospettiva prevalentemente viticola o enologica, e in certi casi, come in Italia, adotta un linguaggio adatto per interloquire con lettori esperti, conoscitori del mondo del vino.

Da questo deriva che le informazioni veicolate online sul Sangiovese non mostrano i tratti "canonici" dello storytelling convenzionalmente considerato. La prospettiva comunicazionale adottata predilige la trasmissione di informazioni e nozioni, con un fine prevalentemente *esplicativo* delle caratteristiche del vitigno. Non c'è una vera e propria narrazione di storie, sono marginali i riferimenti a episodi fittizi o realmente accaduti a persone e contestualizzati in un certo tempo e in un certo luogo. Sappiamo invece che lo storytelling, inteso come narrazione, in cui ci sono personaggi, protagonisti, gesta, azioni, in cui c'è una trama, uno schema, un disegno e un filo conduttore, in cui ci sono un inizio e una fine, una storia con il suo epilogo, è uno strumento straordinario per favorire il coinvolgimento e la memorabilità nelle persone. È ormai noto in letteratura che le *storie* sono potenti strumenti per far sì che gli esseri umani memorizzino, recuperino e trasmettano le informazioni, e rappresentano la forma più profonda per interagire e comunicare in contesti sociali (Kelley e Littman 2006).

Considerando in particolare la sfera del consumo e la percezione di prodotti e brand da parte degli individui, si è dimostrato come il riconoscimento di una narrazione attraverso la selezione di alcune frasi abbia un effetto sui destinatari della storia a più livelli, essendo in grado di attivarne il sistema emotivo, cognitivo e le loro credenze (Adaval et al., 2007). Più recentemente, si è constatato che le storie possono influenzare gli atteggiamenti, le intenzioni e anche i comportamenti d'acquisto dei consumatori (Van Laer et al., 2013).

Nell'ambito del Sangiovese, preme evidenziare che le informazioni cui accedono i consumatori con maggiore probabilità sono presenti su portali verticalmente specializzati nel business vitivinicolo, talvolta su testate giornalistiche (quotidiani generalisti o riviste di settore). Sono quindi in genere fonti terze rispetto ai soggetti che hanno un interesse economico diretto, come le imprese produttrici o in qualche

modo anche i consorzi di tutela delle indicazioni geografiche a base di Sangiovese. Appare dunque plausibile che la finalità della comunicazione sia esplicativa e abbia un tono "didascalico", per fornire al lettore le informazioni di base sul vitigno e sulle sue caratteristiche in maniera oggettiva. Il tono risulta invece maggiormente "promozionale" allorquando si leggono testi prodotti e diffusi dalle aziende.

Relativamente ai contenuti veicolati, si è constatato che quando si parla di Sangiovese all'estero, nei paesi anglofoni indagati, si fa sostanzialmente riferimento alla Toscana e si mette in luce che il vitigno è l'ingrediente principale di vini di prestigio e di qualità elevata in grado di connotarli in maniera marcata da un punto di vista organolettico. In questi paesi emerge chiaramente che il Sangiovese è un "ambasciatore" dell'italianità e dello stile caratteristico delle nostre produzioni vitivinicole. Anche nei paesi in cui si coltiva il vitigno e si realizzano vini con esso, come in California e in Australia, i vini italiani (toscani in special modo) rappresentano un punto di riferimento e sono oggetto di imitazione da parte dei vitivinicoltori locali. Si parla invece molto meno di Emilia Romagna, anzi spesso dicendo che in questa regione il Sangiovese può dare origine a vini di qualità molto elevata, ma anche molto povera.

Situazione opposta si rintraccia nelle fonti in italiano, in cui abbiamo appurato che di Emilia Romagna si parla moltissimo, tanto da costituire uno dei principali temi dei testi. In primis, è opportuno considerare che la DOC di questa regione contiene il termine Sangiovese (Romagna Sangiovese o Sangiovese di Romagna) ed è quindi plausibile che un navigatore atterri con maggiore probabilità su siti che parlino della denominazione. Secondariamente, facendo riferimento a quanto già detto poco sopra, è verosimile che per esempio in Toscana si parli di più delle denominazioni (più note) e meno del vitigno.

Un'ultima considerazione riguarda la numerosità delle fonti che possono essere potenzialmente consultate da un consumatore alla ricerca di informazioni sul Sangiovese. Abbiamo visto che per quanto riguarda le ricerche in lingua inglese, il secondo idioma più diffuso al mondo dopo il cinese mandarino<sup>11</sup>, esiste una situazione di forte "concentrazione" delle fonti. In base all'analisi effettuata in Australia, Canada, Regno Unito e USA, abbiamo verificato che, indipendentemente dalle keyword utilizzate per le ricerche, un quarto degli "atterraggi" avviene sui medesimi 25 url e addirittura sugli stessi 9 siti (che presentano più pagine in cui si parla di Sangiovese). In sostanza esiste un gruppo che potremmo definire di *gatekeeper* dell'informazione, in grado di veicolare gran parte dei flussi di persone che effettuano ricerche online e far in modo che leggano quanto da questi trasmesso. Questo suggerisce che, nel caso in cui si volesse proporre una narrazione del Sangiovese con specifiche caratteristiche, sarebbe opportuno interloquire con i gestori di questi siti concordando con questi i contenuti da presentare.

---

<sup>11</sup> <https://www.travel365.it/lingua-piu-parlata-al-mondo.htm>

## **BIBLIOGRAFIA**

Antonopoulos, N., Veglis, A., & Emmanouloudis, A. (2017). Online Marketing for Media: The Case of Greek News Websites. *International Journal of Marketing Studies*, Vol. 9, n. 2, 104.

ISTAT (2010). VI Censimento Generale dell'Agricoltura, [www.istat.it](http://www.istat.it)

Lancia, F. (2012). T-Lab pathways to thematic analysis, disponibile su <http://www.tlab.it/en/tpathways.php> [ultimo accesso effettuato 15/10/2017].

Lancia, F. (2014). T-Lab user's manual, <http://tlab.it/en/bibliography.php> [ultimo accesso effettuato 15/10/2017].

Mastrobuoni G., Peracchi F., Tetenov A. (2014). Price as a Signal of Product Quality: Some Experimental Evidence, *Journal of Wine Economics*, Vol. 9, No. 2, pp. 135-152.

Oczkowski E. (2016). Identifying the Effects of Objective and Subjective Quality on Wine Prices, *Journal of Wine Economics*, Vol. 11, No. 2, pp. 249-260.

Penco P. (2018). Under the Tuscan brand, *Wine Business*, <https://winebusiness.wordpress.com/2018/08/05/under-the-tuscan-brand/> [ultimo accesso effettuato 04/09/2018].

Stemler, S. (2001), "An overview of content analysis", *Practical Assessment, Research & Evaluation*, Vol. 7, n. 17, pp. 1-10.