MARKET RESEARCH AND DATA ANALYSIS

Il TEXT MINING

Prof.ssa Silvia Ranfagni

Netnography

- **Netnography** is a **qualitative methodology** proposed by Kozinets (2002).
  - It adopts **ethnographic techniques** to **analyze** and **understand consumer behaviors** emerging from texts produced by online communities.
  - Like ethnography, it is based on the **observation of consumer interactive processes**, **but it uses computer-mediated texts** rather than data collected from live encounters (Arnauld and Wallendorf, 1994).

In our methodological approach, from netnography we borrowed → the observation technique applied to online communities and → the phases of an online community research (entrée, data collection, data analysis, data interpretation).

Text mining

- **Text mining** is a **quantitative methodology** used to **extract information** from relatively large amounts of **textual data** (Witten, 2005).
  - It draws from «corpus linguistics», using **software applications** to extract **new types of information** (e.g. word frequency, semantic categories), thus going a step beyond simple information retrieval (Hearst, 1999).
  - It can be used to analyze brand association, **investigating language in texts produced by consumers** (Rickman and Cosenza, 2007; Chen's, 2009; Archak, Ghose and Iperrotis 2007).

In our methodological approach from text mining we borrowed → software applications to identify brand associations emerging from texts produced by fashion bloggers.

Text mining application – Brand association

- Brand associations are a **synthesis** of consumer **brand knowledge** (Anderson, 1983; Keller, 1993) and are **components** of consumer's **brand image** (Biel, 1991).
- Companies look for **strong**, **positive** and **unique brand associations** (Broniarczyk and Gershoff, 2003; Bridges et al., 2000; Chen, 2001). The perception of **brand uniqueness** produces **brand differentiation**, has a positive impact on **consumer choices** (Carpenter et al., 1994) and on **brand performance** (Romaniuk and Gaillard, 2007).
- Companies look for a **match** (Brown, 1950; Tyler, 1957; Venkatraman, 1989) between **brand image** and **brand identity** (Aaker, 2003; Keller, 2003). Not only uniqueness, but also matching produces a positive impact on brand performance.

Research context – The online community

- **Online communities** are seen as **marketplaces** where **consumers** and **users interact** (Muniz and O'Guinn, 2001; Szmigin et al. 2005), **produce** and mutually **exchange information** (Cova, 1997).
- **Online communities** are **social contexts,** where researchers can study the **complex interactions of consumers with** markets and in particular, with companies and their **brands** (De Valck, 2005).
- **Consumers of fashion** now interact by means of digital platforms (Rickman and Cosenza, 2007, Boyd Thomas et al. 2007) → The **online communities** of **fashion bloggers** are important contexts to analyze brand associations.

ELSEVIER

Journal of Interactive Marketing

Volume 35, August 2016, Pages 70–85

How to Measure Alignment in Perceptions of Brand Personality Within Online Communities: Interdisciplinary Insights

Silvia Ranfagni[a], Belinda Crawford Camiciottoli[b], Monica Faraoni[a]

Show more

Understanding brand personality

- **Brand personality** can be a powerful tool to evoke emotions (Biel 1993), build trust and loyalty (Fournier 1998), and enhance consumer preference (Aaker 1999). Thus, it increases the uniqueness of brands which, in turn, contributes to brand equity (Biel 1993; Ogilvy 1985).
- It is seen as a "**set of human characteristics associated with a brand**" (Aaker 1997, 347), which combine *physical* and *functional* attributes with inner features of brands expressed as *traits of personality* (Keller 1993; Plummer 1985; Batra, Lehmann, and Singh 1993).

Introduction: understanding brand personality

Toyota, Kodak | *Apple, Madza* | *Ford, Sony* | *Hermes, Mercedes* | *Nike, Marlboro*

Introduction: understanding brand personality

- Because company-defined brand personality takes on meanings and subjective interpretations when filtered through the minds of consumers (Ivens and Valta 2012), it is of crucial importance to determine whether the **brand personality communicated** by a company is aligned with **what consumers** actually **perceive**.
- In light of the potential risks described above, we believe that it is essential to determine **how alignment between company-defined and consumer-perceived brand personality can be measured and evaluated**.

Introduction: understanding brand personality

- How can the **degree of alignment between company-defined vs. consumer-perceived brand personality** (i.e., consumer-brand alignment) be measured?
- How **can similarity in brand personality between brands** (i.e., inter-brand alignment) be measured?
- How can c**onsumer perception of similarity in personality between brands** (i.e., consumer-interbrand alignment) be measured?

Methodology
Data collection and sample construction

To address the research questions, it was necessary to collect textual data from two different sources:

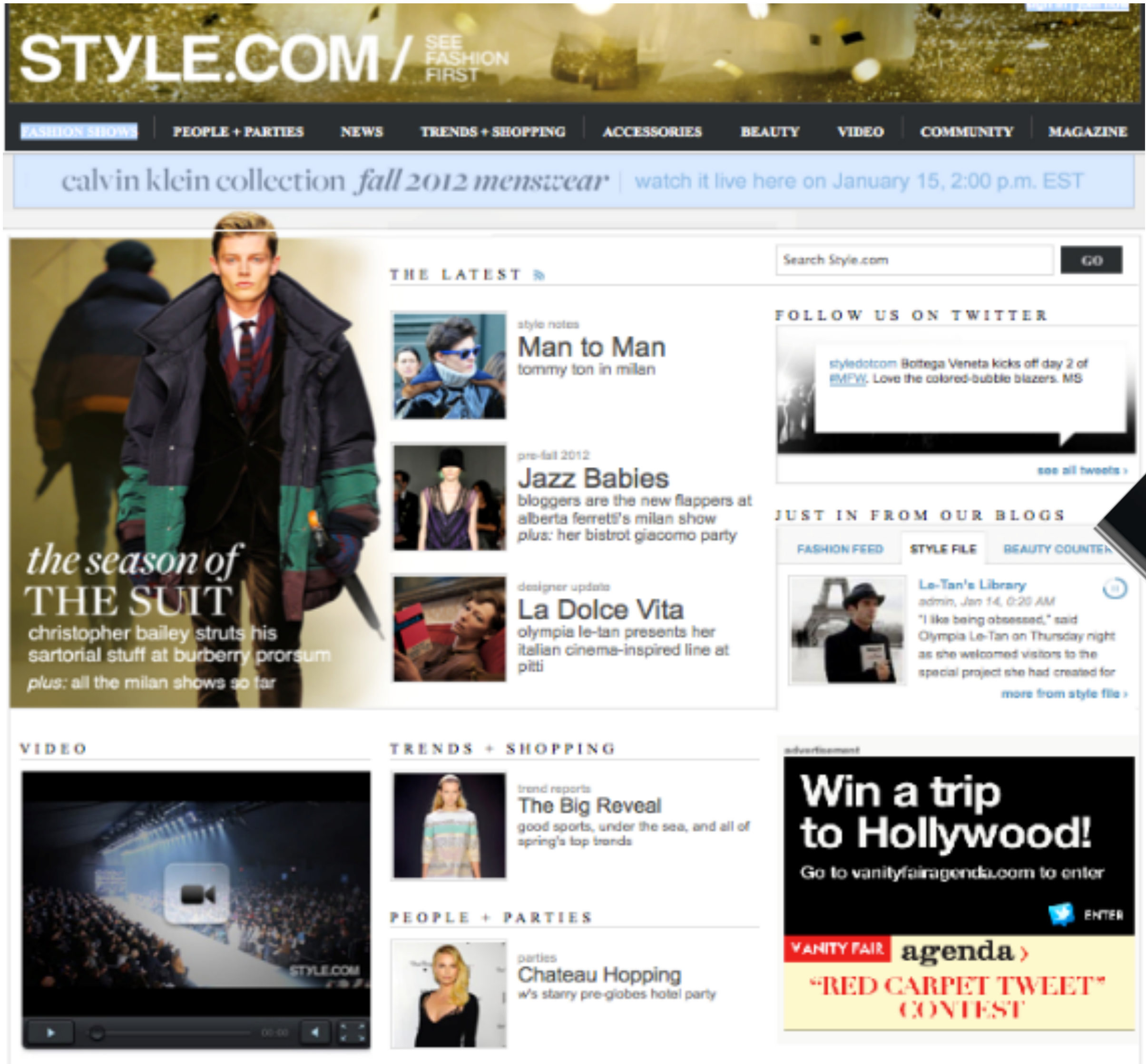
1)a **popular blog** that is extensively used by the online community of fashion consumers to exchange opinions and perceptions about brands (the *blog dataset*)

2) the **websites** and/or **Facebook pages** of fashion companies that offer promotional descriptions of their brands (the *company dataset*).

Methodology
Data collection and sample construction

- The blog dataset was compiled from **Style.com**
- Among the thousands of fashion blogs present in the blogosphere, **Style.com** holds a high ranking in terms of Alex traffic data, membership and incoming links (cf. Bardzell et al. 2009)
- The **posts/comments** are archived for a relatively long period of time (up to approximately four years)
- The corresponding texts were collected into separate files representing **335** different fashion brands. The posts and comments that compile the blog dataset covered a timeframe spanning from August 2008 to August 2013

Methodology
Data collection and sample construction



Methodology
Data collection and sample construction

- We collected blog posts/comments like these

The Man Behind Marilyn's Subway Grate Dress, Giorgio Armani Presents The Paris Photo Festival, Hermès Makes Saris, And More…

October 11, 2011 11:43 am

William Travilla is not a designer most people are familiar with, despite having designed many of Marilyn Monroe's most memorable dresses, including her white halter-neck dress for *The Seven Year Itch*. In a new book, *Dressing Marilyn*, the public gets a rare glimpse into Monroe's costumes and Travilla's process of creating them. [*Vogue U.K.*]

Giorgio Armani is known for doing water-inspired collections, so naturally, he will show an exhibition at the Paris Photo Festival called *Acqua*, showcasing water-themed artwork. Armani is also the official sponsor the for the 15th annual affair taking place in November. [*WWD*]

Hermès has created a line of silk saris with its trademark prints for the standalone stores they are launching in India. "It is symbolic of the relationship of the brand to India, to be Indian in India and to share some activities with our customers," says Hermès chief executive officer Patrick Thomas. [*Grazia Daily*]

DESIGNER UPDATE, Q&A

Everyone's A Critic

February 25, 2011 8:01 pm

Well, at Dolce & Gabbana's Fall 2011 womenswear show, at least. Thanks to a new, dedicated Wi-Fi network at their Milan venue, the Metropol, audience members at Sunday's runway event will be able to log onto a customized Web page (previewed at left) and comment on the action in real time. IPad to the ready, Bryanboy! Comments will stream along with the show on monitors above the catwalk and on the label's online live-stream. (Which, by the way, will be visible right here on Style.com, Sunday at 8 a.m. EST.) The designers have shown quite a willingness to embrace technology the last few years, whether by inviting bloggers into their front row or going full-throttle on Twitter—hello @stefanogabbana—and the new comment system is their latest foray into the digital realm. "We wanted to find a new way to get an immediate and spontaneous feedback to the collection and also a different way to allow people inside the hall to interact among themselves," Domenico Dolce and Stefano Gabbana told Style.com. Comments can also be posted on the brand's Facebook page and made via Twitter. "At the end of the day," the designers said, "what matters more for us is what people think." Now, showgoers, you can think out loud.

Photo: Courtesy of Dolce & Gabbana

Methodology
Data collection and sample construction

Tory Hits Seoul, DVF Hits Vienna, Stefano Rides In Style, And More...

June 30, 2010, 11:16 am

Tory Burch has opened her largest store to date—in Seoul. South Korea, brace for Reva fever in 5, 4, 3, 2... [WWD]

Jean Paul Gaultier, Paul Smith, Giorgio Armani, and Vivienne Westwood are among the designers who have signed up to costume Snow White and the Seven Designers, a pantomime show that hits London this October. The seven designers of the title—Dapper, Snappy, Snazzy, Natty, Classy, Dizzy, and Taupey—will fight, according to Vogue U.K., "ugly interiors." Taupey to the rescue! [Vogue U.K.]

Diane von Furstenberg may be headed to Vienna for this year's Life Ball, but she's bringing a touch of NYC with her. She's arranged for Radio City's Rockettes to perform at the fête, clad in DVF rompers from her recent Resort collection (pictured). [WWD]

More intrigue at T: New editor in chief Sally Singer is said to be bringing in her own fashion director, a perceived slight to longtime T staffer (and former editor in chief candidate) Anne Christensen. [Gatecrasher]

And thank God for Twitter, without which we might never know that Stefano Gabbana is now riding around town on a brand-new leopard-print Vespa. [@giampaolosegura via Refinery29]

Photo: Courtesy of Diane von Furstenberg

We copied blog posts/comments in a word text file and we created a fashion blog corpus, containing the subcorpora (word file texts), one for each fashion brand.

Yea, Nay, Or Eh? Giorgio Girls

April 27, 2010, 10:30 am

Iron Man himself may be a vermilion shade of superhero, but at last night's L.A. premiere of Iron Man 2, Valentino red wasn't on the agenda. The film's two leading ladies, Gwyneth Paltrow (reprising her role from the first flick) and Scarlett Johansson, each opted for a cool, wintry white. Armani was the man of the evening: Johansson picked an undulating Armani Privé cocktail gown, while Paltrow went hard-edged in a Giorgio Armani tailored blazer and shorts suit with jet black accessories. So who wore it better? Are you feeling ScarJo's marquee glamour, or do you prefer Gwynnie's city-girl chic?

Photos: Matt Sayles / AP Photo (Scarlett); Kyle Rover / Startraks Photo (Gwyneth)

Methodology
Data collection and sample construction

- During this process, we discovered that many of the less well-known brands did not have **websites** or **Facebook pages.**
- There were also a number of **websites with blocked text** that could not be copied and pasted into external files.
- Some company websites contained mostly **images without verbal descriptions** of brands or products, and a few had textual material in languages other than English.
- All of these brands were eliminated from the sample as it was essential to have strict matching between the blog and the company datasets for our research aims.
- Given the importance of the interactional dimension in online consumer communities, we also decided to r**emove from the sample all the brands whose files did not contain any user comments in response to the initial posts**, i.e., where consumers failed to engage in ‘conversations’ about the brands.
- Finally, we eliminated brands whose corresponding blog text files did not contain a sufficient amount of text (<1000 words).
- After the various phases of filtering described above, there were **113 fashion brands** represented across the two parallel datasets.

Methodology
Data extraction

The text files contained in both the blog and company datasets were submitted to a series of procedures to systematically analyze adjectives as the linguistic expression of brand personality.

In the study we assume that **perceptions of brand personality** can emerge from:

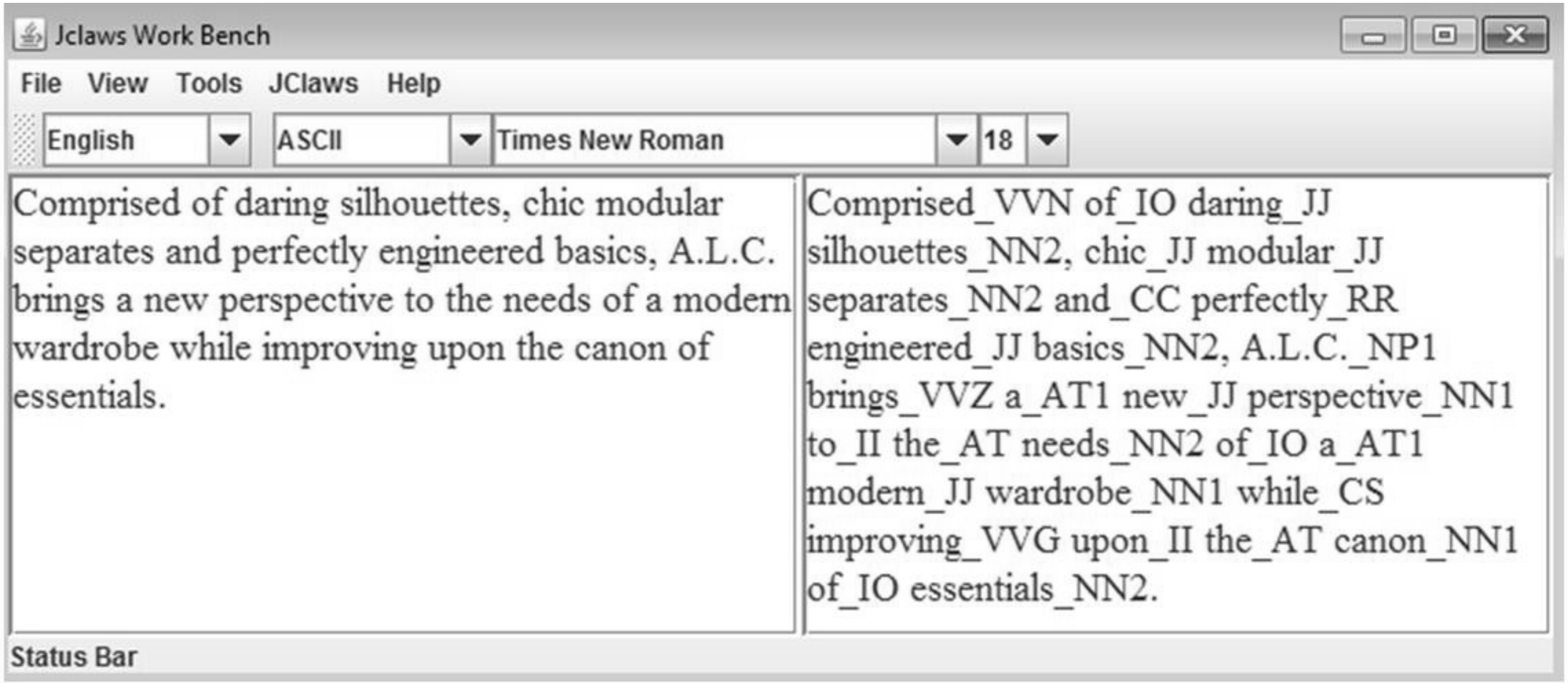
1) the **adjectives** used in **texts** produced by **consumers** during spontaneous online interactions to exchange opinions about brands, and

2) the **adjectives** found in **texts** produced by **fashion companies** through which they define the personalities of their brands.

Methodology
Data extraction

- The 113 files contained in the **company dataset** were first run through the **CLAWS4** (Constituent Likelihood Automatic Word Tagging System) part-of-speech tagger. This software automatically (a) identifies the part-of-speech of each word and (b) tags it accordingly.
- The CLAWS tagger was developed by UCREL (University Centre for Computer Corpus Research on Language) of Lancaster University (UK) and is described as having an accuracy rate of approximately 95%.

Methodology
Data extraction



Tag legend: VVN=past participle lexical verb, IO=preposition of, JJ=general adjective, NN2=plural common noun, CC=conjunction, RR=general adverb, NP1=singular proper noun, VVZ= lexical verb, AT1=singular article, NN1=singular common noun, II=general preposition, AT=article, CS=subordinating conjunction, VVG=-ing participle of lexical verb

Methodology
Data extraction

- The tagged files were then processed with the text analysis software suite *WordSmith Tools* (Scott 2010) to automatically **retrieve** and **analyze** all the adjectives across the files by entering the general adjective tag (JJ) as the search item. The initial **JJ adjective** tag search retrieved **17,347** items across the company dataset files
- First of all, because we were interested only in adjectives that companies used to convey aspects of brand personality, it was necessary to **remove all neutral adjectives**
- We removed all adjectives that were used in merely descriptive contexts, for example, those relating to **color, size/dimension, shape and nationality**.

Methodology - Data extraction

C

File Edit View Compute Settings Windows Help

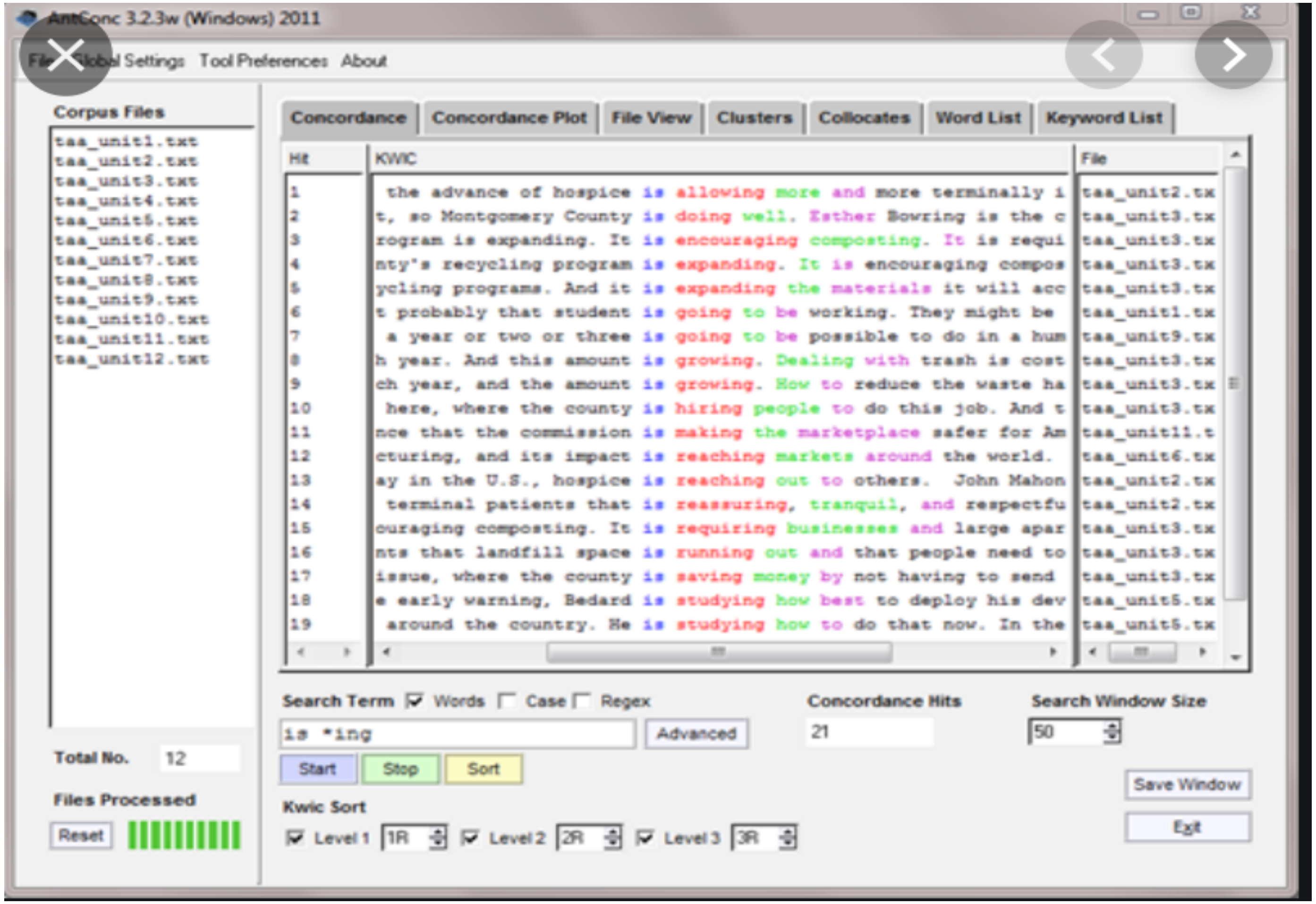
N	Concordance	File
1	iconic_JJ style_NN1 accessible_JJ to_II all_DB ,_, Rachel_NP1	g\rachel zoe.cls
2	of_IO her_PPHO1 highly_RR acclaimed_JJ ready-to-wear_JJ collection_NN1 ._.	g\vera wang.cls
3	to_TO design_VVI for_IF the_AT active_JJ luxury_JJ brand_NN1 Bogner_NN1 ._.	tim coppens.cls
4	-_- Collection_NN1 very_RG active_JJ ,_, sporty_JJ ,_, nothing_PN1 to_TO	e westwood.cls
5	their_APPGE personal_JJ aesthetic_JJ sensibilities_NN2 and_CC broader_JJR	timo weiland.cls
6	for_IF a_AT1 sophisticated_JJ aesthetic_JJ at_II an_AT1 attainable_JJ price_NN1	g\tory burch.cls
7	Hilfiger_NP1 's_GE iconic_JJ aesthetic_JJ by_II fusing_VVG traditional_JJ	mmy hilfiger.cls
8	own_DA unique_JJ aesthetic_JJ and_CC perspective_NN1 to_II the_AT	\rag & bone.cls
9	._. The_AT VPL_NP1 aesthetic_JJ remains_NN2 rooted_VVN in_II the_AT	de 3 tag\vpl.cls
10	a_AT1 D.I.Y_NP1 ._. aesthetic_JJ ._. The_AT media_NN called_VVN	e westwood.cls
11	With_IW a_AT1 modern_JJ aesthetic_JJ and_CC a_AT1 frequent_JJ play_NN1	a mccartney.cls
12	with_II33 the_AT modern_JJ aesthetic_JJ of_IO the_AT brand_NN1 -_- there_EX	ace comm b.cls
13	clean_JJ and_CC edgy_JJ aesthetic_JJ ._. I_PPIS1 love_VV0 the_AT	de 3 tag\tibi.cls
14	's_VBZ signature_NN1 aesthetic_JJ can_VM be_VBI found_VVN in_II	igal azrouël.cls
15	'_GE unique_JJ aesthetic_JJ was_VBDZ captured_VVN in_II a_AT1	g\theyskens.cls
16	romantic_JJ in_II an_AT1 airy_JJ ,_, buoyant_JJ manner_NN1 ._.	tino comm b.cls
17	to_TO capture_VVI such_DA amazing_JJ expressions_NN2 and_CC	enty8twelve.cls
18	._. Two_MC of_IO our_APPGE amazing_JJ summer_NNT1 shoes_NN2 make_VV0	aria cornejo.cls
19	,_, but_CCB there_EX are_VBR amazing_JJ colors_NN2 and_CC nuances_NN2	tag\thakoon.cls
20	miss_VVI all_DB these_DD2 amazing_JJ new_JJ colors_NN2 and_CC	g\vera wang.cls
21	for_IF all_DB of_IO the_AT amazing_JJ new_JJ arrivals_NN2 !_! "_" If_CS	g\rachel roy.cls
22	all_DB manner_NN1 of_IO appealing_JJ embellished_JJ looks_NN2 ._. Here_RL	tino comm b.cls
23	perception_NN1 of_IO architectural_JJ spaces_NN2 ._. At_II Hangar_NN1	berto cavalli.cls
24	,_, modern_JJ ,_, architectural_JJ ,_, minimal_JJ ,_, feminine_JJ ,_,	aria cornejo.cls
25	,_, extracted_VVN from_II architectural_JJ quality_NN1 of_IO vintage_NN1	de 3 tag\vpl.cls

Methodology - Data extraction

- These lists were used to identify the adjectives that were also present in the *blog dataset* as a way to determine **alignments** in the perception of **brand personality among fashion consumers**
- This process was facilitated and rendered systematic through the use of **another software** application for text analysis, i.e., *AntConc* (Anthony 2011), which is able to perform automatic searches on multiple items within a given text file and then display them in lists
- The overlapping adjectives retrieved, it was then necessary to verify that they were actually used by **bloggers** to express a personality facet of the particular brand in question. In fact, in many cases the adjectives qualified other entities mentioned in the blog post/comment.

Methodology - Data extraction

AntConc (Anthony 2011)



Methodology - Data extraction

This completion of this process allowed us to:

1) identify all the **adjective types** that each company used to express facets of brand personality in its web-based communications, and

2) determine which of **those adjectives** had also been used by the fashion bloggers to express their perceptions of the brand’s personality.

Data analysis - CBA

Consumer-brand alignment (CBA)

The CBA ratio measures the *degree of alignment* between **brand personality as defined** by companies and **as perceived** by the bloggers who represent an online community of fashion consumers. For each brand, raw frequencies of common adjectives between the company and blog datasets were tallied and then normalized as the number of occurrences per 1000 words in each blog file. The higher the CBA ratio, the greater the degree of alignment.

Data analysis - CBA

Company	Adjective types extracted from website (N)	Adjectives expressed in the blog files (N)	CBA ratio
A.L.C. (1234)	chic, contemporary, daring, engineered, modern, modular, new (7)	chic, modern, new (3)	**2.43**
Balenciaga (4898)	airy, antique, assertive, beautiful, bold, chic, clean, contemporary, delicate, demure, different, elegant, enchanting, essential, exquisite, feminine, forceful, fragrant, honest, iconic, impassioned, inaccessible, incisive, innate, iridescent, juvenile, military, modern, new, obscure, old-fashioned, opulent, original, peppery, precious, progressive, provocative, pure, radical, reflective, rigorous, romantic, sensual, sexy, sharp, sharp-edged, singular, soft, strong, unambiguous, unexpected, unique, unruly, urbane, whimsical, wild, youthful (57)	beautiful/2, chic, contemporary, elegant, iridescent, new/2, sexy/3, unexpected (12)	**2.45**
Altuzarra	aesthetic, body-conscious, amazing, beautiful, feminine, handmade, light, new, refined, strong, stunning, sumptuous, young (13)	light, new, strong (3)	0.65
Joie	aesthetic, casual, chic, contemporary, fresh, luxurious, modern, new, soft, sophisticated, timeless, unparalleled (12)	casual, contemporary, chic (3)	1.79
Alexis Mabille	arabesque, attractive, beautiful, captivating, celebrated, chic, cute, distinctive, edgy, elegant, feminine, festive, light, modern-day, natural, new, precious, sharp, Sicilian, sophisticated, trim, unexpected (22)	beautiful, chic/2, light, sharp, sophisticated, trim (7)	1.39

Among the 113 brands, in **22 cases** there were no common adjectives. These were eliminated from the sample and, as a consequence, all subsequent analyses refer to the remaining **91 brands.**

Data analysis - CBA

Company	Adjective types extracted from website (N)	Adjectives expressed in the blog files (N)	CBA ratio
Banana Republic (2424)	accessible, amazing, beautiful, bold, colorful, comfortable, different, first-class, fresh, incredible, modern, new, perfect, timeless (14)	beautiful, new/3 (4)	1.65
Alexander McQueen	acclaimed, contemporary, contrasting, crafted, effortless, embellished, everyday, fine, higher-end, iconic, impeccable, light, recognisable, rich, romantic, traditional (16)	romantic, contemporary (2)	0.25
Alberta Ferretti	accessible, aesthetic, affordable, architectural, bohemian, bold, breezy, bright, chic, classic, clean, colorful, contemporary, cosy, decorative, delicate, demi-couture, different, distinctive, elegant, ethereal, exclusive, fabulous, fashionable, feminine, figure-conscious, floating, fresh, functional, handmade, harmonious, high-end, high-quality, innovative, intuitive, invisible, iridescent, irresistible, lacy, light, luxury, magnificent, masculine, modern, muted, new, precious, precise, prestigious, pure, romantic, seductive, sensual, sensuous, sentimental, simple, sophisticated, special, spectacular, style-conscious, sweet, timeless, unique, urban, vibrant, whimsical (66)	accessible, architectural, bright, chic, colourful, contemporary, different, feminine/2, fresh, innovative, light, modern, new/2, pure, romantic, simple, sophisticated, special, unique, urban/2 (23)	3.17
Azzaro	adventurous, aesthetic, aquatic, aromatic, assertive, astounding, attractive, audacious, authentic, avant-garde, beautiful, bold, brilliant, casual, charismatic, chic, contemporary, cosy, crafted, customized, dazzling, different, distinct, distinctive, easy, elegant, emblematic, enchanting, enigmatic, essential, everyday, exceptional, fascinating, feminine, fine, frank, fresh, functional, glamorous, haute-couture, hedonistic, hesperidean, innate, Italian-style, laid-back, Latin, luxurious, masculine, Mediterranean, natural, new, noble, novel, original, ostentatious, pioneering, poetic, powerful, present-day, pure, quintessential, rare, rebellious, resplendent, revitalizing, sensual, sensuous, sexy, sharp, silky, simple, sleek, slender, smooth, soft, solemn, sparkling, spiced, spicy, suave, sublime, sun-infused, sunny, supple, timeless, trendy, ultimate, unadulterated, unforgettable, unique, vibrant, virile, wild, woody, young-at-heart (95)	chic, different, elegant/4, sleek (7)	3.45

Data analysis - IBA

Interbrand alignment (IBA)

The IBA ratio was calculated as the percentage of *intersecting adjectives* between brands in relation their total number of adjectives. To measure IBA, from the lists that contained the adjectives found in the web-based communications of each company, we identified *sets of intersecting adjectives* types across the brands. The higher the IBA ratio, the higher level of similarity in the brand personality communicated by companies.

Data analysis - IBA

Brands	A.L.C.	Calvin Klein	Christian Lacroix	Gucci	Hardy Amies	Henrik Vibskov	Jason Wu	John Varvatos
A.L.C.	50.00 (0.00)	7.69 (0.18)	**20.00** (0.38)	4.00 (0.53)	**15.79** (0.18)	10.00 (0.19)	**16.00** (0.34)	6.67 (0.02)
Calvin Klein	7.69 (0.18)	50.00 (0.00)	4.55 (0.28)	10.71 (0.50)	12.90 (0.12)	4.55 (0.28)	10.81 (0.01)	7.41 (0.14)
Christian Lacroix	20.00 (0.38)	4.55 (0.61)	50.00 (0.00)	3.13 (0.97)	20.00 (0.75)	33.33 (0.00)	14.29 (0.83)	18.18 (0.42)
Gucci	4.00 (0.53)	10.71 (0.61)	3.13 (0.97)	50.00 (0.00)	6.67 (0.51)	3.13 (0.97)	10.81 (0.53)	3.96 (0.46)
Hardy Amies	15.79 (0.18)	12.90 (0.04)	20.00 (0.75)	6.67 (0.51)	50.00 (0.00)	13.33 (0.50)	16.67 (0.13)	20.00 (0.17)
Henrik Vibskov	10.00 (0.19)	4.55 (0.61)	33.33 (0.03)	3.13 (0.97)	13.33 (0.50)	50.00 (0.00)	9.52 (0.55)	18.18 (0.42)
Jason Wu	16.00 (0.34)	10.81 (0.01)	14.29 (0.83)	10.81 (0.53)	16.67 (0.13)	9.52 (0.55)	50.00 (0.00)	7.69 (0.13)
John Varvatos	6.67 (0.02)	7.41 (0.14)	18.18 (0.42)	3.96 (0.46)	20.00 (0.17)	18.18 (0.42)	7.69 (0.13)	50.00 (0.00)

LOW difference of **Delta IBA** associated with **high IBA** indicate situations of high but also of balanced similarity in how the two company communicate the brand personality

Data analysis - IBA

- The resulting IBA ratios ranged from 0 to 37.5%, allowing us to establish three levels of **IBA** as follows: a) low: <12.5% (81.15% of pairs); b) medium: 12.5-25%; (8.56% of pairs); and c) high: 25-37.5% (0.07% of pairs).
- The values in parentheses that appear under IBA correspond to ΔIBA, that is the difference (in absolute value) between the number of intersecting adjectives in relation to the total number of adjectives for each brand. Low differences associated with high IBA values indicate situations of high, but also of balanced similarity in how the two companies communicate brand personality.
- The IBA ratio of the pair A.L.C. - Joie is 21.5%. A.L.C. and Joie share 4 types of adjectives (*chic, contemporary, modern, new*) out of a total of 19 types of adjectives (7 for A.L.C and 12 for Joie). The ΔIBA is 0.24 as the absolute difference between 0.57 (4/7) and 0.33 (4/12).

Data analysis - CIBA

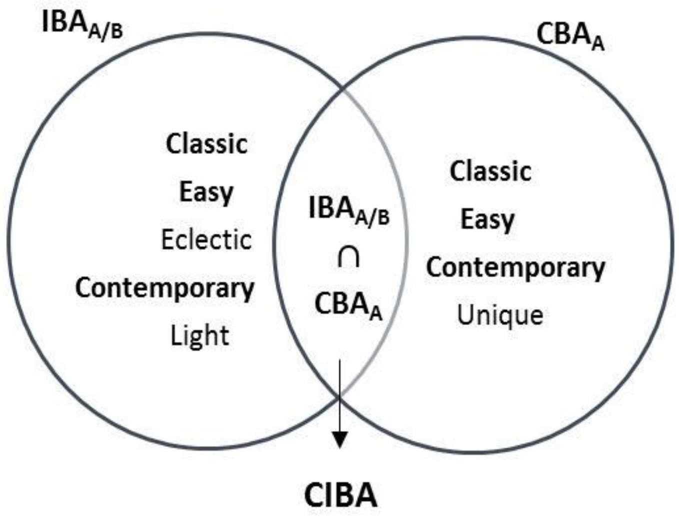
Consumer-interbrand alignment (CIBA)

The CIBA ratio measures the similarity in personality perceived by consumers across brands. **To calculate CIBA, we determined the number the adjectives within the intersecting sets described above (IBA) that were also expressed by consumers in the corresponding blog files (CBA).** These frequencies were again normalized to number of occurrences per 1000 words. This yielded the CIBA ratio which determines how many of the intersecting adjectives across brands (IBA) are also perceived by consumers. The higher the CIBA ratio, the greater the number of intersecting adjectives that form the perceived brand personality. Figure 3 shows a simulation of the CIBA ratio based on two brands.

Data analysis - CIBA

Company$_A$ = **chic**, **elegant**, new	**Company$_B$**= **chic**, unique, **elegant**
Consumer$_A$ = chic (1)	**Consumer$_B$ = chic (3), unique, elegant**
CIBA$_A$= 1/2500*1000 = 0,4	**CIBA$_B$= 4/2000*1000 = 2**
CIBD$_A$= 0	**CIBD$_B$= 1/2000*1000 = 0,5**
CBA$_A$ = 1/2500*1000 = 0,4	**CBA$_B$= 5/2000*1000 = 2,5**

Data analysis - CIBA



$CIBA = IBA_{A/B} \cap CBA_A \rightarrow$
[classic + easy + contemporary] / words in blog file*1000

$CIBD = IBA_{A/B} \not\cap CBA_A \rightarrow$
[unique] / words in blog file*1000

Legend:

- $IBA_{A/B}$ = intersecting adjective types between Brand A and Brand B
- CBA_A = common adjectives between the company and blog datasets of Brand A

As the figure shows, in addition to CIBA, we also distinguished the opposite ratio CIBD (Consumer Interbrand Disalignment), that is, the adjectives perceived by consumers in the corresponding blog files (CBA), but that do not belong to the intersecting sets (IBA), again normalized to number of occurrences per 1000 words

Data analysis - CIBA

	Brand A	Brand B	$IBA_{A/B}$	Δ IBA	$CIBA_A$	$CIBA_B$	$\Delta CIBA_{A-B}$	$CIBD_A$	$CIBD_B$	$\Delta CIBD_{A-B}$	CBA_A	CBA_B	$\Delta\ CBA_{A-B}$
1	A.L.C.	Joie	21.05	0.24	2.43	1.19	1.24	0.00	0.61	-0.61	2.43	1.80	0.63
2	A. Ferretti	Dolce & Gabb.	19.26	0.02	0.96	1.06	-0.10	2.21	1.34	0.87	3.17	2.40	0.77
3	A.Ferretti	Valentino	18.05	0.06	0.82	0.94	-0.12	2.35	1.33	1.02	3.17	2.27	0.90
4	Banana Repub.	Matohu	28.57	0.29	1.65	0.00	1.65	0.00	0.29	-0.29	1.65	0.29	1.35
5	Banana Repub.	Vera Wang	21.21	0.18	1.65	0.31	1.34	0.00	0.31	-0.31	1.65	0.61	1.04
6	Calvin Klein	Joie	22.58	0.21	1.09	0.00	1.09	0.96	1.80	-0.83	2.05	1.80	0.26
7	Calvin Klein	N. Rodriguez	18.91	0.25	1.09	0.80	0.29	0.96	0.13	0.83	2.05	0.93	1.12
8	C. Herrera	Jason Wu	18.70	0.10	0.37	0.35	0.02	0.00	1.32	-1.32	0.38	1.68	-1.30
9	C. Ronson	O. Ceremony	19.30	0.13	0.25	0.07	0.18	0.52	0.29	0.23	0.77	0.36	0.42
10	C. Lacroix	O. Theyskens	21.43	0.73	0.36	0.35	0.00	0.00	0.27	-0.27	0.36	0.62	-0.27

Data analysis - CIBA

- **$CIBA_A$ > $CIBA_B$**: In this situation, the intersecting adjectives between Brand A and B ($IBA_{A/B}$) contribute to consumer-perceived personality more for A than for B. More specifically, for every 1000 words in the corresponding blog files, the number of intersecting adjectives used by consumers when referring to A exceeds those used to refer to B. Thus, the perceived similarity is higher for A than for BSeveral cases illustrate this situation: A.L.C vs. Joie (ΔCIBA = 1.24), Banana Republic vs. Matohu (ΔCIBA = 1.65), Calvin Klein vs. Joie (ΔCIBA= 1.09) and Giorgio Armani vs. Gucci (ΔCIBA= 0.74), which is.
- **$CIBA_B$ > $CIBA_A$**: This represents the opposite situation in which the intersecting adjectives between two brands contribute to consumer-perceived personality more for B than for A. In particular, for every 1000 words in the blog files, there are more intersecting adjectives associated with B than with A, meaning that the perceived similarity is higher for Brand B than for Brand A. From this perspective, we can see that intersecting adjectives characterize more consumers' perception of Rebecca Taylor vs. Hardy Amies (ΔCIBA = -1.68), Tory Burch vs. Peter Som (ΔCIBA = -0.79), Henrik Vibskov vs. Christian Lacroix (ΔCIBA = -0.74).
- **$CIBA_A \approx CIBA_B$**: In this situation, the intersecting adjectives between two brands are used by consumers to refer to their brand personalities in a similar way. Thus, for every 1000 words in the blog files, the intersecting adjectives used by consumers for Brand A and Brand B tend to be the same. This situation of equilibrium characterizes various pairs of brands: Christian Lacroix-Olivier Theyskens (ΔCIBA= 0.0002), Matohu-Jen Kao (ΔCIBA= - 0.02), Henrik Vibskov-John Varvatos (ΔCIBA= 0.02), Hardy Amies - Olivier Theyskens (ΔCIBA= -0.03), Alberta Ferretti - Dolce & Gabbana (ΔCIBA= -0.10). In all of them, $\Delta CIBA_{A-B}$ has a value of close to zero.

Data analysis - CIBD

- **$CIBD_A$>$CIBD_B$**: The non-intersecting adjectives form the perceived personality of A more than B. We can see that they form the perception of Giorgio Armani more than Gucci (ΔCIBD = 1.30), Joie more than Vera Wang (ΔCIBD = 1.49), and Reed Krakoff more than Suno (ΔCIBD = 0.62). In each of these cases, the $CIBD_A$ value is more than twice the value of $CIBD_B$ and there are relatively high positive values of Δ $CIBD_{(A-B)}$.
- **$CIBD_A$ < $CIBD_B$**: The non-intersecting adjectives characterize the perceived personality of B more than A. As can been seen from table 4, they characterize the perception of Pringles of Scotland more than Prada (ΔCIBD = -0.49), Temperly London more than Jason Wu (ΔCIBD = -2.33), and Steven Alan more than Henrik Vibskov (ΔCIBD = -0.28). In each of these, the $CIBD_B$ value is equal to more than twice the value of $CIBD_A$ and there are relatively high negative values of ΔCIBD.
- **$CIBD_A$ ≈ $CIBD_B$**: The non-intersecting adjectives form the perception that consumers have of brands A and B in a similar way. This balanced situation is not very frequent, but can be found in the pair Jason Wu-Joie seen in table 4 (ΔCIBD= 0.07).

Data Interpretation

$CBA_A > CBA_B$
$CIBA_A > CIBA_B$
$CIBD_A \quad CIBD_B$

- **IBA is high**, but the intersecting adjectives form the perceived personality of A more than B ($CIBA_A > CIBA_B$) and $CBA_A > CBA_B$.
- In this case, if the additional adjectives recognized for A are intersecting adjectives *(positive $\Delta CIBA_{A\text{-}B}$)* and are more than the additional *non-intersecting adjectives* recognized for B (negative $\Delta CIBD_{A\text{-}B}$), i.e., $\Delta CIBA_{A\text{-}B} > |\Delta CIBD_{A\text{-}B}|$, then we can infer that A is better able to communicate the intersecting adjectives than B is able to communicate non-intersecting ones. Thus, the differentiating power of A is greater than B and it is based on a *perceived differentiation* (Keller 2012) of adjectives common to A and B.

Data Interpretation (2)

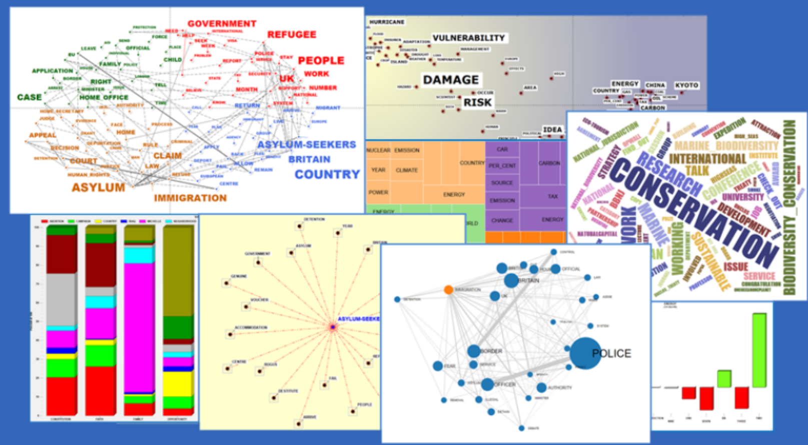
$CBA_A < CBA_B$
$CIBA_A > CIBA_B$
$CIBD_A \quad CIBD_B$

- IBA is high, $CIBA_A$ > $CIBA_B$, but CBA_B>CBA_A. In this case, the adjectives globally expressed for B are more than those expressed for A.
- Even if $CIBA_A$>$CIBA_B$, the additional intersecting adjectives recognized for A (positive $\Delta CIBA_{A\text{-}B}$) are fewer than the additional non-intersecting adjectives recognized for B (negative $\Delta CIBD_{A\text{-}B}$), i.e., $\Delta CIBA_{A\text{-}B} < |\Delta CIBD_{A\text{-}B}|$. We can infer that B can communicate *non-intersecting adjectives* more than A is able to communicate intersecting adjectives. In this case, the differentiating power of B is greater than A, and it is based on an *effective differentiation* (Keller 2012) of brand personality.

Text mining software: T-LAB

https://www.tlab.it/?lang=it

Un ambiente software per l'analisi dei testi e per il text mining completo e facile da utilizzare



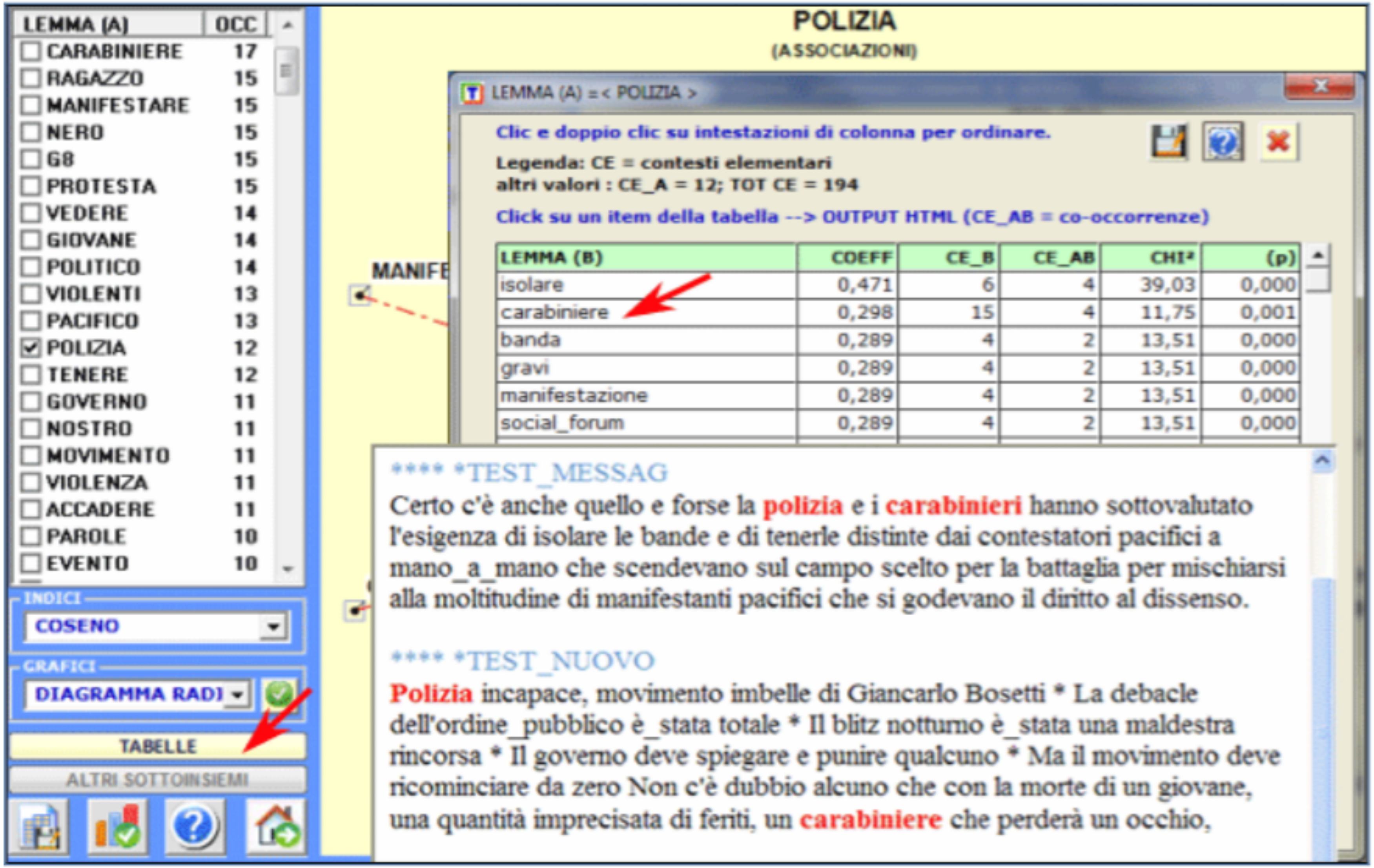
Text mining software: T-LAB

Text mining software: T-LAB

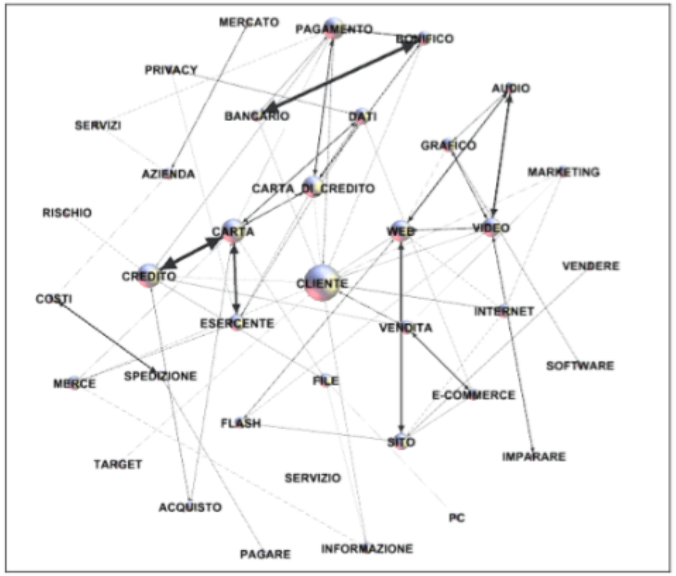
- Associazioni di Parole

Questo strumento **T-LAB** ci consente di verificare come i contesti di **co-occorrenza** determinano il significato locale delle **parole chiave**.

Text mining software: T-LAB



Text mining software: T-LAB



(A) - Associazioni del Primo Ordine

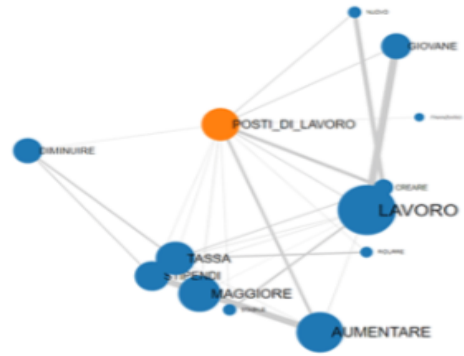
- Analisi delle Sequenze e Network Analysis

Questo strumento **T-LAB** tiene conto delle **posizioni** delle varie unità lessicali all'interno delle frasi e ci permette di rappresentare ed esplorare qualsiasi testo come una **rete** di relazioni.

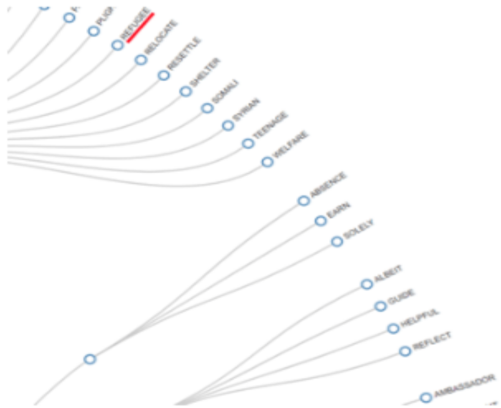
Ciò significa, dopo aver eseguito questo tipo di analisi, l'utilizzatore può verificare le relazioni tra i nodi della rete (cioè le parole chiave) a diversi livelli: a) in relazioni del tipo uno-a-uno; b) all'interno di 'ego network'; c) all'interno delle 'comunità' a cui appartengono; d) all'interno dell'intera rete costituita dal testo in analisi.

RELAZIONI DEL TIPO UNO-AD-UNO

EGO-NETWORK



COMUNITA'

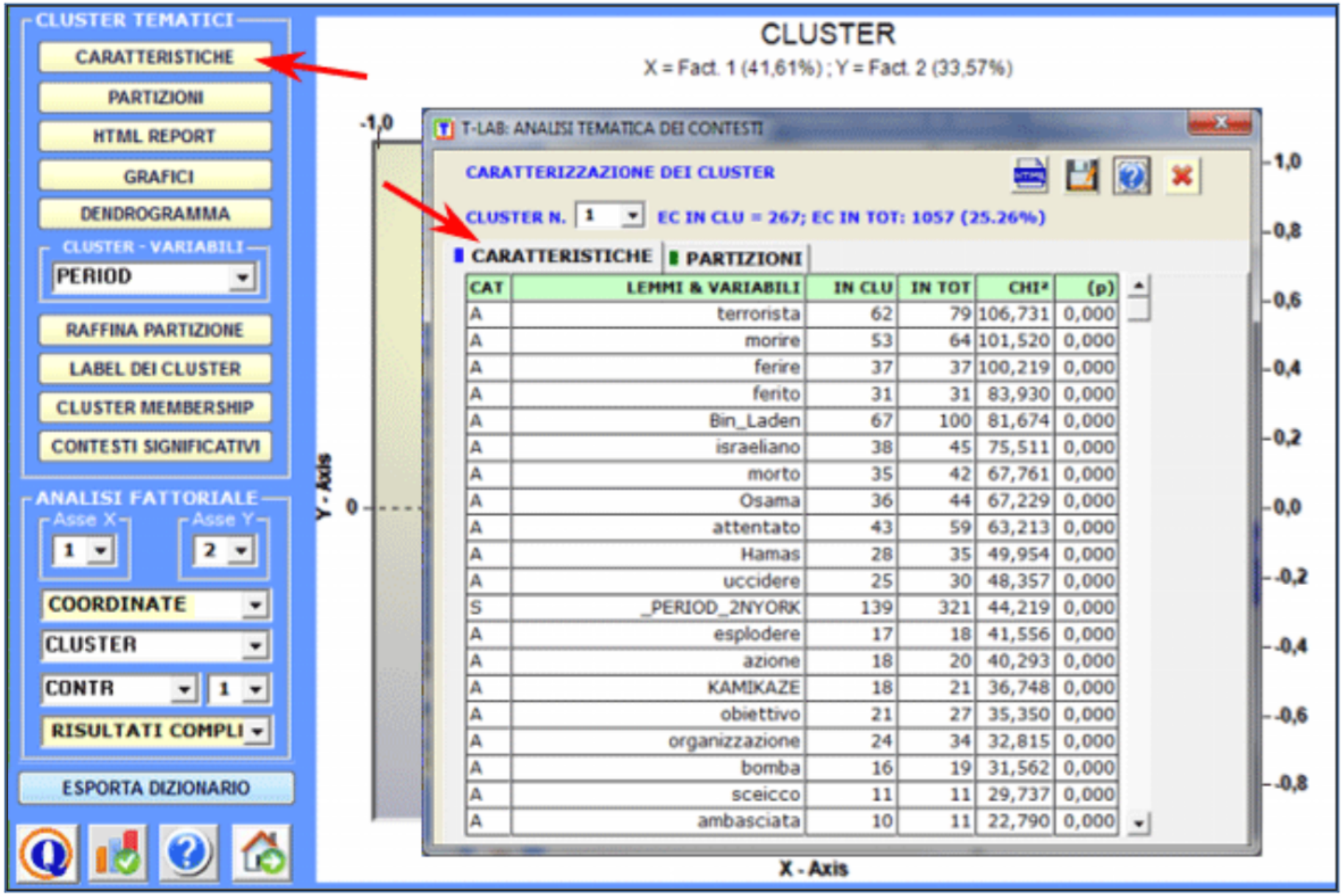


INTERA RETE

Inoltre, facendo clic sull'opzione **GRAPH MAKER**, l'utente può creare diversi tipi di grafici utilizzando elenchi personalizzati di parole chiave (vedi sotto).

Text mining software: T-LAB

- Esplorare le caratteristiche dei cluster



CAT	LEMMI & VARIABILI	IN CLU	IN TOT	CHI²	(p)
A	terrorista	62	79	106,731	0,000
A	morire	53	64	101,520	0,000
A	ferire	37	37	100,219	0,000
A	ferito	31	31	83,930	0,000
A	Bin_Laden	67	100	81,674	0,000
A	israeliano	38	45	75,511	0,000
A	morto	35	42	67,761	0,000
A	Osama	36	44	67,229	0,000
A	attentato	43	59	63,213	0,000
A	Hamas	28	35	49,954	0,000
A	uccidere	25	30	48,357	0,000
S	_PERIOD_2NYORK	139	321	44,219	0,000
A	esplodere	17	18	41,556	0,000
A	azione	18	20	40,293	0,000
A	KAMIKAZE	18	21	36,748	0,000
A	obiettivo	21	27	35,350	0,000
A	organizzazione	24	34	32,815	0,000
A	bomba	16	19	31,562	0,000
A	sceicco	11	11	29,737	0,000
A	ambasciata	10	11	22,790	0,000

Text mining software: T-LAB

LEMMA	CHI SQUARE	WORD	OCC
terrorista	106.731	terrorista	27
terrorista	106.731	terroristi	35
morire	101.52	morendo	1
morire	101.52	morire	5
morire	101.52	morti	34
morire	101.52	morto	4
morire	101.52	muoiono	9
ferire	100.219	ferì	1
ferire	100.219	ferisce	2

SCORE (209.171)

il 9 agosto 2001 un terrorista suicida si fa saltare per_aria nella pizzeria Sbarro, 15 morti, donne, bambini, una famiglia intera di cinque persone; il 4 settembre 20 feriti per un kamikaze davanti all'ospedale Bikur Holim; il 1° dicembre triplo attentato nella via Ben Yehouda, 11 morti e oltre 180 feriti; il 22 gennaio 2001 nella via Jaffa, due morti e 40 feriti;