



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Università degli Studi di Padova
Dipartimento di Studi Linguistici e Letterari

Corso di Laurea Magistrale
in Strategie di Comunicazione
Classe LM-92

Tesi di Laurea

*Analisi dei discorsi dei Papi rivolti alle Università,
da papa Paolo VI a papa Francesco*

Relatore

Prof. Stefano Sbalchiero

Laureando

Nicola Chistè

Matricola n° 2003451/LMSGC

Anno Accademico 2021/2022

INDICE

INTRODUZIONE	7
CAPITOLO 1	
LA STORIA DEI PAPI, IL VATICANO E IL DICASTERO PER LA COMUNICAZIONE	
1.1 Il Papa	11
1.1.1. Il Papa e la sua elezione	11
1.1.2. I Papi non canonici	12
1.1.3. Il Conclave	13
1.2 I Papi eletti nella storia della Chiesa cattolica: statistiche	15
1.3 I Papi e la storia del XX secolo	16
1.4 Papa Francesco	20
1.5 Lo stato della Città del Vaticano	21
1.6 Il Dicastero per la Comunicazione	23
CAPITOLO 2	
RICERCA SOCIALE: DEFINIZIONE, METODOLOGIA, METODI, TECNICHE E APPROCCI	
2.1 La ricerca sociale	25
2.1.1. Definizione generale	25

2.1.2. I paradigmi	28
2.2 Metodologie, metodi, tecniche della ricerca sociale	30
2.2.1 La metodologia	30
2.2.2 I metodi	31
2.2.3 Le tecniche	32
2.3 Gli approcci alla ricerca sociale	32
2.3.1 L'approccio quantitativo o standard	32
2.3.2 L'approccio qualitativo o non standard	36
2.3.3 La gestione dei dati nei due approcci	37
2.3.4 Considerazioni	38
2.3.5 Metodi misti	40

CAPITOLO 3

L'ANALISI DEL CONTENUTO. STORIA, DEFINIZIONI E APPLICAZIONI

3.1 L'analisi del contenuto	43
3.1.1. La storia	43
3.1.2. Le definizioni	46
3.1.3. Il <i>corpus</i>	49
3.2 La codifica e gli approcci classici e moderni	49
3.2.1 Le unità di analisi e le categorie concettuali nel processo di codifica	49
3.2.2 L'analisi del contenuto classica	51
3.2.3 L'analisi del contenuto moderna e il metodo di Reinert	54
3.2.3.1 L'approccio moderno dell'analisi del contenuto	54
3.2.3.2 Il metodo di Reinert	56

CAPITOLO 4

METODOLOGIA E ANALISI DEL *CORPUS* CON IRaMuTeQ

4.1 Metodologia	59
4.1.1. La raccolta dei discorsi dei papi	59
4.1.2. La costruzione del <i>corpus</i>	62
4.1.3. L'analisi del contenuto classica sul <i>corpus</i>	64
4.1.4. Il <i>software</i> IRaMuTeQ	65
4.1.5. Le analisi con il <i>software</i> IRaMuTeQ	65
4.1.5.1 L'analisi lessicometrica	65
4.1.5.2 I mondi lessicali	69
4.1.5.3 Le forme presenti con maggior frequenza nei discorsi	81
4.1.5.4 L'analisi delle corrispondenze	84
4.1.5.5 L'analisi delle similitudini	86
4.1.5.6 La proporzione e l'intensità delle classi semantiche negli anni	89

CONCLUSIONI	95
--------------------	----

FONTI

Bibliografia	99
Risorse web	103

APPENDICI

I nomi di tutti i papi eletti	105
I protagonisti della <i>content analysis</i>	121
L'utilizzo del <i>software</i> Iramuteq	123

INTRODUZIONE

Il lavoro svolto propone una ricerca che si colloca entro la prospettiva della *content analysis* (Sbalchiero, 2018), e si basa su un'analisi del contenuto dei discorsi che i papi, dal 1963 al 2021, hanno rivolto alle Università. I discorsi raccolti si riferiscono a quattro papi, da Papa Paolo VI all'attuale Papa Francesco, con duecentosessantasette discorsi analizzati.

L'analisi del contenuto è un metodo di ricerca quanti-qualitativo che consente di estrapolare informazioni da un insieme di testi, che costituisce il *corpus* di riferimento. Il testo scritto, che spazia dall'opera letteraria all'articolo giornalistico, dal discorso politico ai testi di canzoni, dal colloquio clinico al messaggio pubblicitario, diviene, quindi, il materiale empirico di riferimento per lo studio di determinati fenomeni sociali. In questo senso, le informazioni vengono ricavate dai testi, quindi, dalle parole, ovvero dai dati della presente ricostruzione: "*texts as data*", si potrebbe anche dire altrimenti.

L'obiettivo principale della ricerca è quello di comprendere ed interpretare le informazioni contenute nel *corpus* di riferimento, dapprima attraverso un'analisi quantitativa (basata essenzialmente su un'analisi lessicometrica: frequenze delle parole, occorrenze e co-occorrenze e valutazione della ricchezza lessicale), e di seguito tramite un'analisi che adotta una prospettiva quanti-qualitativa (che si basa sulla comprensione dei contenuti dei testi). Il fine è di individuare, nei discorsi papali, specificità, tematiche di interesse e argomenti ricorrenti anche lungo l'asse temporale, per cercare di comprendere come si pone la chiesa nei confronti del mondo accademico, della scienza e, in generale, della ricerca scientifica.

Uno degli obiettivi di secondo livello, mira a valutare, allo scopo di migliorare, per quanto possibile, la comunicazione e, quindi, le strategie comunicative adottate dai differenti papi nei vari contesti avendo la possibilità di analizzare corpora di grandi dimensioni composti da discorsi pronunciati in contesti istituzionali differenti.

La ricerca si suddivide in quattro capitoli.

Il primo capitolo è di carattere introduttivo e incentrato sul ruolo e sulla storia dei papi presi in considerazione, nonché sul processo che ha condotto alla loro nomina a Vescovi di Roma. È una parte fondamentale che permette di enucleare quegli elementi utili alla

comprensione dei vari discorsi nei loro contesti e co-testi. Una parte approfondisce la storia dello Stato della Città del Vaticano, gli antipapi (i papi non canonici), il conclave e le vicissitudini storiche dall'anno 1922, nomina di Papa Pio XI, fino al 2021 con la figura di Papa Francesco che ha istituito il Dicastero della Comunicazione.

Il secondo capitolo approfondisce la ricerca sociale ed i paradigmi, la metodologia utilizzata nella ricerca scientifica con riferimento a metodi e tecniche. L'ultima parte del secondo capitolo è dedicata agli approcci della ricerca sociale con particolare riferimento agli approcci standard, non standard ed i *mixed method*. Importantissimo quest'ultimo per spiegare e per interpretare con consapevolezza un fenomeno sociale.

Nel terzo capitolo si tratta dell'analisi contenuto e delle varie definizioni date nel corso degli anni, con un *focus* particolare sull'analisi del contenuto classica e moderna. La spiegazione del metodo di Reinert introdurrà la *topic detection* che, attraverso l'analisi delle co-occorrenze del *corpus*, identifica le "classi semantiche", quelle che lui chiama "mondi lessicali".

Il quarto capitolo è metodologico, e affronta le diverse fasi del disegno della ricerca, dalla costruzione del *corpus*, all'analisi dei dati testuali attraverso l'analisi del contenuto con il *software* IRaMuTeQ. Al centro delle analisi è la procedura di costruzione dei *cluster* semantici per l'interpretazione dei mondi lessicali ed i termini legati ai vari *topics*, seguita dall'analisi delle corrispondenze, dall'analisi delle similitudini e dalla valutazione della proporzione e dell'intensità delle classi semantiche negli anni.

Le interpretazioni e la contestualizzazione dei risultati ottenuti chiuderanno il quarto capitolo, dando spazio alle conclusioni finali.

Vale la pena soffermarci sulla scelta del materiale empirico di questa ricerca. La scelta di operare sui discorsi dei papi è stata dettata da esigenze pragmatiche e da una valutazione delle possibilità offerte dal processo di digitalizzazione del materiale scritto. Basti pensare, facendo una ricerca in rete degli eventuali argomenti da analizzare, che tanti sono gli studi, soprattutto recenti, che risultano focalizzati o sulla politica (dai discorsi di propaganda ai post sui social) o sulla pubblicità (messaggi pubblicitari e persuasivi). Poche risultano essere, invece, le ricerche indirizzate all'ambito religioso e, nello specifico, ai discorsi papali. Penso, ad esempio, al caso dei discorsi indirizzati al mondo accademico, affrontati in questa ricerca: essi possono risultare di grande interesse nell'ampissimo dibattito sul ruolo della chiesa, e suoi rapporti con le istituzioni e il

mondo della scienza. Si tratta di discorsi che sono accuratamente strutturati, parimenti ai discorsi dei capi politici visibili sui siti istituzionali di riferimento, sulla pagina internet dedicata al Vaticano www.vatican.va.

CAPITOLO 1

La storia dei Papi, il Vaticano e il Dicastero della Comunicazione

1.1. Il Papa

1.1.1. Il Papa e la sua elezione

«Il Papa è il Vescovo della Diocesi di Roma, è il romano Pontefice e quale successore di Pietro, è il perpetuo e visibile principio e fondamento dell'unità sia dei vescovi sia della moltitudine dei fedeli».¹ È la massima autorità della chiesa cattolica, è chiamato anche sua Santità o Santo Padre: il suo ufficio è il papato e la sua giurisdizione ha il nome di Santa Sede. Il papato è la più antica istituzione del mondo occidentale che è perdurata nei secoli nonostante i conflitti con gli imperatori, gli scismi, le scomuniche, la riforma e la controriforma, la lotta contro l'illuminismo e il mondo moderno. Il termine papa deriva dal greco *pàppas*, cioè padre nell'uso comune, e la parola pontefice significa colui che getta un ponte. Il papa esercita il potere legislativo, giudiziario e amministrativo della Chiesa Cattolica, le funzioni di suprema guida della Chiesa per ciò che riguarda i principi teologici, la morale e la disciplina e, in materia di dottrina e di morale, le sue affermazioni godono del privilegio dell'infalibilità.

Attualmente il papa viene eletto dal Consiglio dei Cardinali attraverso il conclave e di fatto il papa è sempre uno dei cardinali che partecipa all'elezione. La votazione è valida e quindi canonica, con i due terzi dei voti favorevoli dei cardinali presenti e viene eseguita tramite voto segreto. La fumata bianca, dal tetto della Cappella Sistina, è il segnale che comunica l'avvenuta elezione del nuovo papa. Numerose però sono state le trasformazioni nelle modalità di elezione del Santo Padre nei vari secoli. Inizialmente era eletto dal popolo e dal clero romano e per alcuni anni è stato anche nominato direttamente dall'Imperatore romano. Con il passare dei secoli la componente laica fu eliminata dalla parte elettiva e dal 1059 l'elettorato attivo è composto solo dai cardinali e dai vescovi.

¹ Catechismo della Chiesa cattolica, parte prima, capitolo terzo, articolo 9, paragrafo 4, numero 882
https://www.vatican.va/archive/catechism_it/p123a9p4_it.htm

Nel 1270 Papa Gregorio X II istituì il Conclave e una serie di norme molto restrittive (*Ubi Periculum*), a seguito dei fatti successi a Viterbo (sede vacante del papato) dove il tempo di elezione papale perdurò per due anni. Il regolamento stabiliva che i cardinali dovevano riunirsi in un'area chiusa e il cibo veniva loro somministrato solo attraverso una piccola finestra. Se dopo tre giorni i cardinali non eleggevano il papa avrebbero ricevuto una sola portata di cibo al giorno e dopo cinque giorni, avrebbero avuto soltanto pane, vino e acqua. Durante il conclave i cardinali non potevano ricevere nessuna rendita ecclesiastica. Nel 1492 venne celebrato il primo conclave nella Cappella Sistina e dal 1878 Roma divenne sede fissa. Dal 1970 ai cardinali che superano gli ottant'anni è stato tolto il diritto di voto.

1.1.2. I Papi non canonici

Dei duecentosessantasei papi eletti dalla Chiesa Cattolica, circa quaranta sono stati i papi non canonici. Si definiscono non canonici o antipapa quei papi che sono stati eletti dalla Chiesa Cattolica ma non attraverso le procedure ufficiali del diritto canonico². I papi non canonici erano spesso a capo di una delegazione che insorgeva contro il pontefice e che faceva nascere uno scisma nella chiesa cattolica. Spesso erano l'espressione di cardinali o vescovi che volevano liberarsi dall'influenza politica del papa ufficiale. Il primo antipapa è stato Ippolito in contrasto a Callisto nel 217-235 d.c. e uno dei più longevi anti papati è stato quello di Anacleto II, dal 1130 sino al 1138 eletto in contrasto a Innocenzo II, il vero papa. L'ultimo antipapa ufficiale è stato Felice V nel 1449. Spesso gli antipapi ma soprattutto gli scismi della Chiesa Cattolica, nascevano dopo i concili ecumenici che avvenivano per cercare di riunire tutti i cristiani. È bene ricordare che nei periodi in cui professavano due papi non era evidente quale fosse l'antipapa e quale il pontefice ufficiale. Anche nel secolo scorso ci furono degli antipapi, eletti per autoproclamazione o da un gruppo ristretto di adepti. La Chiesa cattolica definisce questi papi non canonici come degli eretici.

² Ordinamento normativo della Chiesa cattolica, che regola l'attività dei fedeli e delle strutture ecclesiastiche nel mondo, nonché le relazioni inter-ecclesiastiche e quelle con la società esterna. Fonte <https://www.treccani.it/enciclopedia/diritto-canonico/>

1.1.3. Il Conclave

Il conclave indica sia la sala dove i cardinali si riuniscono per l'elezione del nuovo papa sia la riunione stessa, deriva dal latino *cum clave*, che significa chiuso con la chiave o sottochiave. Il termine nasce da un evento storico accaduto a Viterbo nel 1270, allora sede papale, dove gli abitanti, stanchi di aspettare l'elezione del nuovo papa (quasi due anni), rinchiusero a chiave i cardinali nel palazzo papale e tolsero parte del tetto per obbligarli a decidere in fretta. Il nuovo pontefice eletto fu Gregorio X cent'ottantaquattresimo papa della chiesa cattolica.

Le procedure del conclave negli anni hanno subito molte variazioni e attualmente sono regolate dalla costituzione apostolica di Giovanni Paolo II del 1996 denominata *Universi Dominici Gregis*. Il regolamento prevede che il giorno di inizio del conclave, tutti i cardinali si riuniscano nella basilica di San Pietro, per assistere ad una messa e nel pomeriggio, con una pubblica processione, si rechino presso la Cappella Sistina (luogo adibito alle votazioni del nuovo papa) dove viene pronunciato il giuramento dal più anziano dei cardinali³:

«Nos omnes et singuli in hac electione Summi Pontificis versantes Cardinales electores promittimus, vovemus et iuramus inviolate et ad unguem Nos esse fideliter et diligenter observaturos omnia quae continentur in Constitutione Apostolica Summi Pontificis Ioannis Pauli II, quae a verbis «Universi Dominici Gregis» incipit, data die XXII mensis Februarii anno MCMXCVI. Item promittimus, vovemus et iuramus, quicumque nostrum, Deo sic disponente, Romanus Pontifex erit electus, eum munus Petrinum Pastoris Ecclesiae universae fideliter exsecuturum esse atque spiritualia et temporalia iura libertatemque Sanctae Sedis integre ac strenue asserere atque tueri numquam esse destitutum. Praecipue autem promittimus et iuramus Nos religiosissime et quoad cunctos, sive clericos sive laicos, secretum esse servaturos de iis omnibus, quae ad electionem Romani Pontificis quomodolibet pertinent, et de iis, quae in loco electionis aguntur, scrutinium directe vel indirecte respicientibus; neque idem secretum quoquo modo violaturos sive perdurante novi Pontificis electione, sive etiam post, nisi expressa facultas ab eodem Pontifice tributa sit, itemque nulli consensioni, dissensioni, alique cuilibet

³ Capitolo terzo, articolo 53 Costituzione apostolica *Universi Dominici Gregis*
https://www.vatican.va/content/john-paul-ii/it/apost_constitutions/documents/hf_jp-ii_apc_22021996_universi-dominici-gregis.html

intercessioni, quibus auctoritates saeculares cuiuslibet ordinis et gradus, vel quivis hominum coetus vel personae singulae voluerint sese Pontificis electioni immiscere, auxilium vel favorem praestatueros.» Tradotto dal latino il giuramento recita: «Noi tutti e singoli Cardinali elettori presenti in quest'elezione del Sommo Pontefice promettiamo, ci obblighiamo e giuriamo di osservare fedelmente e scrupolosamente tutte le prescrizioni contenute nella Costituzione apostolica del Sommo Pontefice Giovanni Paolo II, *Universi Dominici Gregis*, emanata il 22 febbraio 1996. Parimenti, promettiamo, ci obblighiamo e giuriamo che chiunque di noi, per divina disposizione, sia eletto Romano Pontefice, si impegnerà a svolgere fedelmente il munus Petrinum di pastore della Chiesa universale e non mancherà di affermare e difendere strenuamente i diritti spirituali e temporali, nonché la libertà della Santa Sede. Soprattutto, promettiamo e giuriamo di osservare con la massima fedeltà e con tutti, sia chierici che laici, il segreto su tutto ciò che in qualsiasi modo riguarda l'elezione del Romano Pontefice e su ciò che avviene sul luogo dell'elezione, concernente direttamente o indirettamente lo scrutinio; di non violare in alcun modo questo segreto sia durante sia dopo l'elezione del nuovo Pontefice, a meno che non ne sia stata concessa esplicita autorizzazione dallo stesso Pontefice; di non prestare mai appoggio o favore a qualsiasi interferenza, opposizione o altra qualsiasi forma di intervento con cui autorità secolari di qualunque ordine e grado, o qualunque gruppo di persone o singoli volessero ingerirsi nell'elezione del Romano Pontefice».

Terminato il giuramento viene pronunciata da parte del maestro delle celebrazioni liturgiche pontificie la formula: «*Extra omnes*», il «*Tutti fuori*». Rimarranno quindi nella sala chiusa a chiave solo i cardinali elettori e nessuno potrà né entrare né uscire dopo questo momento (se non per motivi di malattia).

Per tutta la durata del conclave è vietato comunicare con persone al di fuori dei cardinali elettori (via telefono, posta o altri mezzi di comunicazione tradizionale o digitale) e il segreto delle modalità di elezioni pontificie dovrà essere mantenuto tale per sempre, anche dopo il termine dell'elezione del nuovo papa.

Gli scrutini per l'elezione del papa sono quattro, due al mattino e due al pomeriggio, e ciascun scrutinio si divide in tre fasi: l'*Antescrutinium* dove i cerimonieri distribuiscono le schede elettorali a ciascun cardinale, lo *Scrutinium vere proprieque* con la deposizione da parte dei cardinali della scheda nell'urna posta sotto all'affresco del *Giudizio Universale* di Michelangelo e il *Post-scrutinium* che comprende la conta delle schede, la

verifica della validità del voto (con almeno i due terzi dei cardinali presenti) e la distruzione delle schede nella stufa della Cappella Sistina. Se il quorum non viene raggiunto si passa subito ad una nuova votazione. Se invece un cardinale riceve i voti necessari, l'elezione a pontefice è canonicamente valida. Il decano dei cardinali si rivolge all'eletto chiedendogli se accetta la sua elezione e il nome con cui vorrebbe essere chiamato. Dopo l'accettazione a nuovo pontefice, alle schede che vengono bruciate nella stufa, viene aggiunto un mix di polveri per creare la tipica fumata bianca che esce dal camino della Cappella Sistina (o nera in caso di non elezione). Il nuovo papa si ritira quindi nella sagrestia per indossare le nuove vesti pontificie e si presenta dalla loggia della basilica di San Pietro. La sua prima presentazione in pubblico è preceduta dall'annuncio dell'*Habemus papam* e dalla croce astile, che dà la solenne benedizione *Urbi et Orbi* (a Roma e al mondo). Nella tabella che segue (Tabella 01) i nomi e gli scrutini dei conclavi degli ultimi cento anni.

Anno	Papa eletto	Durata (giorni)	N° scrutini	Cardinali presenti
1922	Pio XI	4	14	53/60
1939	Pio XII	2	3	62/62
1958	Giovanni XXIII	4	11	51/53
1963	Paolo VI	3	6	80/82
1978	Giovanni Paolo I	2	4	111/114
1978	Giovanni Paolo II	3	8	111/111
2005	Benedetto XVI	2	4	115/117
2013	Francesco	2	5	115/117

Tabella 01. Elenco dei conclavi degli ultimi cento anni.

1.2 I Papi eletti nella storia della Chiesa Cattolica: statistiche

Mettere su carta una lista ufficiale dei papi che si sono succeduti nei due mila anni di storia dopo la nascita di Cristo non è così semplice come si potrebbe pensare. La prima difficoltà è proprio sul primo papa: alcuni elenchi partono da Pietro, altri invece dal

secondo, Lino, poiché Pietro era un apostolo di Gesù. Pietro è stato comunque Vescovo di Roma e come tale è giusto inserirlo come primo papa.

Un altro problema sorge nei periodi bui del pontificato e cioè quando, oltre al papa canonico, vi erano altri soggetti che reclamavano il ruolo, la storia a riguardo non è sempre stata trasparente a causa anche della mancanza di scritture ufficiali. Altro problema è stato l'utilizzo dei numeri ordinali per richiamare un papa piuttosto che un altro, perché questa operazione non è sempre stata fatta in maniera sistematica. Vi sono quindi delle incertezze su alcuni papi. Prendendo in considerazione l'attuale Annuario e il sito ufficiale della Santa Sede (<https://www.vatican.va/>)⁴ i papi canonici eletti e quindi ufficiali per la Chiesa Cattolica sono stati duecento sessantasei, duecento cinquantanove dei quali hanno pontificato a Roma, gli altri ad Avignone (sede papale dal 1309 al 1377 con la prima elezione di un papa francese, Clemente V, voluta dall'imperatore francese Filippo IV dopo un diverbio con il precedente papa Bonifacio VIII per la richiesta di un elevato pagamento di tasse da parte del clero francese.).

Solo due papi hanno rinunciato alla carica (Celestino V nel 1294 e Benedetto XVI, papa Ratzinger il 28 febbraio 2013), due invece morirono prima di essere ufficializzati (Stefano II e Adriano V). Il mandato papale in media è durato sette anni, il più longevo quello di Pietro, apostolo di Gesù, circa 34 anni (dal 33 al 67). Il nome in assoluto più utilizzato dai papi è stato Giovanni con ventuno vescovi che hanno scelto questo nome, seguito da Gregorio, sedici scelte, Benedetto quindici, Clemente quattordici, Innocenzo e Leone tredici scelte e infine Pio con dodici scelte. Come provenienza l'Europa ha il maggior numero di papi, l'Italia con duecentodiciassette pontefici eletti, sedici francesi e sette germanici. Solo dieci asiatici, tre africani e uno solo dalle Americhe, tutt'ora in carica, Papa Francesco.

1.3 I Papi e la storia del XX secolo

La I Guerra Mondiale in Europa pone fine ad uno dei periodi di pace più duraturi e per il papato è una storia senza precedenti: spogliato dal potere temporale si deve confrontare

⁴ Elenco ufficiale dal sito della Santa Sede
https://www.vatican.va/holy_father/archivio_papi/documents/elencopapi_it.html

con un conflitto che si combatte sotto i suoi stessi occhi. Benedetto XI, papa dal 1914 al 1922 tenta di denunciare e bloccare il suicidio dell'Europa e la carneficina in corso, ma non riesce nel suo intento. Si dichiara fin da subito neutrale ma i suoi interventi diplomatici, oltre che scontrarsi con gli interessi dei vari governi italiani, non riscuotono neppure consensi nel mondo ecclesiastico. Per questo diventa un papa impopolare. Il periodo storico di papa Pio XI (1922-1939) nato Achille Ratti, succeduto a Benedetto XI, non è migliore perché abbraccia la rivoluzione sovietica e la crisi economica degli Stati Uniti. Papa Ratti cerca così di risolvere soprattutto i problemi interni tra Chiesa e Italia. I patti lateranensi tra Stato pontificio e Mussolini risolvono solo in parte le problematiche di quel periodo perché, all'interno della Chiesa alcuni membri del clero e soprattutto laici e cattolici rimproverano a Pio XI di aver fatto un'alleanza con il fascismo da una parte e di aver rinunciato allo stato pontificio come unico amministratore dall'altra. La parte più importante dei patti lateranensi è la convenzione finanziaria: 1 miliardo di lire in titoli di Stato e 750 milioni di versamento da parte dello Stato italiano a favore del Vaticano come parziale compensazione alla perdita dello Stato pontificio. Nel 1931 Pio XI inaugurò Radio Vaticana, emittente che ad oggi diffonde la sua voce in oltre quaranta lingue e più di sessanta nazioni.

Pio XI è ricordato anche come il papa che ha pronunciato nel 1939 la frase in lingua francese "Noi siamo spiritualmente dei semiti" per sancire definitivamente la fine dell'alleanza con il fascismo che si era allineato al nazionalsocialismo tedesco. Pio XI muore improvvisamente nel 1939 per attacco cardiaco, non riuscendo a leggere il suo testamento spirituale in occasione del suo 17° anno di pontificato e 60° di sacerdozio. Si scoprì solo nel 1959 che il suo discorso avrebbe denunciato la violazione dei patti Lateranensi e le persecuzioni razziali in Germania. Il suo testamento è rimasto segreto fino al pontificato di papa Giovanni XXIII quando nel 1959 vennero pubblicate alcune parti. È certo che il testo del discorso fu fatto distruggere dal cardinale Segretario di Stato Pacelli. Lo stesso Pacelli, viene eletto papa Pio XII dal 1939 al 1958 dopo la morte di Pio XI e in pieno conflitto mondiale prosegue la linea dettata da Benedetto XV un suo predecessore: impegno diplomatico per evitare il conflitto e piena neutralità della Santa Sede. Difficile da capire il suo silenzio riguardo le atrocità naziste cui il Vaticano è a conoscenza. Gli anni di Pio XII sono anche gli anni dei media e soprattutto dell'audiovisivo e della radio. Media che, grazie alla nuova rapidità di diffusione delle

informazioni, riescono a far arrivare circa due milioni e cinquecentomila fedeli al Giubileo del 1950 a Roma. Tutto questo contribuisce a presentare il papa come l'unica guida della chiesa cattolica spingendolo così ad intervenire fino alle singole questioni delle chiese locali.

Giovanni XXIII eletto dal 1958 al 1963, il settantatreenne Angelo Roncalli, rappresenta una svolta decisiva nel papato. A differenza di Pio XII ama dialogare con tutti, dai credenti ai non credenti fino alle persone separate, con semplicità, umiltà e rispetto, tutti ricordano le sue visite a bambini orfani e a persone sofferenti, viene chiamato anche per questo "papa buono". Con Giovanni XXIII e, dopo circa un secolo, il papa ricomincia a viaggiare spostandosi da Roma per incontrare i fedeli. Nel gennaio 1959 convoca un concilio ecumenico con i cristiani di tutte le confessioni. La più importante enciclica divulgata, "Pacem in terris"⁵, condanna la guerra che non può essere considerata strumento di giustizia e richiama la collaborazione di tutti gli uomini senza distinzione di religioni per un mondo di pace. La morte di papa Giovanni XXIII viene ricordata come un lutto universale. Paolo VI (1963-1978) succede a papa Giovanni XXIII ma non riscuoterà la stessa popolarità del predecessore. Di particolare importanza, dal punto di vista simbolico e spettacolistico, è l'incontro a Gerusalemme nel 1964 con il patriarca di Costantinopoli Athenagoras I che l'anno successivo si recherà a Roma per ricambiare l'incontro e per annullare definitivamente le scomuniche reciproche pronunciate nel lontano 1054. Paolo VI sostiene spesso le richieste della minoranza della Chiesa, condannando ad esempio la contraccezione e spezzando il fronte compatto della Chiesa. Succede a Paolo VI il pontificato di papa Luciani, Giovanni Paolo I nel 1978, ma dura solo trentatré giorni. Operaio nato a Belluno, persona umile viene ricordato per il suo sorriso empatico ma anche come pontefice poco diplomatico e poco esperto della curia. Alcuni sospetti sulla sua morte parlano di avvelenamento e parte della stampa di quel periodo ha sollecitato la curia ad effettuare l'autopsia sul suo cadavere, negata però dal collegio cardinalizio perché non prevista dal protocollo. La salma di papa Luciani è deposta ancora oggi nelle grotte vaticane e un museo a lui dedicato è stato aperto, dopo la sua morte, nella sua città natale a Canale d'Agordo, in provincia di Belluno. Nello stesso anno, al Conclave dell'ottobre del 1978, dopo otto scrutini e la partecipazione di

⁵ Enciclica completa https://www.vatican.va/content/john-xxiii/it/encyclicals/documents/hf_j-xxiii_enc_11041963_pacem.html

tutti i centoundici cardinali elettori viene eletto papa Giovanni Paolo II, Wojtyła, il papa polacco, il primo pontefice non italiano dal 1523. Dimostra fin da subito una visione moderna, con una duplice modalità di utilizzo del linguaggio, autoritario in seno alla Chiesa (rafforzando il potere gerarchico del clero) ma democratico, a volte populista, per la società civile. Gran viaggiatore, visita il mondo intero inginocchiandosi e baciando il suolo ad ogni suo arrivo. Richiama folle immense ad ogni suo incontro pubblico. Dal 1981 viaggia in una vettura blindata con vetri antisfondamento a causa dell'attentato del 13 maggio in piazza San Pietro a Roma. Giovanni Paolo II inaugura nel 1981 la televisione Vaticana con lo scopo di diffondere ancora di più la parola del Vangelo e nel Natale 1995 sempre sotto il suo pontificato, viene creato e pubblicato il sito internet ufficiale del Vaticano raggiungibile ancor oggi alla pagina www.vatican.va. Giovanni Paolo II muore il 2 aprile del 2005 per aggravamento delle sue condizioni di salute, più di tre milioni di persone raggiungono nei giorni successivi Roma per il suo ultimo saluto. Dopo di lui viene eletto Benedetto XVI, papa Ratzinger il 19 aprile 2005. Duecentosessantacinquesimo papa e settimo pontefice tedesco, nato in Germania a Marktl am Inn nel 1927 figlio di un poliziotto antinazista e di una cuoca. Nel 1941 si arruola nella "Gioventù hitleriana" per obbligo del periodo, ma riesce ad ottenere una dispensa poiché già iscritto al seminario. Durante il suo pontificato molta attenzione pone ai temi della povertà, dell'Africa e dei giovani e lotterà con tutte le sue forze contro la pedofilia del clero inasprendo le norme nei confronti dei preti pedofili. Numerosissime le sue pubblicazioni prima dell'elezione a papa da professore ordinario all'Università di Ratisbona. Anche da papa coltiva il dono della scrittura, rinnovando anche la comunicazione papale. È il primo papa a creare e utilizzare un suo profilo social su *Twitter*. Il 28 febbraio 2013, annuncia la sua volontà a rimettere il suo mandato per mancanza di forze fisiche causate dall'età, con grande sorpresa sia del mondo cattolico, che del clero e del Vaticano stesso. Tiene come casa il Vaticano dove tutt'ora alloggia con i suoi fidati collaboratori.

1.4 Papa Francesco

Jorge Mario Bergoglio, Papa Francesco (in onore a Francesco d'Assisi) è l'attuale pontefice, eletto il 13 marzo 2013 dopo cinque scrutini e due giorni di conclave dei centoquindici cardinali elettori presenti. Nato in Argentina nel 1936 è il primo papa proveniente dall'America e il primo ad appartenere all'ordine religioso dei gesuiti. Fin dalla sua elezione papa Francesco si è distinto per la sua umiltà e la sua semplicità, destando però fra il clero qualche critica. Prima della sua benedizione Urbi et Orbi, ha salutato le persone della piazza di San Pietro con un amichevole "buona sera a tutti" e il giorno seguente al primo incontro con la stampa ha preferito rimanere in piedi fra i giornalisti piuttosto che sedersi sulla lussuosa poltrona a lui riservata.

Suo padre è di origine piemontese (funzionario delle ferrovie dello stato) e sua madre ligure (casalinga). Nel 1928 decisero di partire per Buenos Aires per cercare fortuna. Papa Bergoglio entra in seminario nel 1958, dopo essersi diplomato come chimico ed aver lavorato prima come addetto delle pulizie in una fabbrica e poi come buttafuori in un locale di Cordoba. Si laurea in filosofia nel 1963 e svolge qualche anno di insegnamento di letteratura e teologia divenendo, nel 1979, rettore della Facoltà di teologia e filosofia a San Miguel. Diventa poi direttore spirituale e confessore della chiesa della Compagnia di Gesù di Córdoba dopo un periodo di studi in Germania dove non riesce a terminare la sua tesi di dottorato. Il 20 maggio 1992 papa Giovanni Paolo II lo nomina vescovo di Buenos Aires e nel 1997 è eletto arcivescovo coadiutore di Buenos Aires. Dal 2005 al 2011 è a capo della Conferenza Episcopale Argentina. Anche durante il suo periodo come arcivescovo decide fin da subito di mantenere uno stile di vita sobrio e pacato rinunciando a vivere nella casa del clero e utilizzando i mezzi pubblici per i suoi spostamenti. Ha insegnato ai propri fedeli l'importanza del rispetto per le persone omosessuali anche se si oppone nel 2010 alla legge proposta dalla presidentessa dell'Argentina Cristina Fernández de Kirchner sull'equivalenza del matrimonio eterosessuale e le unioni omosessuali. Per Papa Bergoglio i diritti civili sono distribuiti in maniera disomogenea e i cristiani devono impegnarsi a creare comunità con tutte le persone. Numerose le riforme attuate dall'attuale pontefice, tra le principali ci sono:

- la riforma del Codice penale, con la quale vengono adeguate le leggi del Vaticano alle normative internazionali: è stato abolito l'ergastolo ed è stata introdotta al suo posto una pena da trenta a trentacinque anni di reclusione, vengono introdotti i reati per lo spaccio di stupefacenti, la tortura, i delitti contro i minori, la pedopornografia e la detenzione di materiale pedopornografico, gli atti sessuali con minori, e le pene per i reati di corruzione,
- la riforma economica della Santa Sede con la quale viene istituita nel 2013 una commissione laica per raccogliere informazioni sulle questioni economiche della Santa Sede,
- la condanna dell'uso di armi chimiche nella guerra civile in Siria,
- la tutela dell'ambiente con un invito a tutti i popoli ad aver rispetto per ogni creatura di Dio e per l'ambiente in cui si vive,
- l'istituzione nel 2015, del Dicastero per la Comunicazione e del suo sito ufficiale <https://www.comunicazione.va/it.html>.

1.5 Lo Stato della Città del Vaticano

Lo Stato della Chiesa, che per quasi mille anni era esteso per buona parte dell'Italia centrale è stato annesso allo stato italiano nel 1870 e l'ufficialità dello Stato della Città del Vaticano (chiamato semplicemente anche Vaticano) è avvenuta dopo la firma dei Patti Lateranensi tra il cardinale Pietro Gasparri e Benito Mussolini l'11 febbraio 1929. Il papa era Pio XI e il suo pontificato è stato uno dei più difficili per le questioni politiche ed economiche del dopoguerra (la fine della I Guerra Mondiale, la rivoluzione sovietica e la crisi economica degli Stati Uniti). I Patti Lateranensi, che comprendono due documenti distinti, il trattato e il concordato, stabilirono, con il primo documento che il governo italiano garantiva la sovranità e la libertà del nuovo Stato Vaticano e la Santa Sede riconosceva lo Stato italiano e Roma capitale. Il concordato invece disciplinava i rapporti tra Stato e Chiesa per quel che riguardava le relazioni civili e religiose.

Il Vaticano geograficamente ha ad oggi un'estensione di 0,44 km quadrati e solo quattrocentocinquatré abitanti. È una monarchia assoluta presieduta dal pontefice che ha pieni poteri legislativi, esecutivi e giudiziari. Il papa si avvale però di un Governatorato per

amministrare il territorio del Vaticano e il Cardinale Giuseppe Bertello ne è il capo. Il Segretario di Stato è il Cardinale Pietro Parolin. Lo Stato della Città del Vaticano ha una sua banca, la propria gendarmeria e conia una sua moneta (l'euro a seguito dell'adesione all'Unione Monetaria Europea del 2000). Ha un suo quotidiano, l'Osservatore Romano, una casa editrice, Libreria Editrice Vaticana, gestisce una propria televisione ed una radio. È ad oggi il più piccolo stato sovrano del mondo.

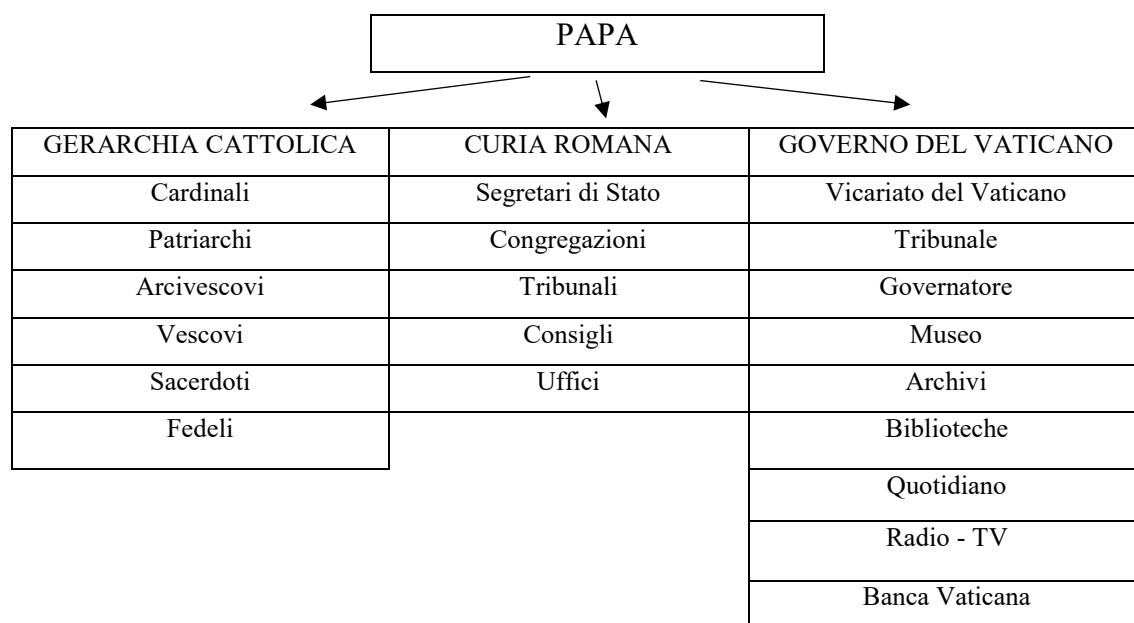


Figura 01. Organigramma dello Stato della Città del Vaticano (fonte sito vaticanstate.va).

Al vertice della gerarchia dello Stato della Città del Vaticano c'è il pontefice, vescovo di Roma che presiede anche il destino della chiesa cattolica. Il papa è circondato dai cardinali, non presenti fisicamente a Roma ma distribuiti territorialmente in tutto il mondo. Possono essere solennemente convocati, per questioni urgenti e di rilevanza per la chiesa cattolica. Sono loro che coordinano le funzioni del papa dopo la sua morte e ne assicurano una nuova elezione. Patriarchi e arcivescovi esercitano poteri regionali ed i vescovi hanno un alto potere ecclesiastico perché presiedono le diocesi di tutto il territorio. La curia romana, che negli anni ha accentrato quasi tutto il potere, segue tutti gli aspetti del governo pontificio, da quello politico e religioso a quello amministrativo e giudiziario. La segreteria di stato occupa una posizione di dominio e si occupa delle relazioni tra papa e la chiesa universale. Le congregazioni esaminano la dottrina della

fede e il buon costume. Il governo del Vaticano riveste un ruolo di secondaria importanza per la Chiesa.

1.6 Il Dicastero per la Comunicazione

Il 27 giugno 2015 Papa Francesco ha istituito il Dicastero per la comunicazione, ente istituzionale della curia romana costituito da cinque direzioni principali: la direzione degli affari generali (cui spetta il compito di gestione del personale, la gestione degli affari comuni delle direzioni e anche gli aspetti legali), la direzione editoriale (che segue gli aspetti relativi al coordinamento dell'editoria vaticana, i piani strategici e gli aspetti dei media digitali), la direzione della sala stampa della Santa Sede (che organizza le conferenze stampa e diffonde i comunicati stampa ufficiali), la direzione tecnologica (che gestisce le piattaforme ed i servizi informatici) e la direzione teologico-pastorale (che promuove la visione teologica ed il messaggio pastorale del pontefice). L'obiettivo di fondo, spiega il Papa nel documento istitutivo, è quello di «rispondere sempre meglio alle esigenze della missione della Chiesa all'interno di un ambito mutevole e rapido come quello dell'informazione multimediale e social, sempre più complessa e interdipendente». Nello specifico, come segnalato dal sito ufficiale <https://www.comunicazione.va>, sono riunite sotto un'unica regia tutte le realtà che si occupano di comunicazione:

- il pontificio consiglio delle comunicazioni sociali che dal 1948 segue e valuta i settori del cinema, della radio, della televisione e della stampa periodica e quotidiana,
- la tipografia poliglotta vaticana, istituita da papa Sisto V nel 1587 che segue la stampa di tutti i documenti della Santa Sede;
- la Libreria Editrice Vaticana che è la casa editrice ufficiale della Santa Sede nata nel 1926,
- l'Osservatore Romano, il quotidiano che dal 1861 segue la vita internazionale ed i dibattiti della Chiesa oltre alle attività pubbliche del pontefice,
- il servizio fotografico che giornalmente, dal 2006, con i suoi professionisti, segue e archivia attraverso servizi fotografici, gli impegni quotidiani svolti dal papa,

- Radio Vaticana, progettata da Guglielmo Marconi, e inaugurata il 12 febbraio 1931 da papa Pio XI che copre oltre sessanta nazioni in tutto il mondo con quaranta lingue e negli ultimi anni viene diffusa anche via web,
- la sala stampa della Santa Sede, che si occupa di diffondere i comunicati stampa della Santa Sede dal 1939 e di organizzare incontri diretti con i giornalisti,
- il centro televisivo Vaticano nato nel 1983 per volontà di Giovanni Paolo II per contribuire alla diffusione universale del Vangelo,
- il servizio Internet Vaticano che cura il sito internet <https://www.vatican.va/> che è online dal giorno di Natale del 1995. Lo stesso giorno viene pubblicato sul sito, per la prima volta, il testo del messaggio natalizio del pontefice Wojtyła.

CAPITOLO 2

Ricerca sociale: definizione, metodologia, metodi, tecniche e approcci

2.1 La ricerca sociale

2.1.1. Definizione generale

«La locuzione ricerca sociale designa un particolare tipo di agire strategico, con il quale il ricercatore si apre a un'esperienza con l'intento di elaborare una risposta (e talvolta più d'una) a una domanda relativa a un dato fenomeno sociale» (Cardano, 2011, p.13 - Ricolfi 1997). La ricerca sociale parte quindi da serie di interrogativi (una o più domande di ricerca) che riguardano aspetti della realtà sociale. L'obiettivo principale è quello di acquisire conoscenza scientifica sui fenomeni della realtà sociale e, contrariamente a quello che spesso viene detto, il fine della ricerca scientifica non è spiegare il reale (che è inconoscibile o conoscibile metafisicamente), ma dare una risposta agli interrogativi sul reale (Boudon, 1980). In secondo luogo, la ricerca sociale mira ad aumentare le conoscenze su un determinato fenomeno e utilizzare il sapere acquisito per dare indicazioni su come intervenire sulla realtà, per migliorarla. La ricerca sociale è poi orientata alla crescita della conoscenza dei comportamenti degli uomini, dei gruppi e delle istituzioni sociali e richiede approccio scientifico, obiettivi differenziati (descrivere fenomeni, comprendere relazioni tra eventi, generalizzare in altri contesti) e necessita di strategie diverse a seconda del contesto e delle condizioni. Infine, utilizza strumenti operativi e coerenti. Per raggiungere gli obiettivi è necessario percorrere un itinerario che parte dal progetto del ricercatore e termina con la restituzione dei risultati. Va anche aggiunto che «controllabilità, ripetibilità e aspetto pubblico delle procedure rappresentano l'unica possibile oggettività della ricerca» (Statera, 1984, p. 250). Si fa poi presente che non esiste una unità di metodo o un metodo migliore, esistono unità di obiettivi e unità di percorso conoscitivo. Quello che distingue sono le tecniche di ricerca e la differenza di strumenti utilizzati per osservare, classificare, analizzare, misurare e collegare gli elementi dello studio. La ricerca, e la ricerca sociale, non avviene nel «vuoto», ma in un contesto definito e ben consolidato. «Per questo motivo è importante conoscerne le caratteristiche principali per posizionarci e collocare le nostre ricerche». (Bryman, 2016).

«Il processo della ricerca sociale segue un processo di circolarità suddiviso in cinque tappe principali» (Bailey 2006, p.15):

- 1) scelta del problema e definizione delle ipotesi,
- 2) formulazione del disegno della ricerca,
- 3) raccolta dati,
- 4) codifica e analisi dei dati,
- 5) interpretazione dei risultati.

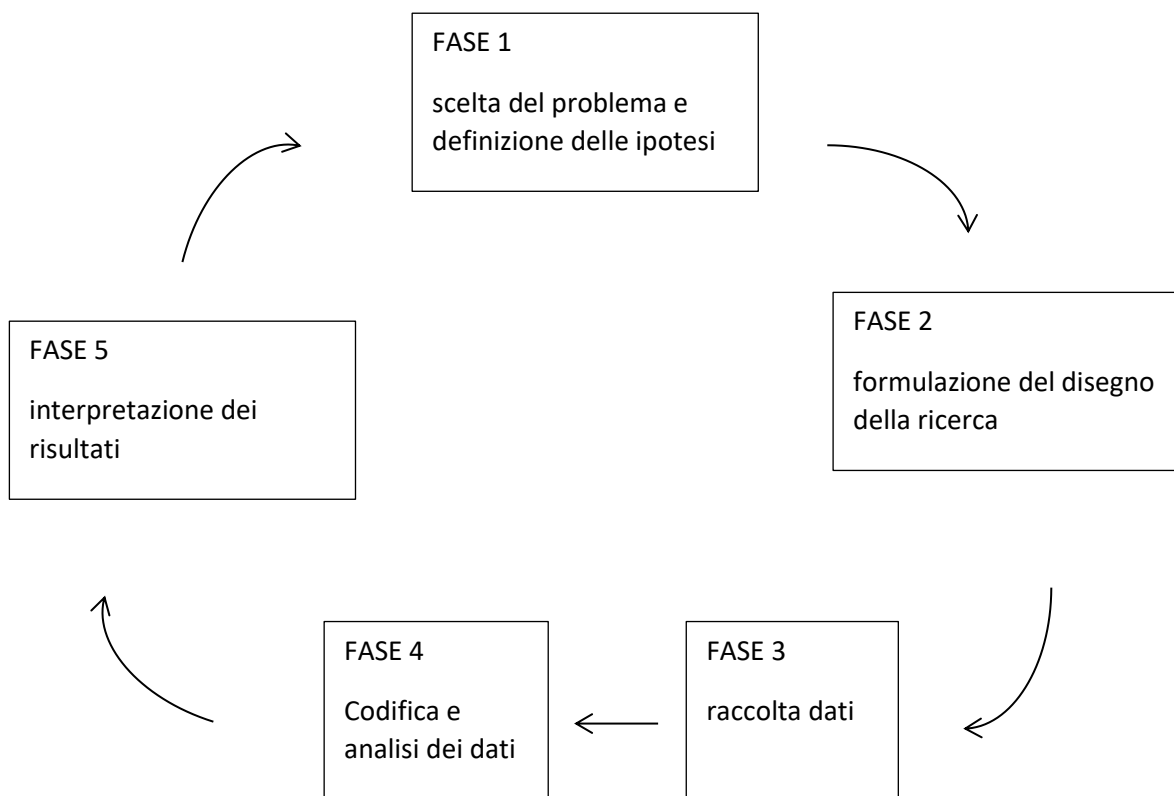


Figura 02. Fasi della ricerca sociale proposte da Bailey 2006, p.15.

Ciascuna di queste fasi dipende dall'altra e si passa alla successiva solo dopo aver soddisfatto quella precedente. Si tratta però di un processo circolare che all'ultima fase non è terminato, può ripartire quando lo studio non ha avuto successo oppure per affinare o ragionare nuovamente sull'ipotesi di ricerca. Spesso è l'analisi dei dati che porta ad una revisione delle ipotesi di ricerca e quindi della ripetizione delle fasi di ricerca. Ma anche se la ricerca ha successo è utile ripetere le fasi, magari con un campione diverso, per

dimostrare che i dati non sono il risultato di una coincidenza o che sono accidentali. La ripetizione esatta dello studio è chiamata replica.

Rispetto alle fasi definite da Bailey, solitamente trattate in maniera lineare e consecutiva, il processo può anche essere suddiviso in “momenti” (Figura 03) e questo permette di considerare la ricerca sociale in un percorso di istanze in *continuum* con lo scopo di comprendere il processo nella sua totalità, senza la necessità di dover scegliere alcune fasi a discapito di altre. «È possibile considerare il processo di analisi del contenuto come una metafora dell’approccio comprendente e quindi di un percorso di ricerca nella sua completezza» (Sbalchiero, 2021).

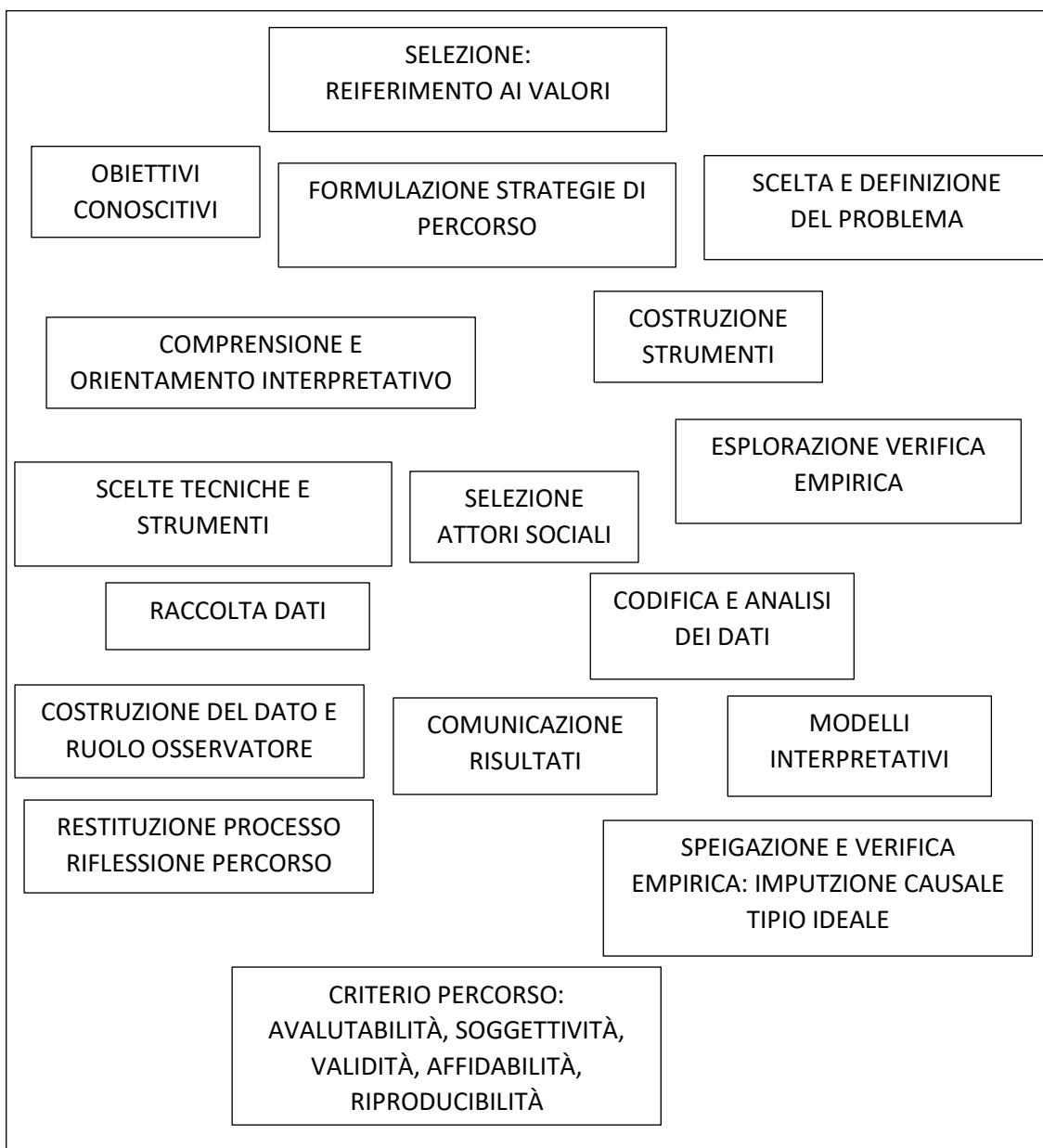


Figura 03. Mappa dei momenti tipici della ricerca comprendente (Sbalchiero, 2021).

I vari momenti, quindi, vengono trattati in maniera incrociata e non in modo consecutivo e possono essere ricondotti ai principi fondamentali della sociologia comprendente di Weber. «Le questioni rispetto all'oggettività delle conoscenze prodotte, il principio dell'avalutatività, e quindi il tema del rapporto tra osservatore e osservato, caratterizzano l'approccio comprendente nei suoi vari momenti e non in una sua specifica fase» (Sbalchiero, 2021, p.106).

La ricerca sociale è quindi un'attività scientifica che implica una dimensione empirica e il suo percorso aiuta a ridurre la complessità dei fenomeni osservati sulla base di un modello in grado di includere diversi punti di vista. Il modello in questione viene definito come una teoria.

Una teoria senza ricerca porterebbe a delle conoscenze astratte, viceversa un'attività di ricerca senza una teoria produrrebbe una massa di dati senza senso, difficili quindi da interpretare per la mancanza di specifiche domande di ricerca (De Rose, 2003).

La teoria costituisce l'inquadramento dell'indagine che serve a formulare ipotesi che dovranno poi essere controllate empiricamente.

Hempel nel suo saggio del 1952 definisce la teoria scientifica in questo modo:

“paragonabile a una complessa rete nello spazio. I suoi termini sono definiti dai nodi, mentre i fili colleganti questi corrispondono, in parte alle definizioni, e, in parte, alle ipotesi fondamentali e derivate dalla teoria. L'intero sistema fluttua, per così dire sul piano dell'osservazione, cui è ancorato mediante le regole interpretative. Queste possono venir concepite come fili non appartenenti alla rete, ma tali che ne connettono alcuni punti con determinate zone del piano di osservazione. Grazie a siffatte connessioni interpretative, la rete risulta utilizzabile come teoria scientifica: da certi dati empirici è possibile risalire, mediante un filo interpretativo, a qualche punto della rete teorica, e qui procedere attraverso definizioni e ipotesi, ad altri punti, dai quali per mezzo di un altro filo interpretativo si può infine ridiscendere al piano dell'osservazione.” (Hempel, 1952, trad. it. 1961, p.46).

2.1.2. I paradigmi

La comunità scientifica cerca di identificarsi, nei vari momenti storici, all'interno di un preciso insieme di teorie che per Kuhn costituiscono un paradigma. Il paradigma è una struttura concettuale, teorica e condivisa fondata su acquisizioni precedenti (fatti, ipotesi

e tecniche). «Rappresenta una visione del mondo che definisce, per il suo detentore, la natura del mondo, il posto dell'individuo in esso e la gamma di possibili relazioni con quel mondo e le sue parti» (Guba, Lincoln, 1994, p. 107). Il paradigma favorisce il consolidamento di un linguaggio comune tra i teorici e tra i praticanti di una medesima disciplina e permette di precisare i principi, le regole e le convenzioni alla base delle loro attività. Stabilisce anche le regole che i ricercatori devono rispettare - pena il loro rigetto da parte del gruppo di appartenenza - rafforza la resistenza al cambiamento (Boivin, 1997) e favorisce poi la convergenza dei ricercatori verso le medesime domande e l'impiego delle stesse metodologie. Sotto un determinato paradigma la scienza è definita normale e il passaggio da un paradigma all'altro determina il cambiamento scientifico.

Un paradigma è fondato su tre quesiti: ontologico, epistemologico e metodologico (Guba, Lincoln, 1994). La questione ontologica si riferisce al "che cosa", l'essenza, la natura della realtà sociale e la sua forma; la questione epistemologica è il rapporto tra il "chi" (studioso) e il "che cosa" (realtà studiata), la conoscibilità della realtà sociale; l'aspetto metodologico, si interroga sul "come", in che modo si può riconoscere la realtà sociale.

La ricerca sociale è caratterizzata da due i paradigmi principali: il positivismo e l'interpretativismo. Le differenze le troviamo nelle risposte che danno alle domande ontologiche (la realtà sociale esiste?), epistemologiche (la realtà sociale si può conoscere?) e metodologiche (come può essere conosciuta la realtà sociale) della ricerca sociale.

Il primo paradigma ad essere usato nelle scienze sociali è stato il paradigma positivista che studia la realtà sociale utilizzando tecniche di osservazione, di misurazione e strumenti di analisi matematica. Il primo vero sociologo positivista è stato Durkheim, che con la sua teoria trattava i fatti sociali come cose effettivamente esistenti, studiabili oggettivamente e senza aspetti soggettivi. Dal punto di vista ontologico la realtà sociale ha esistenza effettiva ed è conoscibile, come se si trattasse di una "cosa". Dal punto vista epistemologico, ricercatore e oggetto di studio non si influenzano in nessun modo e l'obiettivo è quello di spiegare e di formulare leggi naturali e generali immutabili con risultati veri e certi. La metodologia positivista prevede una modalità induttiva nel procedere, dal particolare al generale con esperimenti e tecniche quantitative (raccolta di dati numerici elaborati statisticamente) con utilizzo di variabili.

Dal positivismo si sviluppa il neopositivismo come risposta alle critiche dei positivisti. La visione ontologica, critica il realismo, e presuppone la conoscenza superficiale e imperfetta di una realtà umana esterna. L'epistemologia neopositivista valuta positivamente il rapporto di sovrapposizione tra studioso e studiato anche se dovrebbe essere limitato il più possibile. La metodologia neopositivista rimane simile a quella del positivismo, con un avvicinamento però ai metodi qualitativi.

Sul fronte opposto al positivismo troviamo l'interpretativismo, che si propone di comprendere la realtà piuttosto che spiegarla. Weber è il suo massimo esponente. La sua ontologia si rispecchia con le realtà multiple del costruttivismo e del relativismo, dove non esiste una realtà oggettiva, ma ogni individuo si crea una sua dimensione che è l'unica ad essere conoscibile. Le singole realtà si diversificano nei diversi contesti culturali e quindi non si può parlare di una realtà sociale unica e ogniqualvolta valida. Per quel che riguarda l'epistemologia, nell'interpretativismo, la ricerca sociale si avvicina più ad una scienza interpretativa con l'obiettivo di trovare un significato piuttosto che una scienza che mira alla scoperta di leggi scientifiche. Molto spesso la ricerca sociale si avvale di astrazioni e concetti di generalizzabili che possono essere ricondotti agli idealtipi di Weber. «Lo stile epistemologico di Durkheim viene definito naturalistico e quello di Weber ermeneutico» (Sbalchiero, 2021 p.24). La metodologia prevede l'interazione tra studioso e studiato, il metodo utilizzato è quello induttivo dal particolare al generale. Le tecniche utilizzate sono principalmente qualitative, come le interviste strutturate e semi strutturate, il focus group, l'osservazione partecipante etnografica e l'osservazione a distanza.

2.2 Metodologie, metodi, tecniche della ricerca sociale

2.2.1 La metodologia

La metodologia è un sottoinsieme dell'epistemologia (la scienza della conoscenza scientifica) e può essere definita la scienza della scoperta, dal greco *metà odos* (strada, via, con la quale) e *logos* (discorso, parola). Infatti, la metodologia è il discorso, lo studio, la riflessione sui metodi utilizzati da una disciplina per acquisire conoscenza. Riferita alla

ricerca sociale, risponde alla domanda: come può essere conoscibile, scientificamente, la realtà sociale? Percorrendo quali strade? Con quali metodi?

La metodologia è la disciplina che studia gli assunti, le procedure, i metodi di spiegazione utilizzati in un determinato ambito conoscitivo. Evidenzia quali sono i pericoli ed i punti di forza riguardo i metodi adottati, generalizza la conoscenza metodologica e rende disponibile alla comunità scientifica il lavoro empirico del ricercatore. (Lazarsfeld, 1967).

La metodologia è una riflessione sui metodi e sulle tecniche da adottare per conoscere ed interpretare la realtà, stabilisce la logica ed i principi fondamentali con i relativi punti di forza e punti deboli per interpretare un fenomeno.

Il rigore metodologico rende affidabile una ricerca sociale quando segue un percorso preciso e ben strutturato (De Rose, 2018):

1. scelta dell'approccio metodologico che meglio risponde alla domanda di ricerca;
2. definizione chiara e riproducibile dei metodi e delle tecniche;
3. conduzione seria e rigorosa della ricerca sul campo;
4. fermarsi solo quando si è soddisfatti della conoscenza ottenuta (il criterio della saturazione);
5. trasparenza e pubblicazione dei risultati per poter riprodurre il percorso di ricerca.

2.2.2 I metodi

I metodi (al plurale) sono le procedure e le regole utilizzate per rispondere alle domande di ricerca. Guidano il processo di ricerca e consentono di scegliere tra le tecniche esistenti, quelle più adatte per raggiungere gli obiettivi conoscitivi oppure di crearne e codificarne delle nuove. Si tratta del percorso che il ricercatore dovrà scegliere e seguire a livello empirico, è la "strada da compiere" tenendo in considerazione quattro aspetti:

- 1) la modalità cognitiva prevalente utilizzata per fare la ricerca sul campo (ascoltare, domandare, osservare, leggere, riflettere o modificare);
- 2) le conoscenze scientifiche e le convinzioni etiche del ricercatore;
- 3) le scelte pratiche utili ad affrontare i problemi che pone ogni ricerca;
- 4) le fasi, le attività introdotte e le risorse.

Si può quindi concludere utilizzando una metafora: «fare ricerca sociale è come avere tra le mani un indirizzo, rappresentato dall'obiettivo finale; il metodo rappresenta il percorso di massima per raggiungere l'abitazione, dalla posizione in cui ci trova». (Sbalchiero 2021, p. 105).

2.2.3 Le tecniche

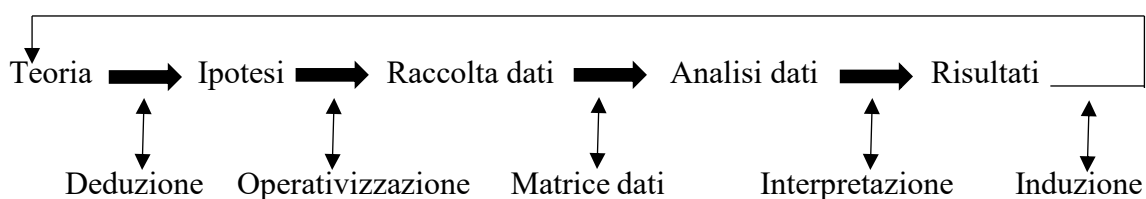
Se il metodo è un'operazione che implica scelte discrezionali, la tecnica è squisitamente esecutiva. In passato si è (erroneamente) creduto che la ricerca sociale fosse basata essenzialmente su tecniche in grado di produrre risultati depurati dalla discrezionalità del ricercatore. Le tecniche invece sono «delle specifiche procedure operative, riconosciute dalla comunità scientifica e trasmissibili per insegnamento» (Corbetta 2003, p. 9), che vengono utilizzate per acquisire e controllare i risultati empirici di una ricerca scientifica. Paradossalmente in una ricerca sociale non si può fare a meno delle tecniche. Diverso è per l'aspetto metodologico che può essere definito come un'autoriflessione sui mezzi che hanno trovato conferma nella prassi. Ma tra metodologia e tecniche non esistono vere e proprie separazioni o confini definiti, esiste invece una sorta di continuum con riflessioni metodologiche che portano a concetti più operativi e viceversa. «Alcuni esempi di tecniche sono, per esempio, [...] le tecniche di intervista e questionario, le procedure di analisi statistica dei dati» (Corbetta, 2019c, pp. 7-8). Si distinguono due tecniche principali in due momenti precisi della ricerca sociale: ci sono «tecniche per la raccolta dei dati» (Marradi, 1988), che possono variare a seconda della loro formalizzazione (si pensi all'osservazione partecipante, all'intervista o all'indagine campionaria), e le «tecniche di analisi dei dati, quelle procedure specifiche per pervenire a una sintesi e a una restituzione delle evidenze empiriche prodotte» (Marradi, 1990).

2.3 Gli approcci alla ricerca sociale

2.3.1 L'approccio quantitativo o standard

Le ricerche basate su analisi statistiche (test per inferenza, correlazioni e tecniche multivariate) e la formulazione di asserti su relazioni tra proprietà che non dipendono da

conoscenze e valutazioni personali, sono caratteri dell'approccio quantitativo. Per quel che riguarda la ricerca sociale l'approccio quantitativo è più formalizzato rispetto al qualitativo, sia per rilevazione e analisi dei dati, sia dal punto di vista dei passaggi da seguire durante la ricerca. Il percorso quantitativo è il percorso tipico o standard ed è costituito da cinque fasi (teoria, ipotesi, raccolta dati, analisi dati e risultati) con i relativi cinque processi (deduzione, operativizzazione, organizzazione o matrice dati, interpretazione e induzione) (Corbetta, 2003).



La prima fase è quella della teoria, che attraverso un processo di deduzione, arriva alla seconda fase, quella dell'ipotesi. Per elaborare un'ipotesi si parte da una teoria cioè da una serie di preposizioni organicamente collegate. La teoria è generale e l'ipotesi è specifica. L'ipotesi implica una relazione tra due o più concetti e si colloca ad un livello di astrazione più basso rispetto alla teoria. Un tentativo di spiegare un fenomeno sociale che deve essere verificato.

La fase successiva è quella di operativizzazione, un processo molto delicato che servirà per la raccolta dei dati e quindi della scelta delle variabili che sono molto importanti perché influiscono sull'esito dello studio. In questa fase il ricercatore trasforma concetti astratti in entità empiricamente rilevabili. I concetti sono dei contenuti semantici, dei segni linguistici, delle rappresentazioni mentali che legano teoria e mondo empirico osservabile. «I concetti sono i mattoni della teoria e la teoria ha un senso per la ricerca empirica solo nella misura in cui riesce a connettersi proficuamente con il mondo empirico» (Corbetta 2003, p.28); «i concetti sono gli strumenti, e gli unici strumenti, per stabilire una tale connessione» (Blumer 1969, p.143). Operativizzando i concetti si realizza la traduzione di una teoria che viene verificata passando dai concetti astratti alla loro applicazione come proprietà specifica, detta unità di analisi. Per fare un esempio, l'altezza è un concetto, l'altezza di una persona è una proprietà e l'altezza di una persona misurata con un metro è una variabile. L'unità di analisi, nell'approccio quantitativo,

rappresenta l'oggetto al quale si riferiscono le proprietà studiate. Nel mio esempio, quindi, sono le singole persone le unità di analisi, l'individuo, ma negli studi su larga scala, che confrontano grandi aree geografiche o ampi aggregati di persone, le unità di analisi possono essere o un gruppo di individui o un aggregato di persone legate alla territorialità oppure anche un evento sociale.

Corbetta stabilisce quattro fasi nel processo di operativizzazione:

1. individuazione delle proprietà concrete dei caratteri delle unità di analisi che rappresentano il concetto astratto;
2. stabilire i diversi stati delle proprietà individuate;
3. dare una definizione operativa;
4. applicare delle scelte (discrezionali del ricercatore che individua proprietà e variabili).

Per far un esempio pratico, è possibile rilevare il concetto astratto del "livello culturale" attraverso le proprietà di ogni singolo individuo come il titolo di studio o il numero di articoli scientifici scritti o il numero di libri letto. Se prendiamo come esempio la proprietà titolo di studio la possiamo classificare attraverso una modalità e un valore. La modalità si riferisce al titolo di studio in possesso (nessuno, elementare, medie inferiori, medie superiori, laurea, dottorato) e il valore è il numero assegnato alla modalità (zero per nessun titolo di studio fino a cinque per il dottorato). Modalità e valore rappresentano la variabile. La quarta fase è quella relativa all'analisi dei dati, è in questa fase, nella ricerca quantitativa, che viene creata la matrice dei dati utilizzata poi per le varie analisi statistiche. Le principali analisi statistiche da effettuare iniziano con la costruzione della tabella di frequenza e con il conteggio delle unità statistiche, il calcolo della moda, delle frequenze cumulate e della mediana. Si passa poi al calcolo dei quartili, della deviazione standard e poi se vogliamo osservare la relazione tra due variabili, il calcolo della regressione lineare e quindi se esiste una legge empirica che lega una variabile indipendente ad una dipendente. Quando due variabili sono in relazione, la variabile indipendente è vista come la variabile che influenza cioè la causa; invece, quella dipendente è l'effetto e quindi la variabile influenzata. L'ultima fase è rappresentata dall'interpretazione delle analisi condotte sulla matrice dei dati costruita precedentemente e quindi nella descrizione dei risultati. Il processo poi è ricorsivo e riparte dalla teoria iniziale mediante un processo induttivo dove vengono confrontati i risultati ottenuti nella

precedente teoria. Fasi e processi ben precisi, standardizzati e generalizzabili che guidano il ricercatore durante il suo percorso di studio.

La principale tecnica utilizzata nella ricerca quantitativa è la raccolta di dati attraverso strumenti standardizzati e strutturati quali il questionario. Il questionario costituito da una serie di domande, detti stimoli, che possono avere risposta chiusa (scelta multipla, set di risposte, scala *likert*) o a risposta aperta (testo breve), anche se questa ultima modalità si allinea meglio all'approccio qualitativo. Le domande chiuse poi (per esempio sì e no) sono le più semplici ma non danno molto in termini di ricerca (è utile quindi a volte inserire anche la scelta "non so" oppure "altro" con la possibilità di inserire del testo). Per la formulazione delle domande è necessario tenere presente le quattro fasi del processo cognitivo e nello specifico che:

- 1) porre una domanda in maniera oscura o poco accessibile significa mettere in crisi il processo cognitivo fin dall'esordio;
- 2) porre una domanda su avvenimenti lontani nel tempo comporta un faticoso recupero di informazioni;
- 3) di fronte alla richiesta di un'opinione l'intervistato è costretto a fare i conti con il senso comune e con il giudizio dell'eventuale intervistatore;
- 4) proporre un elenco di alternative strutturate in cui non è presente l'opinione dell'intervistato equivale a metterlo in crisi nell'ultima fase del processo.

Le tecniche di rilevazione del questionario possono avvenire attraverso la tradizionale spedizione postale, face to face tra intervistatore e intervistato o nella modalità *cawi* (*Computer Assisted Web Interviewing*). Nel primo caso si ha un tempo elevato per la trasmissione e la ricezione dei dati, un alto numero di risposte mancanti e, come vantaggio, nessun tipo di influenza da parte dell'intervistatore. Con la modalità faccia a faccia, ci può influenzare le risposte date a causa della presenza dell'intervistatore ma un'alta qualità del dato. Con la tecnica di rilevazione *cawi*, infine, si ha un basso costo di spesa e di durata temporale di raccolta del dato e una qualità media dello stesso.

2.3.2 L'approccio qualitativo o non standard

A differenza dell'approccio quantitativo dove lo scopo è quello di rilevare dati per fini statistici per pervenire a spiegazioni e quindi generalizzare (spiegare per comprendere), nell'approccio qualitativo l'obiettivo principale è la comprensione dei fenomeni nei vari contesti seguendo da vicino la vita quotidiana degli attori sociali. Il ricercatore ha un ruolo fondamentale, che non è quello passivo e distaccato dell'approccio quantitativo, ma è pienamente coinvolto e attivo nel documentare e valorizzare i contesti con un orientamento interpretativo e induttivo, dal particolare quindi fino ad un livello più ampio e generale. Implica anche grande capacità del ricercatore in termini di conoscenze, flessibilità, intuizione e tempistiche. L'approccio qualitativo è anche definito non standard perché non ha procedure fisse da seguire, la soggettività del ricercatore è centrale e non vi è generalizzabilità dei casi studiati. La ricerca qualitativa nasce a partire dagli anni Sessanta, ed i suoi metodi hanno avuto il loro apprendistato nella sociologia urbana, delle professioni e della devianza. Un forte impulso fu dato da Robert E. Park e William I. Thomas (i fondatori della Scuola di Chicago⁶) per trovare risposte ai problemi dell'immigrazione, dell'integrazione e della devianza sociale. Anche la società contemporanea si caratterizza per un'alta concentrazione di etnie, lingue, e rituali. Tutte caratteristiche che rendono particolarmente difficile l'uso di tecniche quantitative quali il questionario. Nasce quindi l'esigenza di trovare metodi e tecniche meno standardizzati, flessibili e capaci di adattarsi meglio all'unità di analisi studiata senza imporre un modello informatico fisso e poco resiliente ai diversi attori sociali. Due sono i motivi della crescente diffusione dei metodi qualitativi: «la loro flessibilità e capacità di adattarsi alle situazioni» (Silverman, 2002) e «l'utilizzo di programmi informatici che hanno migliorato la precisione e l'attendibilità delle classificazioni del ricercatore e la rigosità dei suoi ragionamenti oltre a garantire l'ispezionabilità dei dati, considerata una delle carenze storiche della ricerca qualitativa» (Ricolfi, 1995). Le principali tecniche utilizzate nell'approccio non standard sono: l'osservazione partecipante, il focus group e l'intervista discorsiva. L'osservazione partecipante vede il coinvolgimento diretto sul campo del ricercatore e la raccolta di dati testuali, foto, video e audio, i campi di

⁶ La prima scuola di sociologia urbana negli Stati Uniti d'America, nata il 1914.

applicazione spaziano dagli studi di comunità, sub culture, a fenomeni dei quali si conosce poco. Il focus group è un metodo qualitativo di rilevazione basato sulla discussione tra un gruppo di persone con dei moderatori e un argomento da approfondire. Si utilizza per capire e studiare problemi sociali complessi e i campi di ricerca spaziano da quelli sanitari (per studiare malattie o stili di vita) al marketing (per il lancio di un nuovo prodotto) all'ambito politico (per studiare le previsioni di voto). L'intervista discorsiva, infine, ha come finalità quelle conoscitive e per comprendere "punti di vista", è guidata da un moderatore secondo uno schema non standard ma flessibile. Viene utilizzata in campo bibliografico (per la raccolta di dati inerenti a storie di vita e racconti), in ambito sanitario (per capire il percorso che ha portato ad un esito) e in campo ermeneutico (per la comprensione dei significati acquisiti durante il processo di socializzazione). Sia il focus group che l'intervista discorsiva raccolgono, oltre che opinioni e comportamenti, anche gli stati d'animo delle persone intervistate rispetto alle loro opinioni. I principali limiti e le criticità della ricerca qualitativa (soprattutto da parte di chi segue più un approccio quantitativo) sono la soggettività del ricercatore, la difficoltà nella generalizzazione e delle procedure non standard oltre al forte impiego di risorse sia economiche che temporali.

2.3.3 La gestione dei dati nei due approcci

Dalle differenze dei due tipi di approcci descritti nei paragrafi precedenti ne consegue anche una differenziazione per quel che riguarda la gestione dei dati: dalla raccolta, all'analisi e infine nella divulgazione dei risultati. La ricerca quantitativa utilizza le tabelle per strutturare i propri dati prima dell'analisi, tabelle che sono organizzate su righe e colonne e permettono una visualizzazione dei valori in maniera chiara, precisa, sistematica e sintetica. L'organizzazione e la struttura tabellare però è decisa dal ricercatore e, molte volte, non rispecchia lo schema mentale dei soggetti presi in considerazione, causando un impoverimento del dato stesso per quel che riguarda l'informazione. Nella ricerca qualitativa invece, i dati raccolti sono descrizioni, testi o porzioni di testo e riportano parola per parola la voce degli intervistati, facendo da una

parte una sorta di fotografia dei loro pensieri e dall'altra rimediano al problema della mancanza di ricchezza del dato strutturato in forma numerica nelle tabelle.

L'obiettivo dell'approccio quantitativo poi è quello di pervenire a spiegazioni e generalizzare, per giustificare i risultati ottenuti. Cercare quindi il nesso causale tra le variabili anche attraverso leggi universali. L'approccio qualitativo invece si focalizza su categorie concettuali, che nella realtà non esistono ed hanno un livello di astrazione superiore. Dei "tipi ideali" utilizzabili come modelli per interpretare la realtà e descrivere i fenomeni sociali e che costituiscono l'accentuazione unilaterale di alcuni elementi della realtà al fine di caratterizzare gli aspetti salienti del fenomeno indagato (Sbalchiero, 2021). Tipologie che mettono in evidenza le complesse interconnessioni tra fenomeni ricomposte in cause plausibili e adeguate (Pace, 2021). Secondo Weber gli idealtipi permettono di mettere in luce gli elementi significativi della realtà, utili poi per comprendere il significato dei fenomeni sociali.

Per concludere, se la quantità dei dati raccolti nella ricerca quantitativa è direttamente proporzionale alla generalizzabilità dei risultati, nella ricerca qualitativa un maggior numero di soggetti da valutare porta ad un minore approfondimento dei singoli casi e quindi profondità e ampiezza della ricerca sono inversamente correlate. «La ricerca qualitativa, quindi, segue la strada della riduzione dell'estensione del dominio osservato, la focalizzazione di pochi casi, di cui si propone di rilevare i più minuti dettagli» (Cardano, 2011, pp. 16-17).

2.3.4 Considerazioni

Alla luce delle differenze tra ricerca qualitativa e quantitativa la logica dell'indagine scientifica è unitaria. Non possiamo definire dei veri e propri metodi qualitativi e quantitativi, nel senso che il metodo, visto come insieme di passi procedurali e regole scientifiche, è lo stesso. Le differenze sono a livello delle tecniche e delle strategie di ricerca che riassumo nella tabella che segue. Allo stesso tempo la ricerca sociale può essere definita come il punto di incontro tra "*thin description*" (lo studio di tutto ciò che è collocabile in una matrice di dati, statistica sociale e questionari) e "*thick description*" (analisi di tutto ciò che è costitutivo di senso e può essere analizzato essenzialmente con

metodi qualitativi come l'intervista in profondità, focus group, osservazione partecipata). (Sbalchiero, 2021).

Peraltro, non è nemmeno possibile definire quale dei due approcci sia "scientificamente" migliore dell'altro. Corbetta nel suo volume "Metodologie e Tecniche della Ricerca Sociale del 2014" (Corbetta, 2014) individua tre posizioni. La prima conferma l'incompatibilità dei due approcci, e quindi i sostenitori del proprio paradigma confermano la loro correttezza rispetto all'altro approccio. Un secondo punto di vista, neopositivista, vede nella metodologia qualitativa solo un approccio preliminare e stimolante per il ragionamento generale della ricerca. La terza e ultima posizione assegna ugual valore alle due metodologie, qualitative e quantitative, con l'auspicio che la scelta del primo o del secondo approccio avvenga in relazione al contesto e al fenomeno da indagare. Infatti, tutte e due le metodologie, hanno dato il loro contributo alle scienze sociali e quindi è auspicabile che possano essere prese in considerazione congiuntamente per migliorare la comprensione della realtà con i due diversi punti di vista.

PRADIGMA	
POSITIVISTA	INTERPRETATIVISTA
OGGETTO DELLO STUDIO	
Fatti sociali, reali, empirici basati su "dati"	Realtà relativa, costruita
COMPITO DELL'INDAGINE	
Spiegare, dar conto, alla luce delle cause e condizioni che li determinano	Ricerca del significato, comprendere il fenomeno/oggetto di studio
RACCOLTA DATI	
Contesto artificiale basato sul controllo di variabili	Contesto naturale e campione ad hoc
FORMATO DATI	
Numeri in matrice di dati	Audio, video, testi
ANALISI DATI	
Elaborazioni statistiche (conteggi, frequenze e correlazioni)	Descrizione dei fatti, comparazioni
PROCEDURE	

Osservative con tecniche standardizzate e fisse	Legate alla discrezionalità del ricercatore
ATTEGGIAMENTO DEL RICERCATORE	
Distaccato	Pieno coinvolgimento
APPROCCIO	
Pervenire a spiegazioni e generalizzare, spiegare per comprendere. Leggi assolute ed universali	Comprendere per spiegare, non leggi generalizzate ma eventualmente tipizzazioni
STRATEGIA	
QUANTITATIVA	QUALITATIVA
PRINCIPALI METODI	
Raccolta dati basata su strumenti standardizzati quale il questionario	Osservazione partecipante, focus group ed intervista discorsiva

Tabella 02. Confronto del paradigma positivista vs paradigma interpretativista.

2.3.5 Metodi misti

Partendo dalle considerazioni finali del paragrafo precedente si può quindi parlare di metodi misti (o *mixed method*). Si parla di metodi misti quando si combinano tecniche standardizzate e non standardizzate di rilevazione e il disegno di ricerca include quindi sia tecniche qualitative che quantitative. «L'uso congiunto di tecniche qualitative e quantitative, chiamato anche triangolazione, consente di capire la vera posizione di un oggetto lontano guardandolo appunto da direzioni e lati differenti» (Lazarsfeld, 1971, p. 49). Nel disegno della ricerca dei *mixed method* sono incluse, in un processo continuum, sia tecniche qualitative che quantitative. Il suo utilizzo serve sia a spiegare sia ad interpretare un fenomeno. Può essere utile anche per sviluppare e testare nuovi strumenti di ricerca e per capire punti di forza e debolezza dei vari approcci. Con i metodi misti si possono generalizzare dati qualitativi ma ho di fronte vari punti di debolezza come l'elevato tempo richiesto per lo studio di ricerca, la discrepanza effettiva tra i dati dei due approcci e molte volte risultati incomparabili. Per effettuare una ricerca basandosi su metodi misti è importante separare la raccolta dei dati qualitativi dai quantitativi, è utile

inserire domande qualitative in un disegno di ricerca quantitativo (per esempio domande aperte in un questionario standardizzato e strutturato) e condurre le ricerche in maniera indipendente per poi compararle e convergerle. Due sono i metodi misti utilizzati nella ricerca sociale: *Sequential Explanatory Design* e *Sequential Exploratory Design*. Nel primo caso si parte con una fase quantitativa per poi passare ad un approccio qualitativo. I dati qualitativi mi aiuteranno a spiegare i risultati del metodo quantitativo. Nel *Sequential Exploratory Design* si parte con una fase iniziale qualitativa per poi passare a tecniche quantitative. Comprensione e spiegazione di un fenomeno tramite verifica empirica quindi. (Figura 04).

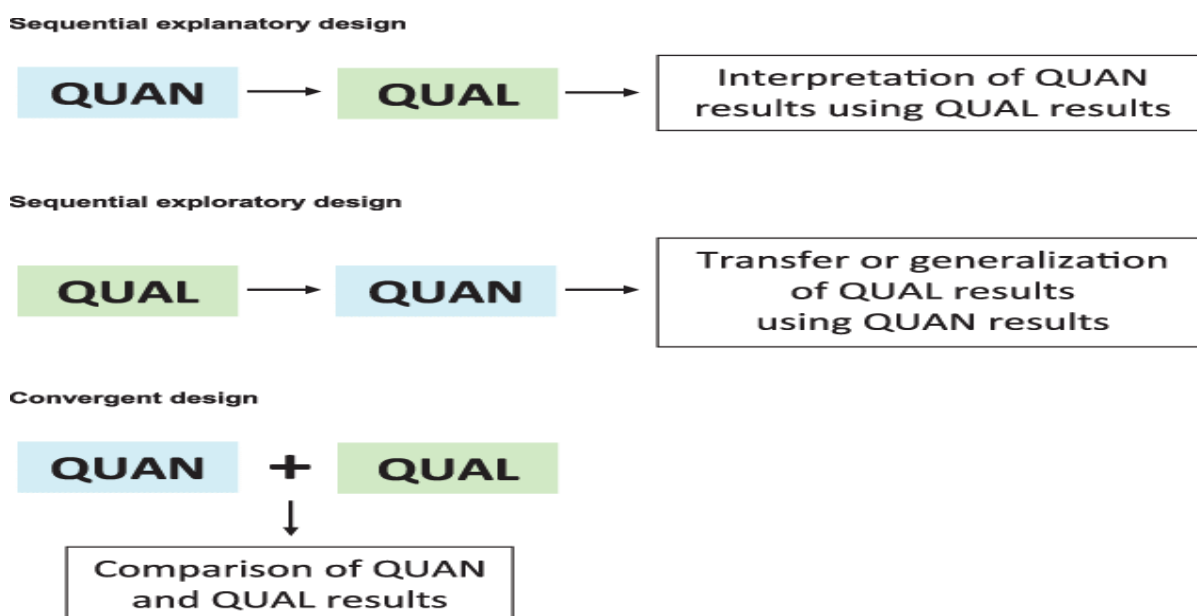


Figura 04. Da “Applied Mixed Methods in Oral Health Research”: Importance and Example of a Training Program B. Nicolau, G. Castonguay, A. Levine, Q.N. Hong Summer Institute 2015 Participants* and P. Pluye.

CAPITOLO 3

L'analisi del contenuto. Storia, definizioni e applicazioni

3.1 L'analisi del contenuto

3.1.1. La storia

«L'analisi del contenuto è un viaggio che conduce dalla parola al numero e ritorno ed è a cavallo tra analisi qualitativa e analisi quantitativa» (Tuzzi, 2003, p.17).

Prima di parlare di definizione, o meglio di definizioni, di analisi del contenuto, è utile cercare di capire l'evoluzione che ha avuto negli anni questa disciplina.

Secondo quanto scritto nel saggio del 1980 di Krippendorff "*Content Analysis: An Introduction to its Methodology*" (tradotto in italiano, giapponese, spagnolo e ungherese), in letteratura il primo esempio di analisi del contenuto (anche se non viene sempre menzionato) risale al XVII secolo, in Svezia, e riguarda l'analisi di testi religiosi. Lo studio, commissionato dalla Chiesa Luterana, aveva come obiettivo quello di capire se l'opera i "Canti di Sion" (costituita da novanta inni religiosi di un autore sconosciuto) conteneva locuzioni in contrasto con l'ortodossia. Un lavoro molto più noto, recente e rilevante di analisi del contenuto, risale però al 1918 con l'opera di William I. Thomas e Florian Znaniecki "*The Polish Peasant in Europe and America*". In questo studio i due ricercatori analizzarono i testi di circa ottocento lettere che gli emigranti polacchi in USA spedivano alle loro famiglie. Fu una ricerca qualitativa con lo scopo di comprendere le condizioni di vita degli emigranti e i mutamenti della vita contadina polacca in un nuovo paese industrializzato. Se da una parte questo lavoro è visto ancora oggi come precursore della sociologia moderna, dall'altra ha riscosso anche numerose critiche sul piano metodologico. Tra le più significative troviamo quelle di Blumer ("*Critiques of research in the social sciences: an appraisal of Thomas and Znaniecki's The Polish peasant in Europe and America*", 1939), che mette in evidenza alcuni aspetti riguardo la mancata descrizione delle procedure e delle tecniche di analisi adottate, la non sistematicità della scelta delle fonti documentali e la soggettività nelle interpretazioni dei testi.

Gli anni seguenti alla pubblicazione dell'opera di Thomas e Znaniecky, segnano un proliferare di ricerche sull'analisi del contenuto, principalmente per il diffondersi della

carta stampata ma anche per le vicende legate alla Prima Guerra Mondiale. L'obiettivo principale, quindi, è quello di indagare nel campo della comunicazione politica e della propaganda. Le analisi del contenuto utilizzate erano basate su strumenti quantitativi molto semplici, come il calcolo della frequenza di determinati argomenti, spazio dedicato alle notizie e lunghezza dei titoli (Amaturo, 1993). Di particolare importanza sono gli studi condotti da Lasswell, ricercatore di politica nella prestigiosa Scuola di Chicago, a partire da quello tratto dalla sua tesi di dottorato del 1927 "Propaganda Technique in the World War", dove metteva in evidenza la carenza di rigore scientifico e metodologico nelle ricerche sulla propaganda politica di quel periodo perché non esplicitavano i criteri di campionamento, la selezione dei materiali e la costruzione degli indici di campionamento. Secondo Lasswell solo un'analisi quantitativa rigorosa poteva dare certezze sull'attendibilità dei risultati (Tuzzi, 2003). Un altro suo studio, sull'analisi di centocinquantesette articoli di undici numeri del periodico "The Galileian" del 1942, portò alla condanna del suo direttore responsabile William D. Peley per propaganda antiamericana. La sua analisi del contenuto, infatti, aveva messo in luce che negli articoli del periodico "The Galileian" vi erano delle analogie con quattordici temi di propaganda nazista. Lo stesso Lasswell insieme a Yacobson pubblicarono poi nel 1949 "The Language of Politics" una ricerca sulla semantica quantitativa con l'analisi degli slogan conati in occasione della festa del Primo Maggio in Unione Sovietica nel periodo tra il 1918 e il 1943. L'idea di fondo era quella di capire i mutamenti della linea politica del Partito Comunista sull'asse temporale.

È Lasswell il padre fondatore del termine *content analysis*, grazie agli studi relativi alle tecniche di persuasione utilizzate durante la Prima Guerra Mondiale ma anche per le ricerche sulla propaganda, definita «il controllo degli atteggiamenti collettivi, attraverso la manipolazione di simboli significativi» (Gili 2010, p. 28). La *content analysis* era nata, in principio, solo per indicare le metodologie di analisi quantitativa della comunicazione politica e, solo in seguito, è stata estesa a qualunque ricerca che riguardava i contenuti di un messaggio (messaggio non solo politico o necessariamente scritto) (Tuzzi, 2003).

La veloce diffusione di questa nuova disciplina necessitava però anche trasparenza e scelte ben definite nel campo metodologico. Furono Lasswell, Berelson insieme a Lazarsfeld e Krippendorff i precursori del tentativo di stabilire delle regole ben precise riguardo i metodi e le tecniche di ricerca da utilizzare nell'analisi del contenuto. I

problemi, infatti, erano legati alla metodologia utilizzata senza standardizzazioni e l'utilizzo di tecniche che spesso non erano rigorose, affidabili e replicabili.

L'utilizzo di questa nuova metodologia ebbe una battuta d'arresto a metà degli anni '50. Sorokin, nel 1956, coniò il termine "quantofrenia": l'eccessivo utilizzo di tecniche guidate a risultati quantitativi e molto slegate dalla dimensione semantica dei testi analizzati. L'analisi era troppo rigida, oggettiva e utilizzava sistemi matematici che facevano perdere sia i valori reali delle situazioni, sia gli aspetti qualitativi di legame tra ricercatore e osservato, oltre ai i valori semantici appunto. L'abbandono del metodo della *content analysis* da parte dei politologi venne poi ripreso, dagli psicologi e dai linguisti, negli anni '90, anche grazie alla diffusione dei primi calcolatori che aiutavano ad automatizzare le procedure. Proprio grazie ai linguisti e agli aspetti teorici e metodologici introdotti dalla scuola francese *Analyse des Données* (relativi alla statistica multivariata), l'approccio lessico-testuale passò dall'analisi delle frequenze alla logica multidimensionale delle analisi e quindi alle relazioni tra le variabili (Benzécri, 1982). Contemporaneamente, si svilupparono tecniche statistiche sempre più complesse e *software* dedicati di analisi testuali, che aiutarono all'elaborazione di aspetti sia quantitativi che qualitativi. Benzécri invece sviluppò delle tecniche di analisi delle corrispondenze tra simboli fonetici, già utilizzate in passato, ma che, grazie ai calcolatori potevano essere applicate a raccolte di testi di grandi dimensioni.

La *content analysis* è una materia in continua evoluzione. Se da una parte può risultare non ben definita, dall'altra raggruppa tutti quegli aspetti legati all'analisi del testo e nello specifico relativi a:

- *text analysis*
- *text mining*
- *text data science*
- *(quantitative/statistical) analysis of textual data*
- *content analysis, discourse analysis*
- *opinion mining, web reputation, sentiment analysis*
- *semantic analysis, semantic networks, semantic neural network*
- *digital methods (digital humanities, digital sociology, ...), distant reading*
- *computational social sciences, computational humanities, social data science*
- *natural language processing, information retrieval*

- *content mapping, content extraction, word mapping*
- *concept extraction, topic detection, document summarization*
- *text categorization, text classification, text clustering*
- *quantitative linguistics, computational linguistics, corpus linguistics*

I campi di applicazione spaziano dall'economia all'antropologia, dalla semiotica alla filosofia, dalla statistica all'informatica, dalle scienze politiche alla comunicazione, dalla linguistica alla storia e in molti altri ambiti dove i testi e quindi il linguaggio sono il *core* dell'indagine.

Si può quindi riprendere la frase di Tuzzi ad inizio capitolo e quindi rimarcare il valore della *content analysis* come «un viaggio che conduce dalla parola al numero e ritorno ed è a cavallo tra analisi qualitativa e analisi quantitativa». Il viaggio della parola al numero rimarca la storia iniziale dell'analisi del contenuto con gli aspetti quantitativi, che analizzati da soli fanno però perdere il reale significato dell'analisi se non associati ad aspetti qualitativi e quindi riferiti alla semantica delle parole. Si può in questo caso parlare di *mixed method* della *content analysis* come migliore metodologia per un'accurata e precisa analisi del contenuto, una simbiosi tra competenze linguistiche e competenze statistico-matematiche.

Dopo questo *excursus* storico vedremo, nel prossimo paragrafo, le varie definizioni date all'analisi del contenuto.

3.1.2 Le definizioni

In letteratura, una delle definizioni utilizzate per la maggiore di analisi del contenuto è: «[...] quel processo di acquisizione, sintesi e restituzione delle informazioni presenti in una comunicazione» (Tuzzi, 2003, p. 15) (Figura 05).



Figura 05. Il processo di analisi del contenuto.

La parte di acquisizione si riferisce alla raccolta del materiale e alla scansione di tutte le informazioni, con la sintesi si organizzano e si elaborano i dati testuali e solo con la restituzione si ha il risultato dell'analisi del contenuto. I *software* e gli hardware recenti hanno avuto un ruolo importante nella velocizzazione di questi tre processi, ma le scelte discrezionali del ricercatore hanno sempre un ruolo determinante. Recentemente, l'analisi del contenuto è stata definita come «l'analisi di materiale empirico, in termini processuali, finalizzato a rispondere alle domande che il ricercatore pone ai testi, coerenti con gli obiettivi della ricerca» (Sbalchiero, 2021, p. 126). Quindi, anche in questo caso, un processo di analisi delle informazioni presenti nella raccolta di testi, ma che è coerente e si focalizza in maniera più specifica con gli obiettivi che il ricercatore definisce nella sua domanda di ricerca. Altre note definizioni di *content analysis* le troviamo con gli studi di Lasswell come «semantica quantitativa» (Lasswell, 1949) o con Berelson che la definisce «una ricerca atta a stabilire le inferenze valide e replicabili a partire dai dati, per arrivare al relativo contesto» (Berelson, 1952). Rositi parla di «insieme di metodi che sono orientati al controllo di determinate ipotesi su fatti di comunicazione [...] e che a tale scopo utilizzano procedure di scomposizione analitica e di classificazione normalmente a destinazione statistica» (Rositi, 1998). Losito invece la definisce in maniera generale come «un insieme di procedure che possono essere applicate su oggetti di ricerca molteplici e differenziati» (Losito, 1993) e anche Amaturò parla di «una tecnica di scomposizione di qualunque tipo di messaggio [...] in elementi costitutivi più semplici, di cui è possibile calcolare la ricorrenza» (Amaturò). Infine, Nobile parla di un processo quantitativo nel quale «ricadono tutte quelle tecniche capaci di produrre dati relativi alla misurazione e al conteggio di elementi testuali» (Nobile, 1997) ma anche «l'insieme di tecniche atte a studiare in maniera standardizzata il contenuto di messaggi rilevanti da un punto di vista sociale, culturale o politico ben specifico» (Nobile, 1997).

In tutte le definizioni presenti in letteratura e, che ho riportato, due sono gli aspetti fondamentali che emergono: l'aspetto linguistico e l'aspetto statistico. Processi multidisciplinari che sono complementari e non esclusivi e servono per il raggiungimento dello scopo di ricerca prefissato anche se nella storia della disciplina, le varie scuole di pensiero, quantitative e qualitative, si sono spesso scontrate e contrapposte soprattutto per quello che riguarda la miglior interpretazione dei risultati. È importante mettere in evidenza che nel processo di analisi del contenuto «le scelte del ricercatore non solo sono

parte integrante del processo, ma ne rappresentano il vero punto di forza» (Tuzzi, 2003, p.19). La possibilità discrezionale del ricercatore quindi risiede nella scelta delle tecniche migliori da utilizzare come strategia efficace per il raggiungimento dell'obiettivo prefissato.

È solo con la definizione di Sbalchiero (Sbalchiero, 2021) però che vengono messi in evidenza gli scopi del ricercatore come obiettivo della ricerca perché, nelle varie definizioni fino ad ora trattate, si parla solamente di tecniche utilizzate nel processo di analisi senza prendere in considerazione l'obiettivo che spinge il ricercatore nella sua analisi. È ancora Sbalchiero a definire l'analisi del contenuto «come una metafora dell'approccio comprendente» (Sbalchiero, 2021, p.126) (visto come comprensione sociologica) per ribadire che «non si tratta di comprendere i comportamenti umani in modo intuitivo e simpatetico, ma di renderli intelligibili in un progetto di conoscenza intellettuale e razionale» (Schnapper, 2005, p. 172). Quindi un processo di analisi visto come un unico percorso di ricerca completo che comprende

«i momenti che vanno dalla selezione delle unità di analisi, alle sintesi interpretative fino alla costruzione di un sistema di categorie per fornire, infine, quegli elementi empirici utili per valutare le relazioni tra tali categorie concettuali, a un livello di astrazione sufficientemente elevato da poter procedere con la costruzione di strumenti adeguati per comprendere e spiegare il fenomeno». (Sbalchiero, 2021 p. 126)

La finalità di base per cui l'analisi del contenuto è nata, è quella di permettere la scomposizione dei messaggi di cui si compone una data comunicazione in modo tale da ricavarne i significati impliciti e latenti insiti nel contenuto (Pandolfini, 2017). I messaggi possono essere di tipo testuale verbale oppure anche fonti audio e audiovisive opportunamente codificate. L'utilizzo dei *software* ha portato ad avere risultati immediati e relativi a grandi quantità di testi. È però importante sottolineare che la credibilità dei risultati è dovuta sia alla trasparenza delle procedure e, alla possibilità di replicarle e controllarle, sia alla capacità e alla conoscenza, da parte del ricercatore, nell'interpretare i risultati. Troppo spesso l'utilizzo del *software* viene visto come l'unica chiave di lettura per comprendere i risultati, piuttosto che uno strumento di supporto, senza dimenticare che alcuni *software* sono nati nell'ambito di una specifica scuola di pensiero, che li sviluppa nell'ottica di rispondere alle proprie esigenze di ricerca (Tuzzi, 2033, p. 157).

3.1.3 Il *corpus*

Fino ad ora abbiamo parlato di *content analysis* in generale senza focalizzarci sul materiale empirico di riferimento per l'analisi del contenuto: il *corpus*. Il *corpus*, infatti, è una «collezione di testi che raccoglie testi coerenti con gli scopi perseguiti dalla ricerca e questa coerenza è valutabile solo discrezionalmente» (Tuzzi, 2003, p. 29). Oltre a essere una raccolta di testi, il *corpus*, deve rispettare caratteristiche di ampiezza, coerenza e omogeneità. Principi indispensabili per un'analisi coerente. «Un *corpus* è una raccolta di testi selezionati e organizzati secondo espliciti criteri, allo scopo di essere usata come campione rappresentativo del linguaggio o di una sua particolare varietà» (McEnery & Wilson, 2001). È quindi un campione rappresentativo di testi affini ed inerenti ad uno specifico argomento, che non può essere assemblato con materiali provenienti da più fonti e con argomenti differenti. Se pensiamo ai messaggi di un politico che divulga tramite *tweet*, il *corpus* non può raccogliere anche i suoi comunicati stampa oppure i testi dei messaggi sms, *WhatsApp*. Stessa cosa vale qualora la ricerca fosse inerente alle opere letterarie di un determinato autore, il *corpus*: comprenderebbe solo questi testi e non anche gli appunti o le tracce che ha utilizzato per scrivere le sue opere. Il fine, quindi, non è meramente statistico, ma deve seguire l'obiettivo e la domanda di ricerca stabiliti. Il materiale raccolto nel *corpus* può poi essere anche suddiviso in sub testi o frammenti, secondo un criterio ragionato di suddivisione deciso dal ricercatore che utilizza un raggruppamento, una classificazione. Se prendiamo per esempio i testi di un'opera letteraria di un autore, possono essere considerate classificazioni o raggruppamenti i paragrafi o i capitoli dell'opera stessa.

3.2 La codifica e gli approcci classici e moderni

3.2.1 Le unità di analisi e le categorie concettuali nel processo di codifica

I momenti principali del processo di codifica nell'analisi del contenuto sono: la selezione delle unità di analisi e la costruzione delle categorie concettuali. Le unità di analisi rappresentano l'oggetto al quale si riferiscono le proprietà studiate, sono l'elemento principale per l'analisi del fenomeno sociale e si differenziano nei due approcci

quantitativo e qualitativo. Nella prima metodologia l'unità di analisi è definita come l'individuo studiato o un gruppo/organizzazione di individui o un evento sociale oppure un messaggio di comunicazione visto come rappresentazione simbolica. Nell'approccio qualitativo invece, l'unità di analisi è una porzione di testo, vista come un paragrafo o un periodo o delle frasi di senso compiuto. Queste porzioni di testo si riferiscono chiaramente agli attori sociali e sono l'espressione delle loro azioni, dei loro comportamenti e dei loro atteggiamenti.

Definite le unità di analisi, attraverso un processo sistematico di lettura e riletura del materiale empirico, si vanno a definire le categorie concettuali, abbinando delle proprietà a porzioni del testo ritenute rilevanti. Due sono i passaggi principali da elaborare: il primo, più operativo, per l'assegnazione di etichette (codici) alle parti più significative del testo (espressioni di spicco o parole vivide con un alto valore semantico) ed un secondo, più astratto, che andrà a definire delle categorie riferite a codici con caratteristiche simili. Per fare un esempio pratico, se il mio materiale è composto da testi che si riferiscono ai comunicati stampa dell'Università di Padova, posso dapprima segnalare tutti i temi che trattano (che saranno le mie etichette) per poi definire delle categorie: didattica, servizio degli studenti, percorsi di laurea, eccetera. Dalle categorie si possono poi definire anche delle macrocategorie che hanno un livello di astrazione superiore rispetto alle categorie. Viene quindi utile e importante ribadire che «il fine dell'analisi del contenuto è la costruzione di categorie concettuali in grado di restituire i contenuti dei testi sotto una nuova forma» (Sbalchiero, 2021) che saranno indispensabili e necessarie per interpretare i risultati e comprendere il fenomeno sociale studiato. Dal punto di vista metodologico due sono i percorsi adottati nell'analisi del contenuto riguardo la costruzione delle categorie concettuali (Figura 06), uno classico che si rifà alla scuola USA degli anni Venti, di tipo giornalistico, in cui la struttura dei concetti del mio *corpus* viene eseguita in maniera essenzialmente manuale e solo parzialmente automatizzata e, uno moderno, nato agli inizi degli anni Sessanta in Francia, dove la codifica delle unità testuali viene eseguita essenzialmente con l'utilizzo automatizzato dei computer e solo parzialmente manuale che si rifà ad un approccio lessicale e alla linguistica quantitativa.

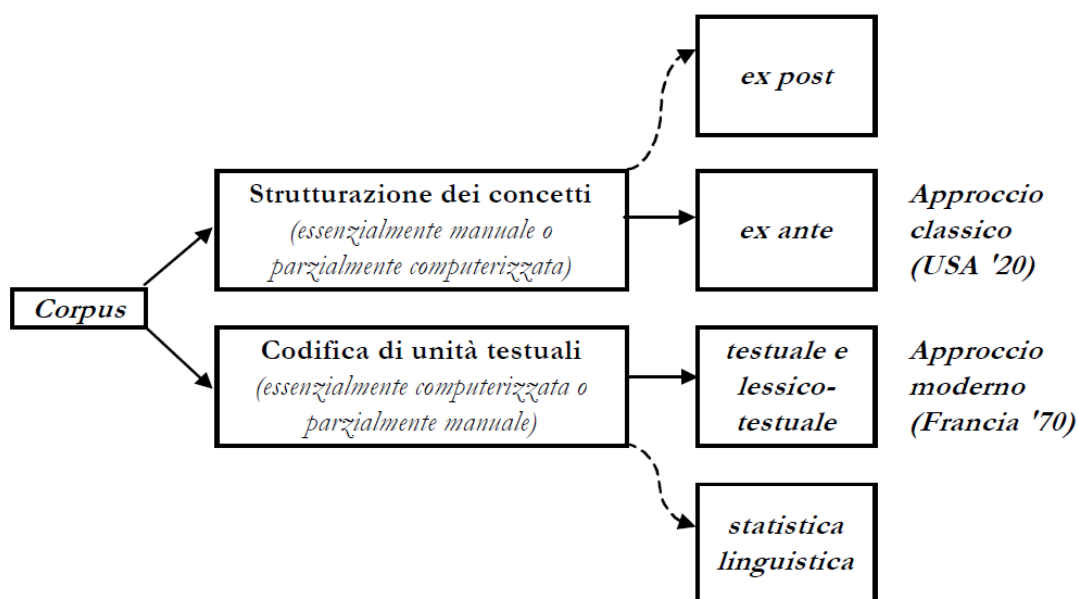


Figura 06. I due percorsi di codifica dell'analisi del contenuto (Tuzzi, 2003, p. 35).

3.2.2 L'analisi del contenuto classica

L'analisi del contenuto classica ci riconduce alle ricerche effettuate alla nascita della disciplina, dai ricercatori degli anni Venti in USA. Ho già parlato quindi nei paragrafi precedenti dell'analisi del contenuto in Svezia dei testi dell'opera del XVII secolo "I Canti di Sion" voluta dalla Chiesa Luterana per verificare se vi erano valori in contrasto con l'ortodossia, delle ricerche di Thomas e Znaniecki del 1918 riguardo le lettere dei contadini polacchi emigrati in America, dell'analisi degli slogan del 1927 di Lasswell e di quelle relative agli articoli del periodico Galileian per capire i collegamenti semantici con ideali antiamericani. Il percorso non è stato omogeneo e molte volte è stato criticato anche a causa del massiccio utilizzo di tecniche quantitative che non davano spazio ad approcci qualitativi, con la conseguenza di avere una netta separazione tra osservatore ed osservato e poca valenza dei contesti. Parallelamente all'eccessivo utilizzo di tecniche che privilegiavano il conteggio delle parole e delle frequenze, vi era poi poca trasparenza sulla metodologia delle analisi. Le ricerche hanno comunque dato vita ad una nuova

disciplina, la *content analysis* che dopo una breve battuta d'arresto negli anni Cinquanta, ha avuto una notevole diffusione e molto interesse grazie al supporto dei *personal computer* nei vari processi di acquisizione, codifica e analisi dei testi, ma anche grazie alla disponibilità di corpora testuali già in formato elettronico. L'analisi del contenuto classica è prevalentemente un'analisi di tipo qualitativo e si utilizza in quei contesti dove il numero di categorie è limitato. L'obiettivo è quello di riconoscere e sintetizzare in maniera perlopiù manuale le informazioni rilevanti presenti nei testi. Oltre alle tecniche manuali vengono utilizzate delle tecniche assistite, *CAQDAS*, *computer assisted qualitative data analysis software*. Le regole per identificare le categorie, ma anche le regole per associare porzioni di testo alle categorie, sono stabilite dal ricercatore.

Per capire bene la metodologia utilizzata nell'analisi del contenuto classica possiamo fare riferimento alle procedure che di solito le persone adottano quando devono riassumere un capitolo di un libro, la trascrizione di un'intervista o un articolo di giornale. Solitamente si sottolinea o si evidenziano le parti principali, che più rappresentano il contenuto del testo. È proprio questo l'obiettivo dell'analisi del contenuto classica: riconoscere e sintetizzare informazioni rilevanti contenute nei testi utilizzando tecniche manuali o *computer assistant* per rilevarle e definire delle categorie (decise dal ricercatore) da associare alle porzioni di testo. Le categorie sono raggruppate in una griglia (vedremo in seguito la prospettiva *ex post* ed *ex ante*) e ogni porzione di testo (che è la nostra unità statistica) ha il suo collegamento alla categoria di riferimento. Possiamo avere un raggruppamento dei concetti di tipo *whitin* (porzioni di testo simili dentro la stessa categoria) o *between* (più possibili diverse e in più categorie) ma è opportuno operare per macrocategorie quindi partire da una categoria generale con un livello di astrazione alto (per esempio categoria animale) per poi affinare in varie sotto categorie (mammiferi, domestici, vertebrati). Essendo l'analisi del contenuto classica perlopiù manuale è normale che si verificano situazioni di ambiguità di assegnazione delle porzioni di testo a una determinata categoria piuttosto che all'altra, è importante quindi, motivare bene le scelte e tralasciare, le situazioni *borderline*, essere obiettivi e utilizzare gli stessi criteri su tutte le porzioni di testo. Il processo di codifica è un processo ricorsivo che permette un perfezionamento continuo e quindi è bene rivedere più volte i frammenti di testo che non vengono categorizzati e quelli su cui si è indecisi. Da tener presente che i codificatori poi sono degli umani, e molto spesso più di uno, ed è importante che si confrontino fra

loro per arrivare a degli accordi e definire le categorie che meglio rispecchiano il contenuto del testo da analizzare. La creazione della griglia per l'analisi del contenuto organizzata per categorie (variabili) è un processo definito dal ricercatore e può essere gestito secondo due processi di strutturazione:

1. *ex ante* quando la griglia con le categorie concettuali è scelta prima della lettura dei testi
2. *ex post* quando la griglia con le categorie concettuali nasce dalla lettura dei testi, quindi dopo

Se la mia ricerca è già stata effettuata o sono di fronte ad uno studio che ha similitudini con altri, posso utilizzare la prospettiva *ex ante*. Ho il vantaggio in questo caso di avere categorie già approvate, con un notevole risparmio di tempo e la possibilità di comparare gli studi simili. C'è da tener presente però che potrei incorrere negli stessi errori avuti nello studio simile e che la prospettiva *ex ante* non è adatta per studi che hanno un'evoluzione veloce o per lo studio di nuove realtà. Se invece siamo di fronte ad un nuovo studio è bene utilizzare il processo *ex post*, creando la mia matrice di dati con le categorie concettuali scelte dopo la lettura dei testi. È un lavoro che occuperà molto tempo, spesso ci saranno molti frammenti di testo non categorizzati e non sempre le categorie scelte dai vari codificatori saranno comparabili.

L'analisi del contenuto classica è quindi un'analisi essenzialmente manuale con metodi qualitativi assistiti da *software*, la raccolta dei dati, la codifica e la loro interpretazione sono un processo continuo che si evolve (secondo la *Grounded Theory*) ed è quindi importante creare categorie generali per poi o affinarle in sottocategorie oppure aggregarle in categorie simili. L'utilizzo del *software* è indispensabile quando siamo di fronte a grandi quantità di testo da analizzare, ed è utile anche per creare, in automatico, mappe concettuali o interconnessioni tra le varie categorie. Il suo utilizzo (visto unicamente come supporto all'analisi del contenuto e che non sostituisce il momento interpretativo) va però inserito in un discorso più ampio e metodologico che come Weber insegna, prevede una riflessione generale sull'insieme delle procedure, dei metodi, delle tecniche, degli strumenti utilizzati ed «è una buona regola tenere a mente che gli strumenti sono indispensabili e forniscono un aiuto prezioso, ma che le analisi procedono guidate dalle intuizioni, dalle interpretazioni, dalle idee del ricercatore» (Sbalchiero, 2021, p. 138).

3.2.3 L'analisi del contenuto moderna e il metodo di Reinert

3.2.3.1 L'approccio moderno dell'analisi del contenuto

L'approccio moderno, o automatico, dell'analisi del contenuto nasce in Francia agli inizi degli anni Settanta, in quegli anni, infatti, l'analisi del contenuto ha una nuova spinta, dopo il periodo di *débâcle* dovuta all'etichettatura di disciplina "quanto frenetica", grazie soprattutto ai linguisti e al nuovo approccio testuale e lessico-testuale. La metodologia introdotta dalla scuola francese *Analyse des Données*, e il nuovo approccio della codifica del testo va oltre all'analisi delle sole frequenze e si orienta alla logica delle relazioni tra le categorie concettuali. Parallelamente i nuovi sistemi automatici garantiscono processi quantitativi utilizzati per far emergere informazioni che difficilmente analisi qualitative darebbero.

L'analisi del contenuto moderna è un processo automatizzato e parzialmente manuale che si prefigge l'obiettivo di sfruttare, a fini statistici, le informazioni rilevate in *corpora* di grandi dimensioni. È un approccio lessico-metrico basato sull'analisi di profili lessicali e quindi sulla distribuzione delle occorrenze delle parole. La disciplina abbraccia numerosi settori dal *computational linguistics* al *quantitative linguistics* dall'*information retrieval* al *natural language processing*, dal *pattern recognition* alla *machine learning*, per lo sviluppo di applicazioni industriali quali i dizionari elettronici, il riconoscimento vocale, la traduzione automatizzata e l'intelligenza artificiale. I *software* che gestiscono corpora testuali lavorano sulle forme grafiche, che sono stringhe di caratteri e si identificano nel concetto di parola e che formano le unità statistiche o unità di analisi testuale. Per comprendere meglio possiamo prendere come esempio questa frase (scioglilingua):

"Caro conte chi ti canta tanto canta che ti incanta".

Il numero di unità statistiche o forme grafiche totali è pari a 10.

"Caro (1) conte (2) chi (3) ti (4) canta (5) tanto (6) canta (7) che (8) ti (9) incanta (10)".

Il numero di modalità o parole diverse è invece pari a 8.

"Caro (1) conte (2) chi (3) ti (4) canta (5) tanto (6) canta (5) che (7) ti (4) incanta (8)"

Sono però i termini inglesi utilizzati per la maggiore e quindi: forme grafiche totali o *word token* e parole differenti dette *word type*.

Modalità	Forma grafica	Occorrenza	Frequenza	Freq. cumulata
1	caro	1	1/10= 0,1=10%	10%
2	conte	1	1/10= 0,1=10%	20%
3	chi	1	1/10= 0,1=10%	30%
4	ti	2	2/10=0,2=20%	50%
5	canta	2	2/10=0,2=20%	70%
6	tanto	1	1/10= 0,1=10%	80%
7	che	1	1/10= 0,1=10%	90%
8	incanta	1	1/10= 0,1=10%	100%
		Tot: 10	Tot: 100%	

Tabella 03. Vocabolario di frequenza.

La tabella (Tabella 03) che riassume *word token*, *word type*, occorrenze e frequenze, è detta vocabolario di frequenza e identifica una rappresentazione del testo preso in esame (che nel mio caso si riferisce alla singola frase, ma che potrebbe essere relativo all'intera raccolta di testi).

Il percorso tipico dell'analisi del contenuto automatica è diviso in quattro fasi principali che vanno dalla normalizzazione del mio *corpus*, con la raccolta di tutti i miei testi in un file con estensione “.txt” e formato utf8 (codifica standard), alla codifica del *corpus* dove vengono eliminati i caratteri presenti solo una volta (detti *hapax*) ed eventualmente le forme grafiche ripetute, alla selezione delle unità testuali (che sono le forme grafiche definite da analizzare) e infine l'analisi statistica vera e propria. Sono molti e in continua evoluzione i *software* per l'analisi automatica dei dati testuali. Per citarne alcuni, NVivo e RQDA *software* per l'analisi qualitativa dei dati testuali o IRaMuTeQ (che utilizzerò per il mio studio) che permette di effettuare analisi statistiche su *corpora* di testi con un approccio semantico e basato sul metodo di classificazione di Max Reinert, che approfondirò nel prossimo paragrafo.

L'analisi del contenuto moderna è quindi orientata all'individuazione di argomenti, *topic detection* e alla categorizzazione di contenuti con analisi statistiche che si orientano alla linguistica. L'utilizzo principale del *software*, come nell'approccio classico, non deve in nessun modo sostituire il ruolo del ricercatore che è il principale protagonista per l'interpretazione dei dati. Da tener presente poi che «non è sulla bontà dello strumento

che risiede l'efficacia dell'analisi, ma sulla capacità del ricercatore di usarlo sviluppando la consapevolezza del ruolo che lo strumento stesso gioca nel processo di ricerca» (Sbalchiero, 2021, p. 139).

3.2.3.2 Il metodo di Reinert

L'approccio informatizzato dell'analisi del contenuto ha portato alla nascita di numerosi *software* utilizzati come strumenti di supporto per la codifica automatica di corpora di grandi dimensioni. L'evoluzione degli aspetti informatici, insieme alla necessità di creare una sintesi tra approcci qualitativi e quantitativi, ha aperto le porte al metodo di Reinert, un approccio semantico del testo basato sulla classificazione di frammenti testuali. Il metodo di Reinert è stato implementato dapprima nel *software* Alceste (*Analyse Lexicale par Contexte d'un Ensemble de Segment de Texte*) e poi in IRaMuTeQ (*Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires*). L'obiettivo principale è quello di estrapolare dal testo i principali argomenti (*topic detection*) attraverso l'analisi delle co-occorrenze, al fine di identificare le varie "classi semantiche" che Reinert chiama "mondi lessicali". Le co-occorrenze sono la presenza concomitante di più parole in un testo che l'algoritmo rileva per la costruzione di una matrice di contingenza. Le unità di analisi prese in considerazione in automatico dal metodo di Reinert possono essere delle frasi, un paragrafo o una parte di testo e sono individuate attraverso due criteri: tramite il numero massimo di parole di cui sono composte e attraverso la punteggiatura. Questo permette di suddividere il *corpus* in parti con lunghezza simile da analizzare. Le classi semantiche o mondi lessicali sono invece una serie di parole associate tra loro in grado di restituire un concetto espresso, sono «espressioni osservabili dei "luoghi di pensiero", degli "spazi mentali" su cui il locutore si sofferma per parlare di un certo argomento» (Reinert, 1998).

L'utilizzo del *software* e quindi dell'algoritmo implementato nel metodo di Reinert aiuta il ricercatore a estrapolare in maniera automatica insiemi di parole rilevanti per ciascuna classe semantica necessarie per interpretarne il contenuto. A differenza di altri *software* meramente qualitativi, qui siamo di fronte al risultato di un'analisi che sarà utile al ricercatore per produrre le proprie interpretazioni. L'utilizzo del *software* e la capacità di

restituire gli stessi risultati replicando la stessa analisi (attendibilità) non devono essere visti come sinonimo di validità della ricerca, ma devono essere utilizzati in una riflessione unitaria insieme agli obiettivi, ai metodi, alle tecniche e alle procedure utilizzate nel percorso di analisi del contenuto, per arrivare poi all'interpretazione dei risultati.

CAPITOLO 4

Metodologia e analisi del *corpus* con IRaMuTeQ

4.1 Metodologia

4.1.1. La raccolta dei discorsi dei papi

La ricerca che ho sviluppato in questa tesi è un'analisi del contenuto dei discorsi dei vari papi che si sono succeduti nella storia della Chiesa. In particolare, ho raccolto tutti gli scritti degli ultimi cento anni, da Papa Pio XI (eletto nel 1922) fino a Papa Francesco, inglobando nella mia raccolta di testi, i discorsi fino al dicembre 2021. Il tema che ho preso in considerazione è l'Università, quindi le argomentazioni dei discorsi che si riferiscono all'ente stesso oppure alle sue componenti, corpo docenti o studenti.

I testi sono stati presi in parte dal sito dell'Associazione Italiana Docenti Universitari (<http://inx.aiduassociazione.it/>) e in parte dal sito dello Stato del Vaticano (<https://www.vatican.va/content/vatican/it.html>), unica fonte ufficiale relativa alla storia e alla vita dei vari papi.

Con il termine discorsi si intende la trascrizione della prolusione che i papi hanno fatto incontrando le varie realtà accademiche, non mi sono occupato quindi dei messaggi di saluto oppure delle lettere che vengono inviate dal pontefice e che sono lette in concomitanza di altre particolari ricorrenze.

La tabella che segue (Tabella 04) raccoglie i testi dei papi scritti con riferimento alle Università negli ultimi cento anni e segnala, per ogni papa, il numero di discorsi prodotti annualmente, gli anni di papato, l'anno di elezione e la media di testi scritti per anno (arrotondata e senza cifre decimali).

Papa	N° testi	Totali	Anno elezione	Anni papato	Media testi per anno
PIO XI	1	1	1922	17	0
PIO XII	4	5	1939	19	0
GIOVANNI XXIII	3	8	1958	5	1
PAOLO VI	30	38	1963	15	2
GIOVANNI PAOLO II	168	206	1978	27	6
BENEDETTO XVI	38	244	2005	8	5
FRANCESCO	31	275	2013	8	4

Tabella 04. I papi ed i testi prodotti con riferimento all'Università negli ultimi cento anni.

Dopo un'attenta valutazione dei discorsi raccolti e vista anche la produzione dei discorsi annuali per ogni papa (Tabella 04), ho deciso di prendere in esame quelli riferiti ai papi che hanno scritto almeno due testi per anno e quindi dal 1963 (elezione di Paolo VI), fino a dicembre 2021 (con l'attuale Papa Francesco). Sono in totale duecentosessantasette testi che andranno a costituire il mio *corpus* (Tabella 05).

Papa	N° testi	Totali	Anno elezione	Anni papato	Media testi per anno
PAOLO VI	30	30	1963	15	2
GIOVANNI PAOLO II	168	198	1978	27	6
BENEDETTO XVI	38	206	2005	8	5
FRANCESCO	31	267	2013	8	4

Tabella 05. I papi ed i testi prodotti nei vari anni del loro papato presi in considerazione.

Rispetto all'oggetto di ricerca il papa più produttivo risulta essere stato Papa Giovanni Paolo II con una media di sei testi per anno, mentre il meno produttivo Paolo VI con due discorsi in media per anno, pur avendo effettuato quindici anni di papato (Figura 07 e Figura 08).

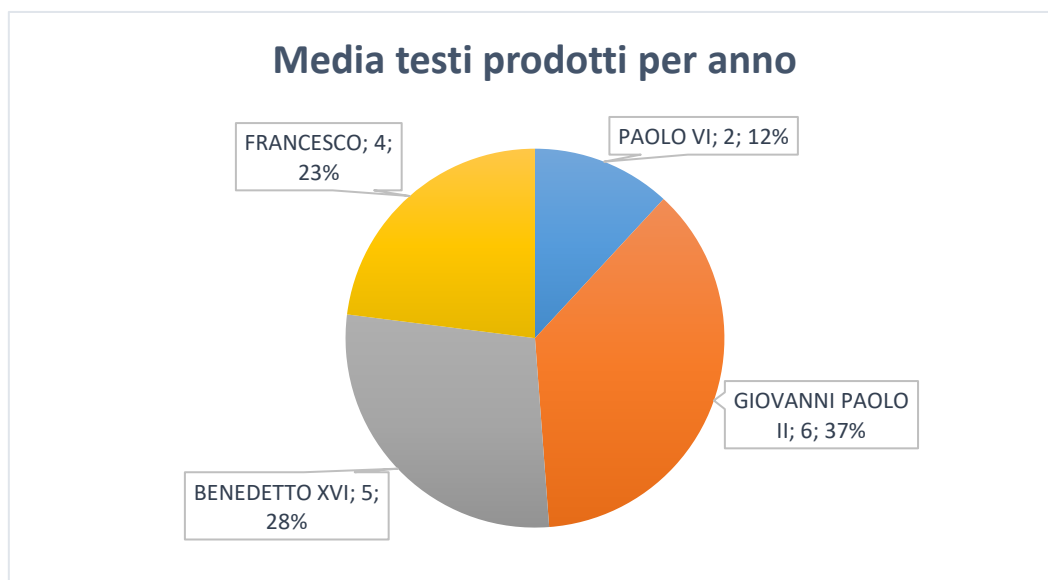


Figura 07. Grafico della media dei testi prodotti annualmente da ogni papa.

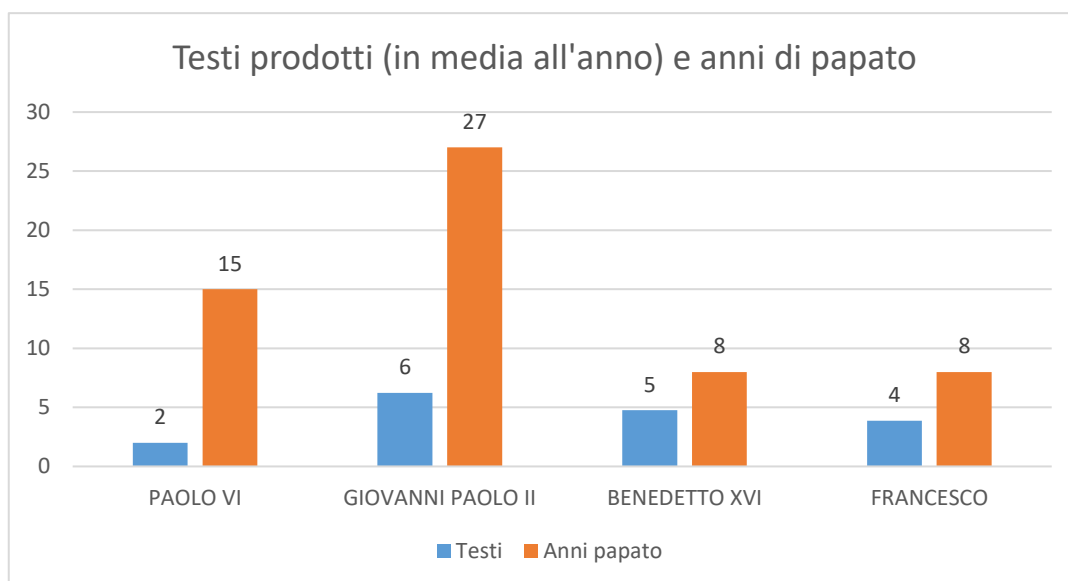


Figura 08. Grafico della media dei testi prodotti e anni di papato per ogni papa.

L'anno di partenza preso in considerazione per questa analisi è, come anticipato, il 1963. Facendo un collegamento con la storia vaticana di quel periodo, gli anni Sessanta sono segnati per l'elezione del conclave di Paolo VI che viene ricordato come un conclave duro per via del Concilio Vaticano II rimasto aperto e privo di una conclusione a causa della morte di Giovanni XXIII. Il Concilio, infatti, appare subito come la vera discriminante per l'elezione e l'ultima possibilità di invertire la strada innovatrice che il Concilio stesso aveva intrapreso. Molte sono le riserve sul candidato a pontefice, il Cardinale Montini, sia da parte degli elettori sia dalla parte politica (tantoché pure il Presidente della Repubblica Italiana Antonio Segni aveva espresso ai cardinali il suo dissenso alla candidatura Montini). Dopo tre giornate e sei scrutini viene comunque eletto il nuovo papa, Montini, che assunse il nome di Paolo VI, nonostante un gruppo di ventidue cardinali mantenne fino all'ultimo il proprio rifiuto.

Gli anni Sessanta sono anche ricordati nella storia perché nelle scuole è stato l'ultimo decennio di utilizzo della penna ad inchiostro che viene poi sostituita dall'introduzione della penna a sfera. Per quel che riguarda l'aspetto tecnologico c'è da tenere presente che negli ultimi cento anni la nascita delle prime macchine di videoscrittura e poi la diffusione dei personal computer hanno portato a notevoli progressi. Nel contempo il supporto ai

discorsi papali è passato dalla carta al digitale, da un metodo di scrittura lento e preciso quindi (come quello su carta) ad uno veloce e sistematico (come quello a computer). Si tratta di differenze rilevanti di cui tener conto nella produttività dei testi che i pontefici hanno avuto ma soprattutto per quel che riguarda il recupero delle informazioni in rete, quelle degli ultimi anni già presenti sui siti di riferimento, quelle più datate invece che hanno avuto un processo laborioso di trascrizione.

Dal marzo 2020, a causa della pandemia, le conferenze e le uscite del Papa (viaggi apostolici), sono sensibilmente diminuite, andando ad incidere anche sulla produzione dei discorsi ufficiali del Pontefice.

4.1.2. La costruzione del *corpus*

Il mio *corpus* raccoglie tutti i discorsi relativi all'oggetto di studio, nello specifico duecentosessantasette testi. Per la costruzione del *corpus* ho utilizzato il programma "Blocco Note" grazie al quale è possibile creare dei *files* con estensione ".txt" e codifica standard "UTF8" (che vale per tutte le lingue e di *default* elimina tutti i caratteri di formattazione del file che potrebbero creare problemi nell'elaborazione).

Le variabili prese in esame sono cinque:

- 1) "papa", che identifica il nome del papa autore del testo
- 2) "anno" riferito all'annata di redazione del testo
- 3) "mese" relativo al mese in cui è stato redatto il testo (associato all'anno)
- 4) "destinatario" relativo al ricevente del discorso
- 5) "argomento" (visto come incontro con attori differenti) e nello specifico la tipologia (variabile) "argomento" si divide in tre tipi (modalità):
 - 1) "vipastorale" l'incontro del papà con le varie realtà esterne dal Vaticano
 - 2) "inpastorale" l'incontro del papa con le varie realtà all'interno del Vaticano
 - 3) "viapostolico" l'incontro del papa con le varie realtà accademiche durante i viaggi apostolici esteri e con una durata di più giorni.

Andando più nel dettaglio la variabile "argomento", la cui tipologia si divide in tre tipi come specificato prima, ha le seguenti caratteristiche:

l'incontro pastorale (chiamato "inpastorale" nel *corpus*) è un incontro del papa con una delegazione di persone accademiche, che è differente dalle udienze, le quali invece sono pubbliche e a cui tutti possono partecipare. Gli incontri pastorali avvengono all'interno del Vaticano e si organizzano, negli ultimi anni, o nella Biblioteca Apostolica o nel cortile di San Damaso del palazzo apostolico o nella sala della Clementina a fianco della basilica di San Pietro. La visita pastorale ("vipastorale" riferita al *corpus*) nella Chiesa cattolica è invece la visita di un vescovo ai luoghi e alle persone della sua diocesi, nel mio caso corrisponde alla visita del papa, esterna al Vaticano, alla realtà accademica. Il viaggio apostolico infine ("viapostolico" nel *corpus*) equivale al viaggio intrapreso da un papa per più giorni in Italia ma nel mio caso all'estero.

Specifico, inoltre, che il sito ufficiale del Vaticano (<https://www.vatican.va/content/vatican/it.html>) da dove ho raccolto parte dei testi, segnala all'inizio di ogni trascrizione se il discorso è riferito ad una visita pastorale o a un viaggio apostolico.

Il *corpus* poi per essere idoneo alla elaborazione con il *software* deve avere delle caratteristiche di qualità riferite al dato, rilevabili nei concetti di:

- ampiezza (numero delle parole differenti usate nel *corpus*)
- dimensione (numero delle parole che compongono il *corpus*)
- accuratezza (la vicinanza del valore trovato a quello del dato reale);
- coerenza (valore di un dato non contraddittorio con gli altri);
- omogeneità (parametri di una stessa popolazione di dati con valori identici);
- esaustività (dato completo di tutte le informazioni).

Le comunità scientifiche hanno definito poi varie dimensioni che si riferiscono ad altri concetti di *data quality*, come la consistenza, la completezza, la correttezza che fanno riferimento alla dimensione della qualità (Batini, Scannapieco, 2006).

Nello specifico, un *corpus*, per essere trattato statisticamente e in maniera automatica con un *software*, deve avere un *token type ratio*, vista come percentuale del rapporto tra numero di *hapax* (parole che compaiono una sola volta nel *corpus*) e il numero delle forme (numero di parole diverse), minore del 50%. Se la percentuale è maggiore, risulterebbe troppo bassa la ricchezza del vocabolario e, quindi, il *corpus* non sarebbe trattabile e idoneo.

4.1.3. L'analisi del contenuto classica sul *corpus*

L'analisi del contenuto classica, come accennato nei capitoli precedenti, è un'analisi di tipo manuale e senza l'ausilio di *software* automatici. Durante la costruzione del *corpus* (avvenuta mediante la copia dei discorsi originali dai siti di riferimento nel *file* con estensione .txt) ho avuto modo di controllare, seppur in maniera rapida e senza una specifica lettura, il contenuto dei vari testi. Ho quindi svolto una prima forma di *content analysis*. Ogni testo, infatti, a seconda dell'autore (un papa nel nostro caso), si identifica con un saluto iniziale, un corpo centrale e una parte finale; quest'ultima di solito si ripete in ogni discorso, perché riguarda atti religiosi. Ogni papa poi ha un suo stile di scrittura, Papa Francesco termina, ad esempio, scrivendo quasi sempre la frase "ricordatevi di pregare per me". Ho notato, inoltre, che il tipo di linguaggio varia a seconda dell'interlocutore e del tema, e si rileva molto bene l'approccio informale di Papa Francesco (che utilizza termini meno ricercati e molto più popolari) rispetto a quello formale e di Benedetto XVI (che nella stesura dei suoi scritti è molto rigoroso e rispettoso nei confronti dei suoi interlocutori). Nei vari discorsi papali, e per quasi tutti i papi, vengono trattati i temi della fede, della scienza e della ricerca: il rapporto tra chiesa e chi si occupa del sapere, come la scienza può aiutare la religione e viceversa, il ruolo e il rapporto della religione con la scienza, il cambiamento culturale e il cristianesimo. Alcuni discorsi poi, anno per anno, si ripetono, per quel che riguarda il destinatario, perché ogni papa è solito andare annualmente al congresso nazionale della Federazione universitaria cattolica italiana (F.U.C.I.), per esempio, piuttosto che all'incontro pastorale con i referenti degli atenei romani. Ripetendosi i destinatari posso immaginare che si ripetano anche i discorsi e il linguaggio utilizzato, ma per questo aspetto mi affido al controllo automatico attraverso il *software*. La mia analisi di tipo classico si è focalizzata su questi semplici particolari, che credo siano importanti e che nei capitoli che seguono andranno interpretati in vista anche del metodo moderno di *content analysis* applicato al *corpus*.

4.1.4. Il *software* IRaMuTeQ

Il *software* utilizzato per il mio lavoro di analisi del contenuto è Iramuteq (il cui nome è così previsto: IRaMuTeQ e significa interfaccia R per l'analisi Multidimensionale del Testo e dei Questionari) è un *software* francese che basa il suo funzionamento sui linguaggi di programmazione R e *Python*. Iramuteq permette, a sua volta, di svolgere delle analisi di tipo statistico con un approccio semantico ai testi: esso è in grado di individuare i *topics* di riferimento e i mondi lessicali, evidenziandoli grazie a rappresentazioni grafiche e fattoriali. In questo modo, il *software* unisce un'analisi di tipo quantitativo-statistica a una di matrice qualitativo-semantica.

Iramuteq è un *software* scaricabile gratuitamente dal sito di riferimento <http://www.iramuteq.org/> e riproduce il metodo di classificazione definito da Reinert e descritto nel capitolo terzo della presente tesi. Consente analisi di similarità sulle forme piene, chiamate forme attive, di un *corpus* suddiviso in segmenti di testo (il cui numero di parole che compone questi segmenti è deciso dal ricercatore) e una classificazione gerarchica discendente che incrocia le forme grafiche piene e i segmenti del testo. Può fare analisi delle specificità a partire da segmenti definiti e analisi delle similitudini sulle forme piene di un *corpus*. È possibile creare dei *sub-corpus* ed esportare i metadati. Tutte le analisi vengono salvate in forma di dato su tabelle e in forma grafica (*wordcloud*, dendrogrammi, analisi delle corrispondenze, *network analysis*).

4.1.5. Le analisi con il *software* IRaMuTeQ

4.1.5.1 L'analisi lessicometrica

La prima analisi effettuata con il *software* è prettamente quantitativa e lessicale ed i dati, sono raccolti e riassunti nella Tabella 06. Il numero di testi rilevato è corretto e conferma il numero di testi che sono stati inseriti nel *corpus*, 267, il numero di segmenti di testo (con un minimo di parole per segmento di testo pari a 40 che ho scelto in fase di importazione) è pari a 10.146, le occorrenze (*word token*) sono uguali a 35.9874 e si riferiscono al totale delle frequenze assolute di ogni parola, il numero di volte che compare ogni parola nel testo.

N° parole per segmenti di testo	40
N° di testi rilevato	267
N° segmenti di testo	10.146
N° occorrenze totali (<i>word token</i>)	359.874
Media occorrenze per testo	1.347
N° forme grafiche (<i>word type</i>)	21.165
N° <i>hapax</i>	9.694
<i>Token type ratio (hapax/word type)%</i>	45,80%

Tabella 06. Indici lessicometrici.

Le forme (*word type*) sono un numero pari a 21.165 e vanno riferite alle parole differenti presenti nel *corpus*. Gli *hapax* sono 10.146 e corrispondono alle parole che compaiono una volta sola e vengono eliminate di *default* dall'analisi (sono rappresentate da errori di battitura, parole che non esistono, accorpamento di lettere ed altro non rappresentativo per l'analisi del contenuto). Posso confermare quindi che il mio *corpus* è idoneo e, è quindi, trattabile statisticamente perché il *token type ratio*, cioè la percentuale del rapporto tra *hapax* e *word type*, è pari al 45,80% e quindi minore del 50%. Questo significa che i segmenti di testo che verranno presi in considerazione per l'analisi sono ricchi di parole diverse e hanno al loro interno poche parole che compaiono una sola volta. Se il *type token ratio* fosse stato maggiore del 50%, il vocabolario sarebbe stato povero di contenuto e con poche frequenze di parole utilizzate, come se prendesse in considerazione pochi o uno solo segmento di testo. Nella mia analisi non verranno prese in considerazione le lemmatizzazioni (i plurali per esempio) perché lo scopo è quello di fare un'analisi tematica dei contenuti e non prettamente lessicale. Se le prendessimo in considerazione, le nostre forme iniziali diminuirebbero e di conseguenza cambierebbe anche la ricchezza del *corpus*.

Le forme attive (o parole piene) con maggior frequenza, sono riportate nella Tabella 07. Non sono, invece, rilevanti allo scopo della mia ricerca gli *hapax* (parole che non esistono e/o errori di battitura) e le forme supplementari (congiunzioni, articoli e altre parole che non sono significative per la mia ricerca di analisi del contenuto).

Forma attiva	Frequenza	Tipo di parola
uomo	1640	nom
università	1572	nom
dio	1278	nom
chiesa	1258	nom
vita	1057	nom
cultura	1018	nom
verità	1010	nom
mondo	872	nom
fede	855	nom
cristo	750	nom
ricerca	654	nom
oggi	653	adv
scienza	529	nom
società	479	nom
cuore	465	nom
studenti	443	nom
grande	434	adj
uomini	433	nom
giovani	430	adj
spirito	428	nom
cattolica	422	adj
servizio	421	nom
umana	417	adj
amore	394	nom
senso	384	nom
incontro	381	nom
particolare	364	adj
comunità	364	nom
proprio	353	adv
storia	345	nom
cari	342	adj

Tabella 07. Forme attive più frequenti del *corpus*.

La forma attiva più frequente è la parola “uomo” ed è presente nel mio *corpus* 1.640 volte, sintatticamente si riferisce ad un nome. Dall’analisi delle forme attive più frequenti possiamo capire in maniera generale ma immediata quale sia l’argomento principale e quali siano le argomentazioni di tutti i discorsi papali. L’uomo resta quindi al centro degli

frequenza di 1.640 e un rango uguale a 1 di frequenza). I vocabolari delle parole creati dal *software* sono, quindi, ordinati per frequenza decrescente e rango crescente. Il grafico vede sulle ascisse il logaritmo del rango delle parole e sulle ordinate il logaritmo del numero di occorrenze delle rispettive parole. Le parole più significative si collocano nella parte centrale della curva, all’inizio (quelle con rango uguale a 1) ritroviamo quelle “troppo comuni” e in coda quelle “troppo rare” e poco utilizzate nel *corpus*.

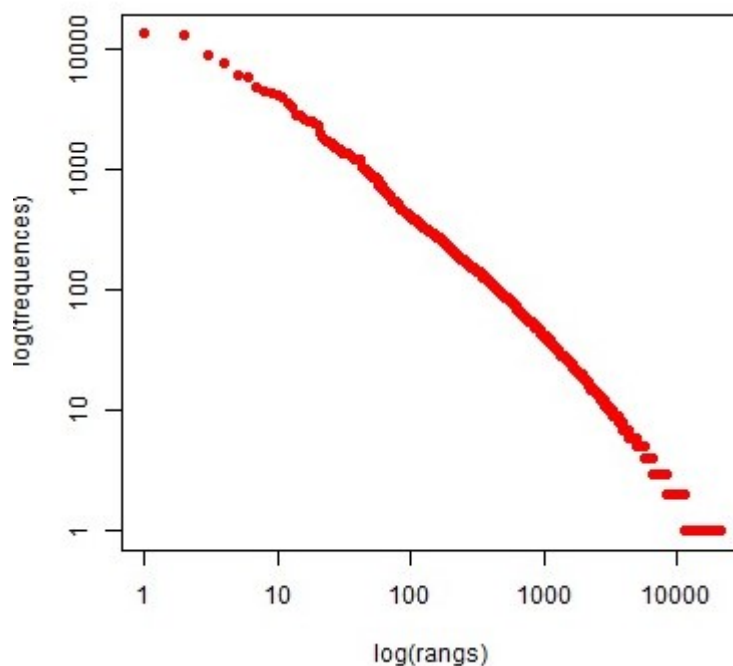


Figura 10. Grafico della legge di Zipf.

4.1.5.2 I mondi lessicali

Focus principale della presente ricostruzione è l’individuazione dei mondi lessicali, *topic detection*, applicata attraverso il metodo Reinert, che permette di costruire una matrice basata su parole e porzioni di testo a cui è successivamente applicata un *cluster analysis*. Il risultato è l’aggregazione in classi semantiche delle porzioni di testo presenti nel *corpus*; intuitivamente, ogni classe rappresenta un mondo lessicale costituito da porzioni di testo simili (che condividono parole) dal punto di vista dei contenuti. Il risultato è un’analisi automatica del contenuto i cui risultati possono essere oggetto di analisi tramite comparazioni con le variabili in analisi (papa, tipologia di discorso, anno).

I risultati ottenuti sono i seguenti e identificati in quattro classi semantiche di seguito rappresentate graficamente (Figura 11).

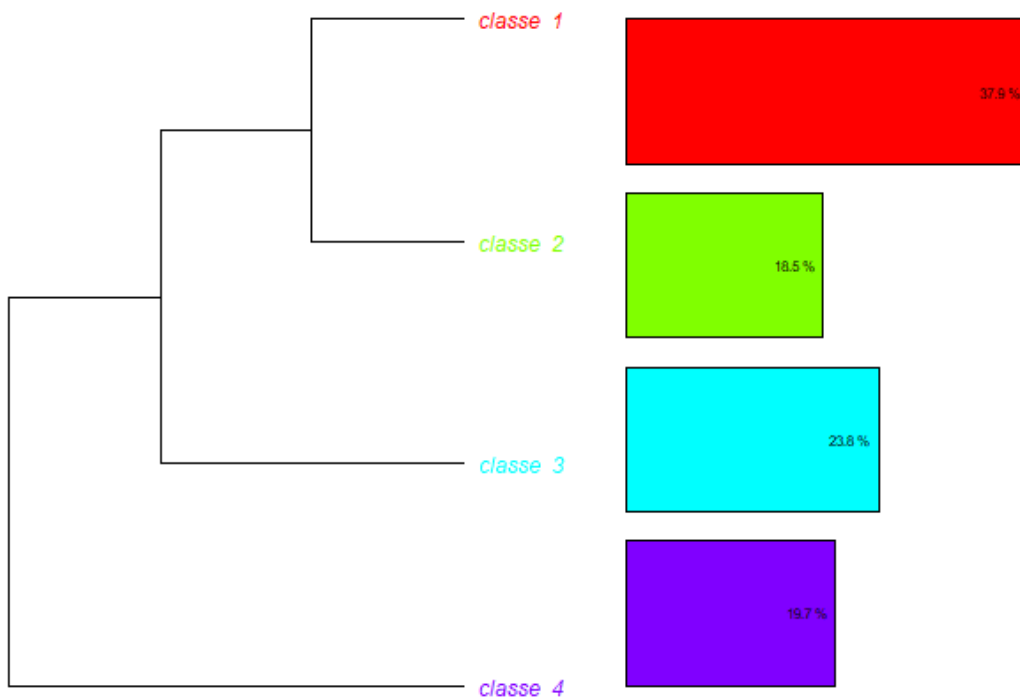


Figura 11. Rappresentazione grafica delle classi semantiche (dendrogramma).

Di questi macro-temi individuati, uno (la classe 4) si differenzia da tutti gli altri e raccoglie il 19,7% delle porzioni di testo classificate; inoltre, è presente un'altra macroarea costituita da due classi che condividono alcune parole (la classe 1 che raccoglie il 37,9% delle porzioni di testo e la classe due con il 18,5%) insieme alla classe 3 che è costituita dal 23,8% delle porzioni di testo totali.

Per ogni classe semantica vi è la possibilità di vedere graficamente (Figura 12) da quali termini essa è costituita. La classe numero quattro, raggruppa i termini come saluto, rettore, ringraziare, caro, benedizione e si riferisce quindi al mondo lessicale dei saluti e delle formalità nei vari discorsi, la classe tre è la classe riferita alla chiesa, a Dio e alla religione (ed è costituita da parole come Dio, Cristo, amore, Gesù), la classe due e la classe uno che sono semanticamente vicine (ma con un chi quadro sufficiente per essere definite come indipendenti) individuano i tematiche legate al sociale e alla persona (classe due) e agli aspetti di ricerca e cultura di dialogo con la chiesa (classe uno).

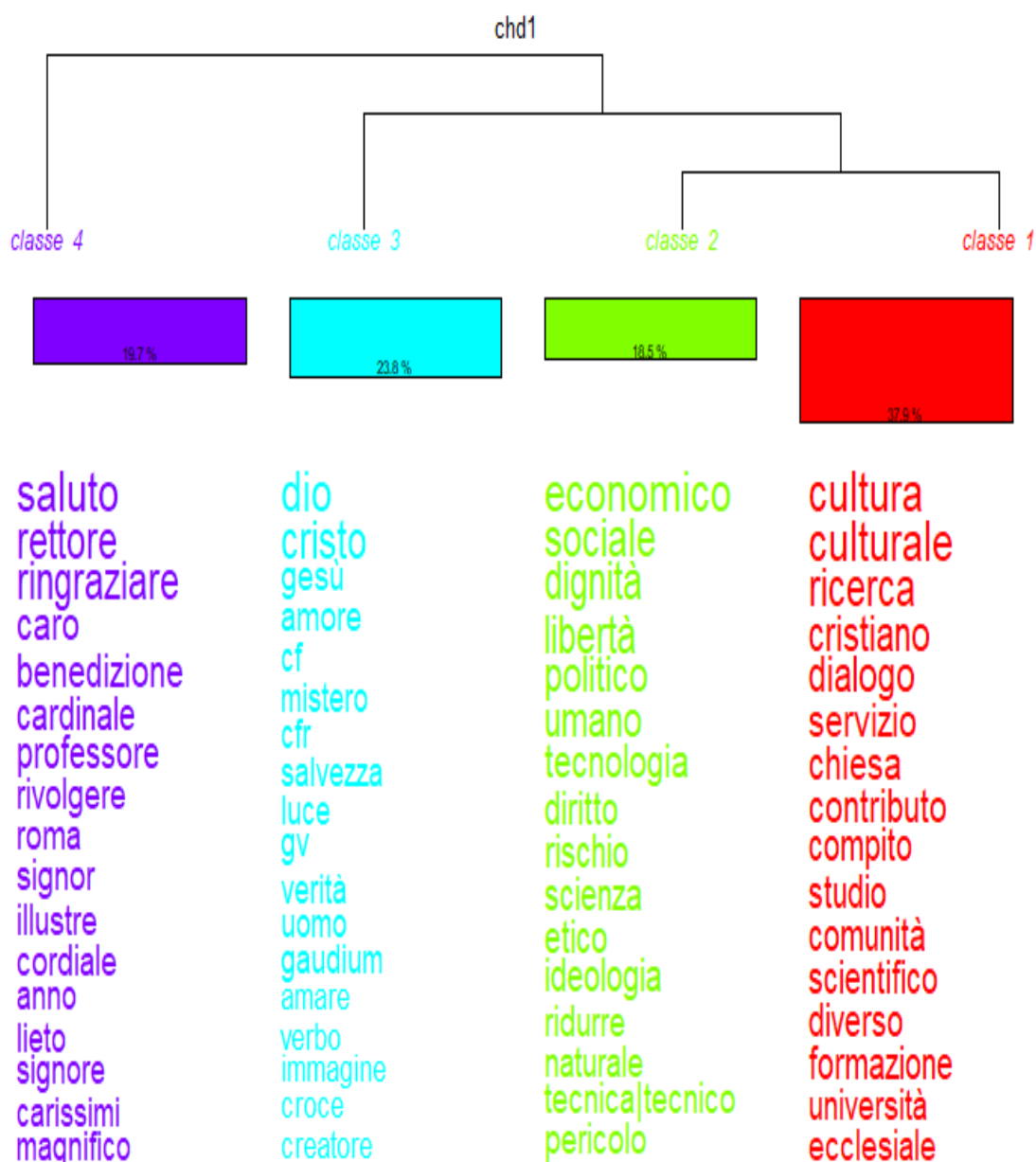


Figura 12. Grafico delle classi e termini legati alla classe semantica.

Dal grafico seguente (Figura 13) si evince come siano rappresentati i temi (classi) dai vari papi. La prima classe semantica, che raccoglie i temi della cultura e della ricerca, è molto presente nei discorsi di Paolo VI; viceversa, tali temi sono poco o per niente rappresentati nei discorsi di Benedetto XVI e di Francesco. La classe semantica caratterizzata dai temi sociali e della persona è maggiormente presente in Francesco. La classe semantica che richiama maggiormente aspetti religiosi è molto presente in Benedetto XVI e Francesco

e, viceversa, sono temi che caratterizzano in misura minore i discorsi di Giovanni Paolo II e Paolo VI. Infine, la classe quattro, relativa alle formalità dei discorsi, è sovra rappresentata da Paolo VI e poco o per niente da Francesco: sembra di poter sostenere, perciò, che il primo abbia preferito comunicare attraverso discorsi formali mentre il secondo, Francesco, con un linguaggio e un approccio maggiormente informali.

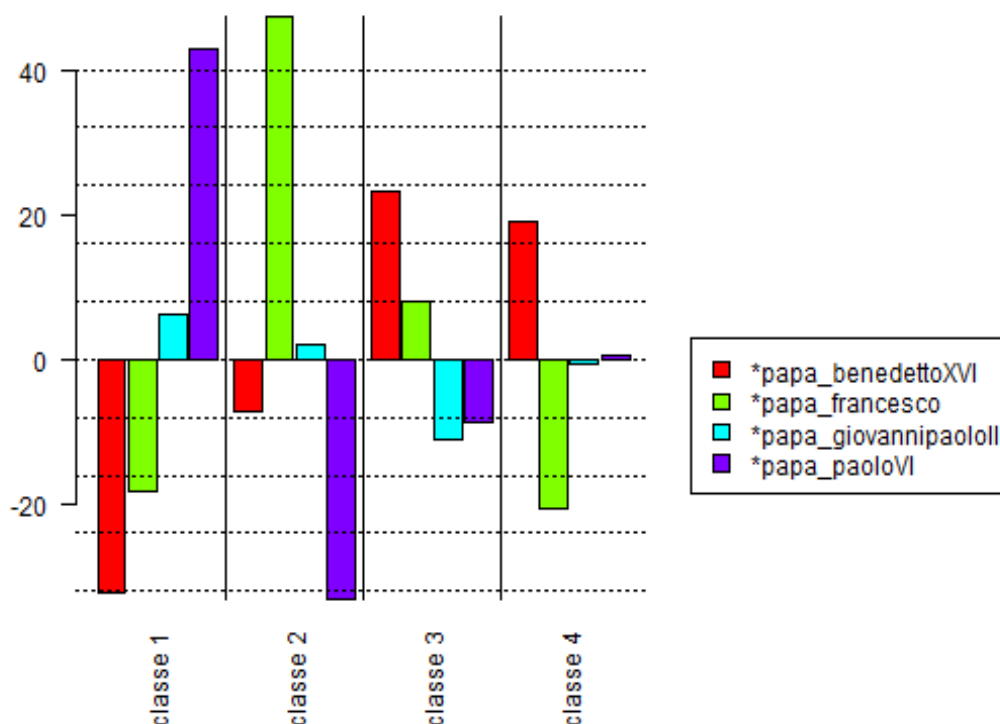


Figura 13. Grafico della distribuzione della variabile papa per ciascuna classe semantiche.

Per capire se le quattro classi identificate dalla procedura di costruzione dei *cluster* semantici si possono scomporre ulteriormente, è stata applicata un'ulteriore analisi ricombinando i parametri in modo da ottenere dei *topics* più specifici, rappresentati nel dendrogramma che segue (Figura 14) con i relativi termini per ogni *topics* (Figura 15). Questi risultati rappresentano la soglia massima di scomposizione dei mondi lessicali del *corpus* oltre la quale il *software* non identifica ulteriori partizioni. I vari temi trovati con i relativi termini danno l'idea della distribuzione delle parole scelte dai papi per parlare di un certo argomento.

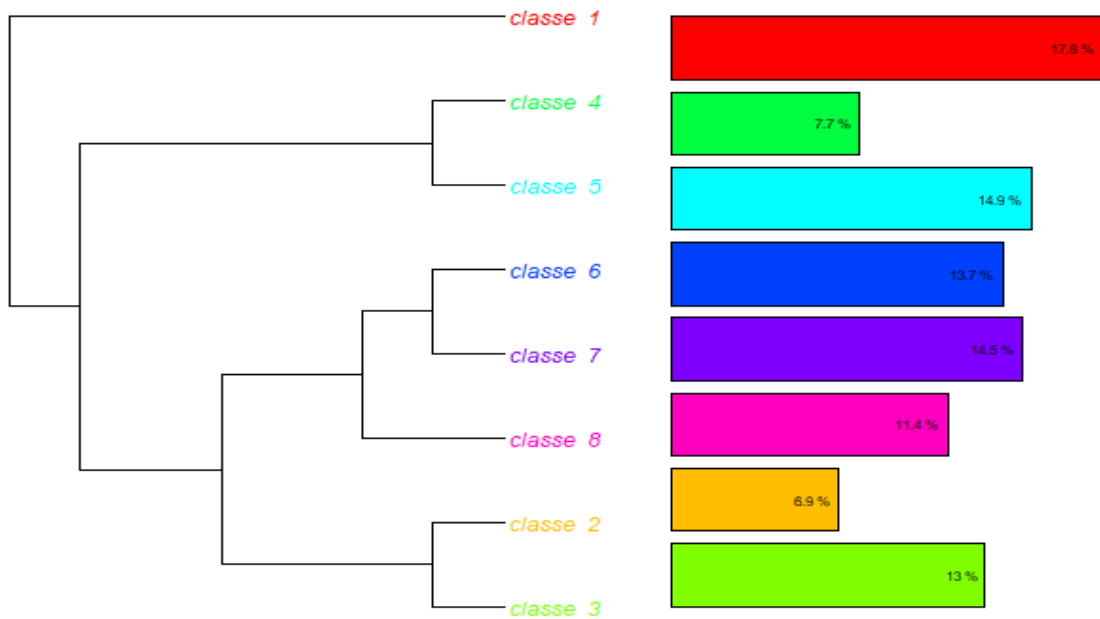


Figura 14. Rappresentazione grafica delle classi semantiche (dendrogramma).

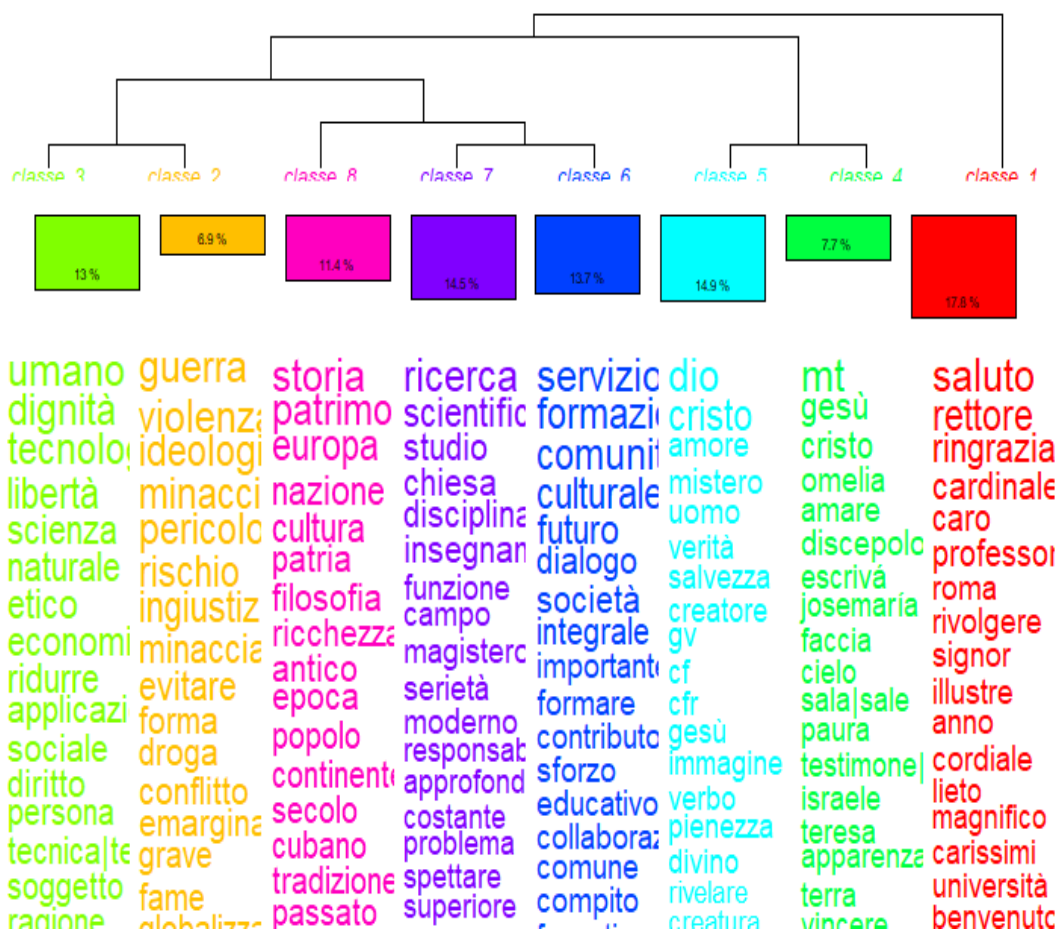


Figura 15. Rappresentazione grafica delle classi semantiche e termini associati.

Questi risultati identificano tre macroaree ben distinte, ognuna delle quali è caratterizzata da specifici *topics* (Figura 16):

1. una prima area è definita dalla classe 1: riguarda i saluti e i ringraziamenti e, quindi, le formalità nei discorsi scritti dai vari papi (rappresentata dai termini come saluto, ringrazia, caro, rettore) e rappresenta il 17,8% delle porzioni di testo.
2. una seconda area religiosa è descritta dai termini delle classi 4 e 5 che in totale rappresentano il 22,6% delle porzioni di testo. Essa è identificata da termini prettamente relativi alla Chiesa come Gesù, Cristo e omelia per la classe 4 e più sul versante filosofico religioso per quel che riguarda i termini della classe 5 (Dio, Cristo, mistero, verità, uomo).
3. una terza area (che si riferisce al 59,2% delle porzioni di testo) è rappresentata dai temi che non hanno a che vedere con la religione e nello specifico da una parte quelli relativi al sociale e la guerra (classe 2 e classe 3) e dall'altra quelli relativi alla storia e al patrimonio culturale, (classe 8), alla ricerca scientifica (classe 7) e alla formazione culturale (classe 6).

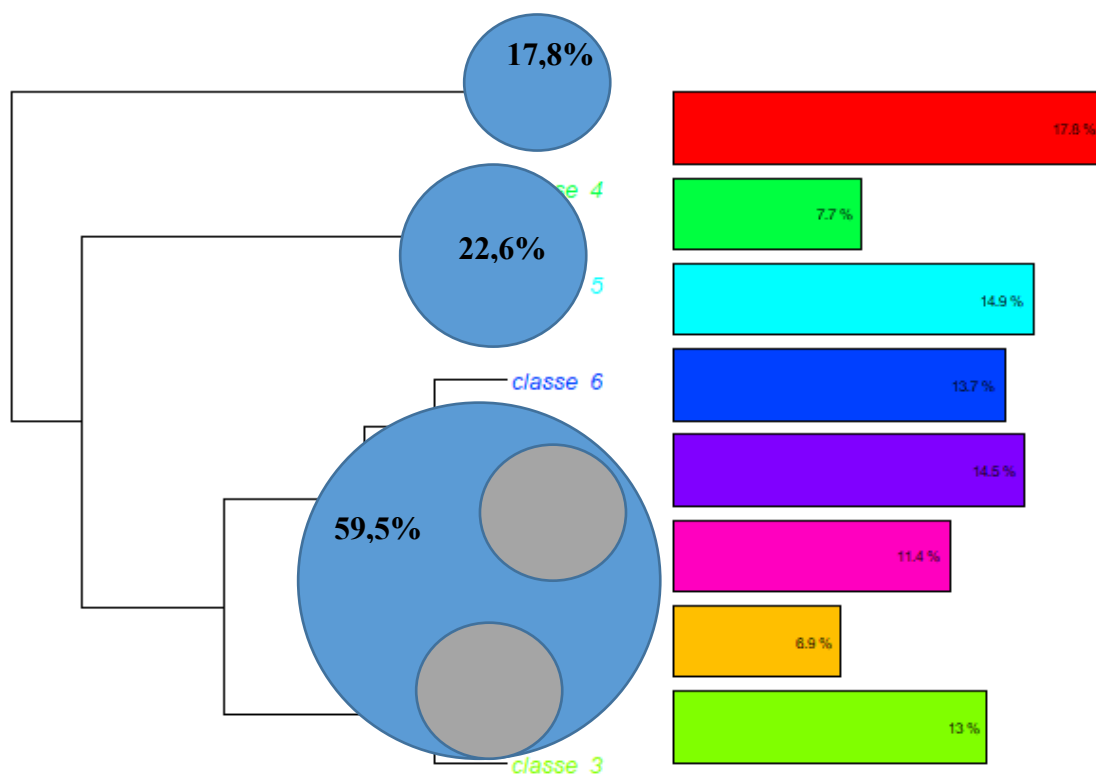


Figura 16. Rappresentazione grafica delle tre macroaree delle classi semantiche.

La terza area, che è la più significativa per quel che riguarda l'oggetto della mia ricerca, aiuta a capire quali sono i temi rilevanti nei testi scritti dai papi e i termini utilizzati per descrivere ogni argomento. Se da una parte sembra scontato che si parli di religione (come per i *topics* relativi alla classe 4 e 5), lo è meno per gli argomenti relativi alla cultura, alla formazione, alla ricerca scientifica e al sociale. Questa terza area è rappresentata dal 59,5% delle porzioni di testo del *corpus* e al suo interno presenta cinque diverse classi. Da una parte la classe numero 2 e la classe numero 3 che rappresentano il 19,9% delle porzioni di testo (semanticamente vicine perché condividono alcune parole) e sono relative ai temi sociali e dall'altra le classi 6,7,8 relative alla formazione (classe 6), alla ricerca scientifica (classe 7) e al patrimonio culturale europeo e nazionale (classe 8) che rappresentano il 39,6% delle porzioni di testo del *corpus*. Dalle percentuali di rappresentazioni delle porzioni di testo appena descritte, posso evincere che i temi legati ai discorsi dei papi, in questa macroarea, hanno un taglio più scientifico-culturale che sociale. Entrando nel dettaglio di ogni mondo lessicale, la classe numero 2, relativa al tema sociale, raggruppa termini quali minaccia, ideologie, violenza e guerra e la classe numero 3, sempre riferita ad argomenti di tipo sociale, è definita da forme quali l'umanità, umano, dignità, libertà e tecnologia. La classe numero 6 raggruppa termini come servizio, formazione, cultura e futuro e rappresenta il 13,7% delle porzioni di testo, è semanticamente vicina alla classe 7 che è definita dai termini quali ricerca, scientifica, studio e chiesa ed è rappresentata dal 14,5% delle porzioni di testo. La classe numero 8, infine, è relativa alla storia e al patrimonio culturale nazionale ed europeo e raggruppa i termini come storia, patrimonio, europea, nazione ed è costituita dall'11,4% delle porzioni di testo. Queste ultime tre classi scientifico-culturali sono egualmente rappresentate in percentuale nel mio *corpus* e sono ben definite da specifici termini anche se condividono delle parole in comune essendo semanticamente e graficamente vicine. Al fine di capire come sono organizzate le tabelle della frequenza delle parole nelle varie classi, riporto quella relativa alla classe semantica numero 3 (Tabella 08). Ogni *cluster*, che si riferisce ad una classe, raggruppa al suo interno tutte le parole rappresentative di quella classe e la relativa percentuale di presenza. La frequenza effettiva della parola "uomo", per esempio, nella classe 3 è di 234 sui 758 totali del *corpus*. La stessa parola sarà poi presente anche in altre classi ma con frequenza minore. In ogni *cluster* le parole non sono ordinate in base alla loro frequenza di presenza nel *corpus*, ma in base

all'associazione di quella parola con quella classe semantica e, nello specifico, in maniera decrescente in base al chi quadrato. Importanti sono poi i valori nell'ultima colonna, il *p-value*, che misura il livello di significatività: più basso è il valore e più alta è la significatività di quella parola in quel mondo lessicale, rispetto agli altri *topics*. Il *p-value* aumenta quindi fino ad arrivare a valori che non sono significativi di quella classe ($p > 0,5$) oppure a valori proprio non significativi (segnalati con la dicitura NS).

posizione	presenza effettiva	presenza totale	chi- quadrato	forma	p-value
0	234	758	234,69	umano	< 0,0001
1	88	186	198,6	dignità	< 0,0001
2	35	40	197,92	tecnologia	< 0,0001
3	107	271	174,12	libertà	< 0,0001
4	192	647	172,26	scienza	< 0,0001
5	54	93	169,47	naturale	< 0,0001
6	64	123	168,74	etico	< 0,0001
7	61	93	163,56	economico	< 0,0001
8	33	44	150,89	ridurre	< 0,0001
9	32	44	140,03	applicazione	< 0,0001

Tabella 08. Frequenza delle parole del *cluster* numero 3.

Il grafico di Figura 17 rappresenta la significatività della parola “umano” rispetto ai vari mondi lessicali. La forma “umano” infatti è molto rappresentata e molto significativa (con valore di *p-value* molto bassi) nella classe 3 che è relativa ai temi legati al sociale e nello specifico all'umanità (uomo, libertà, tecnologia, dignità). È ben raffigurata anche nel *topics* 5 che rappresenta il tema religioso in chiave filosofica. C'è invece poca significatività della parola umano nella classe 1 (relativa alla formalità dei discorsi), nella classe 4 (con termini legati all'argomento religioso) e nella 7 (relativi alla ricerca scientifica). Nelle altre classi ha una significatività media.

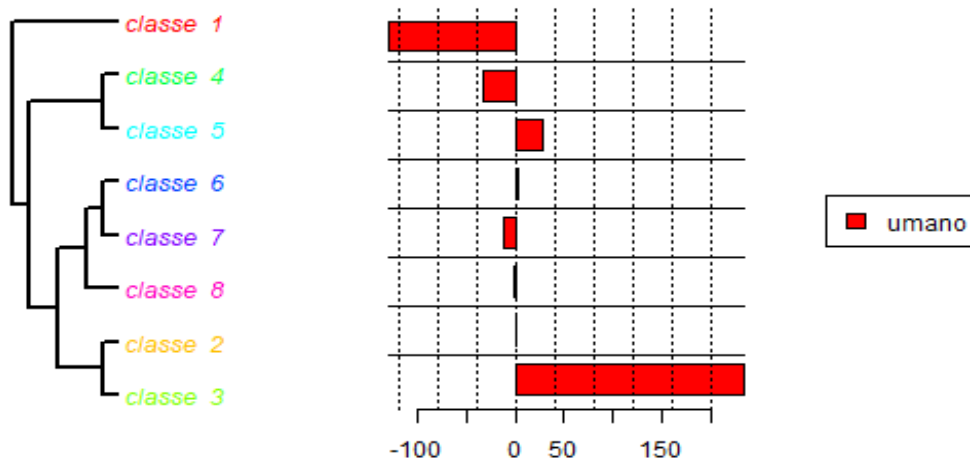


Figura 17. Grafico significatività parola “umano” sulle classi semantiche.

Un'altra importante analisi è quella che risponde alla domanda: “Chi parla di che cosa”, ovvero che permette di comprendere a quali temi e quindi a quale classe è legato ciascun papa. Il tutto è graficamente rappresentato in Figura 18. Nel dettaglio, livelli alti del grafico significano sovra utilizzo delle parole di quella classe, livelli pari a zero corrispondono a un utilizzo medio delle parole di quel mondo lessicale e livelli sotto zero sono segno di un poco delle parole di quella classe. Tutti i papi sono rappresentati in tutte le classi con vari livelli di utilizzo dei termini dei vari *topics*.

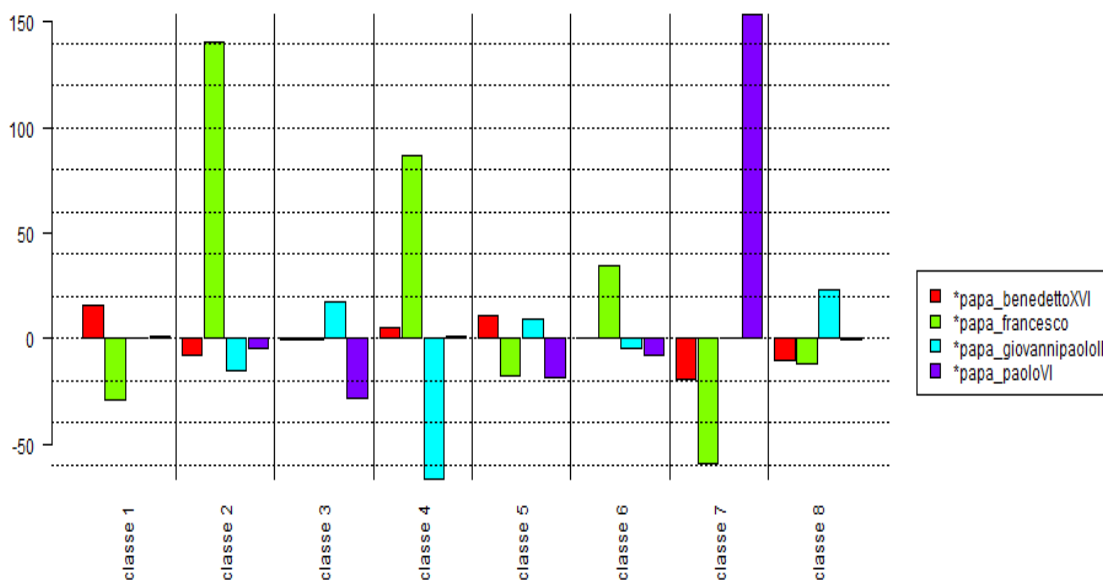


Figura 18. Grafico della distribuzione della variabile papa per ciascuna classe semantiche.

Facendo riferimento al *corpus* della mia ricerca osservo quanto segue:

Papa Francesco (grafico di Figura 19) ha un sovra utilizzo dei termini legati alla classe 2 relativi all'umanità, alla violenza, alla guerra e alla minaccia, a quelli puramente religiosi della classe 4 e a quelli della formazione culturale della classe 6. Per quel che riguarda queste tre classi, si può definire un utilizzo quasi esclusivo rispetto agli altri papi. È però meno legato ai termini della formalità nei discorsi della classe 1, all'aspetto filosofico della religione della classe 5 e poco utilizzati i termini sia dei temi della classe 7 (storico-culturali) che della classe 8 (che riguardano la ricerca scientifica). Ha un livello medio di utilizzo dei termini legati ai *topics* della classe 3 relativa all'umanità e al sociale.

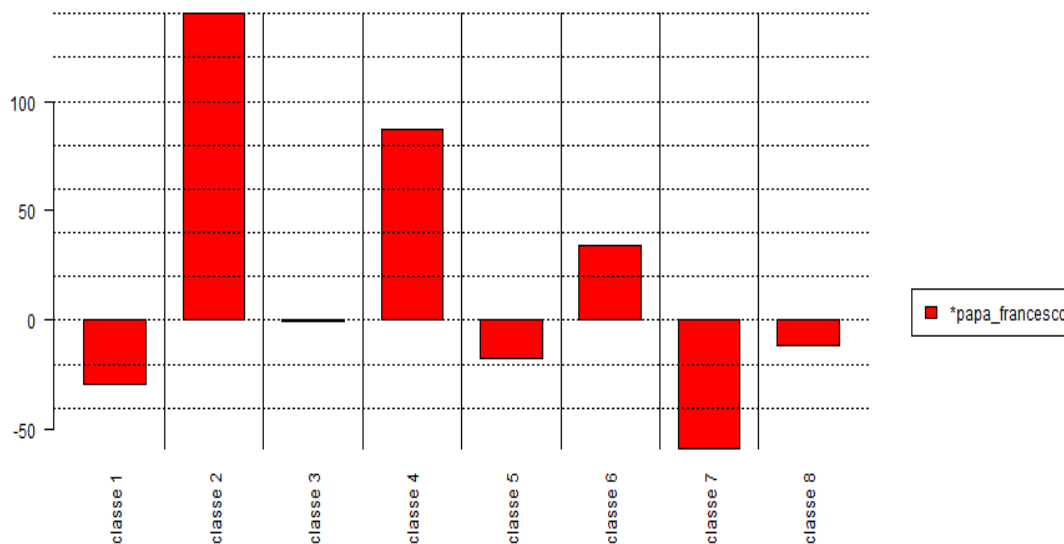


Figura 19. Grafico della distribuzione di Papa Francesco per ciascuna classe semantica.

Papa Benedetto XVI (grafico di Figura 20) ha dei livelli alti di utilizzo dei termini legati ai temi della classe 1, 4 e 5. La classe 1 relativa alle formalità nei discorsi e nei ringraziamenti e le classi 4 e 5 relative agli aspetti religiosi sia dal lato filosofico che quelli puramente legati alla Chiesa. Ha però un utilizzo minore delle parole della classe 2 relativi all'umanità, della classe 7 (storico-culturali) e della classe 8 (che riguardano la ricerca scientifica). Livelli medi di utilizzo dei termini legati alle classi 3 (umanità sociale) e 6 (formazione culturale).

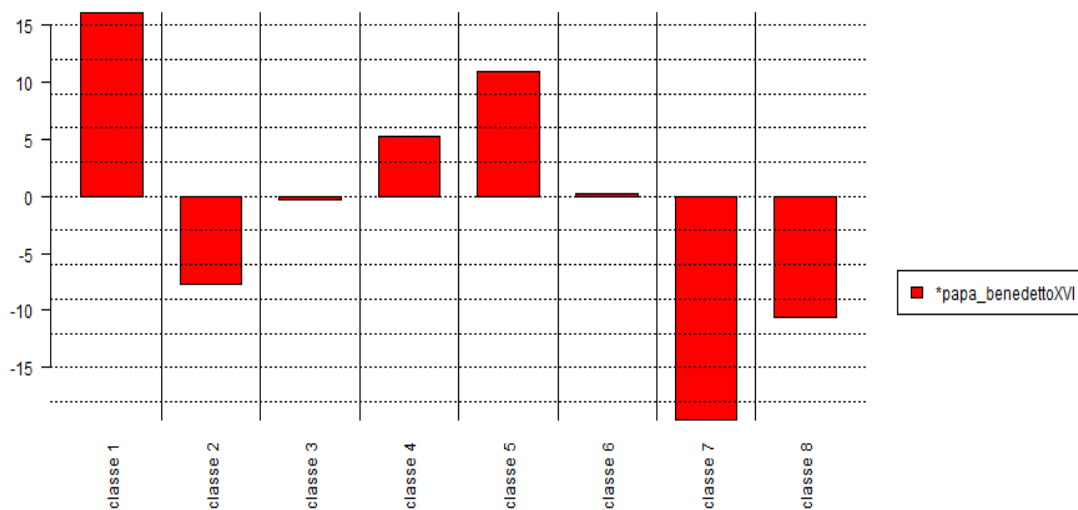


Figura 20. Grafico della distribuzione di papa Benedetto XVI per ciascuna classe semantica.

Come si nota dal grafico di Figura 21 Papa Giovanni Paolo II ha un sovra utilizzo dei termini delle seguenti classi: la classe 3, la classe 5 e la classe 8. La classe 3 che si riferisce agli aspetti sociali dell'uomo, della dignità, della scienza, della libertà, la classe 5 con argomenti filosofico religiosi e la classe 8 riferita alla storia come patrimonio culturale europeo e nazionale.

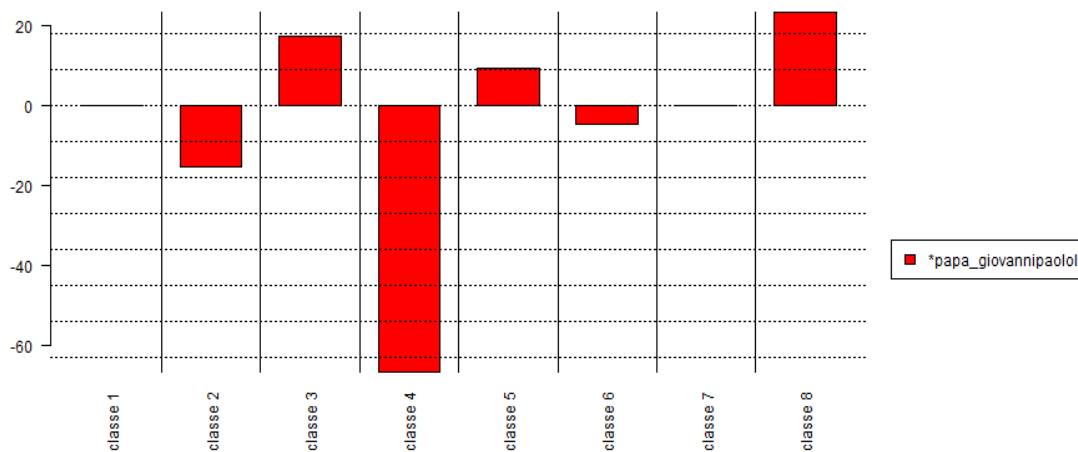


Figura 21. Grafico della distribuzione di papa Giovanni Paolo II per ciascuna classe semantica.

Sono meno marcati, invece, i riferimenti ai termini legati sia alla classe 4, associata alla religione, sia alla classe 2, le cui parole riguardano la guerra, i pericoli e le minacce

dell'umanità. Medio l'utilizzo dei termini collegati alle classi 1 (formalità nella stesura dei discorsi) e 6 (formazione culturale).

Per ultimo, Paolo VI (grafico di Figura 22) è molto legato ai temi relativi alla ricerca scientifica della classe 7 (in maniera esclusiva rispetto agli altri papi) con un sovra utilizzo di questi termini, ma utilizza in maniera minore le parole legate alle classi 2,3,5,6: degli aspetti legati al sociale e i pericoli dell'umanità, degli aspetti filosofici religiosi e a quelli riguardanti la formazione culturale. Un medio utilizzo delle parole legate alle classi 1 (formalità dei discorsi) e 4 (temi religiosi).

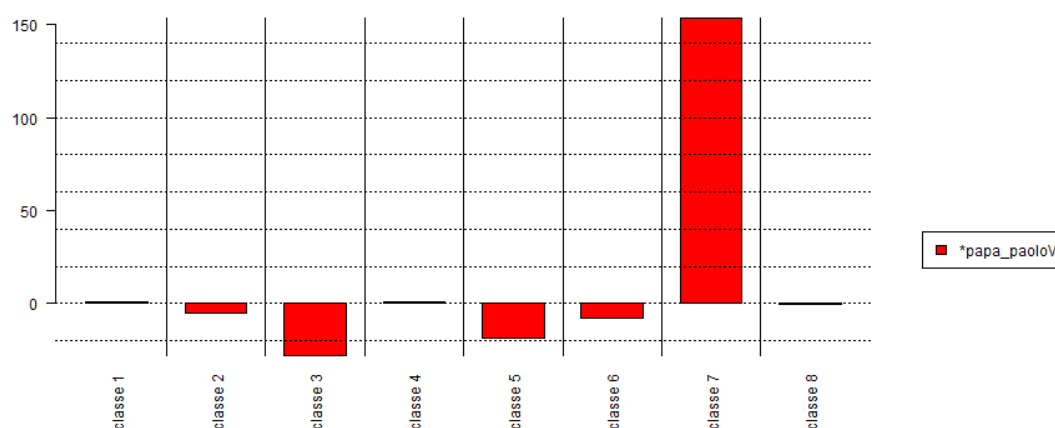


Figura 22. Grafico della distribuzione di papa Paolo VI per ciascuna classe semantica.

Per comprendere a quali temi e quindi a quale classe sono legati gli argomenti trattati nei discorsi nelle uscite del pontefice (domanda, “Dove si parla di che cosa?”) viene in aiuto il grafico di Figura 23 che rappresenta la distribuzione della variabile “argomento” del corpus rispetto ai vari mondi lessicali trovati. Negli incontri pastorali (modalità della variabile “inpastorale”) il papa incontra le persone all’interno del Vaticano, questi incontri sono legati ai temi della formalità (nella stesura dei testi) e agli aspetti religiosi. C’è invece un minor utilizzo dei termini legati alle classi 3,5,8 relative agli aspetti sociali, filosofico-religiosi e riguardanti la ricerca scientifica.

Durante i viaggi apostolici (modalità della variabile “viapostolici”) di più giorni e all’estero, il papa tratta, per la maggiore, i temi legati alla classe 2, le cui parole riguardano la guerra, i pericoli, e le minacce dell’umanità, nonché alla classe 3 con gli aspetti filosofici religiosi e alla classe 8 riferita alla storia come patrimonio culturale europeo e

nazionale. Si nota, invece, un utilizzo meno marcato dei termini delle classi 1 (l'aspetto formale nei testi), 4 (temi religiosi) e 7 (riguardante la ricerca scientifica). Nelle visite pastorali (modalità della variabile "vipastorale"), quando è invitato al di fuori del Vaticano, ma in Italia, sono in sovra utilizzo le parole legate ai mondi lessicali relativi agli aspetti filosofici religiosi e alla ricerca scientifica e poco utilizzati i termini delle classi che riguardano il sociale, i temi a carattere prettamente ecclesiastico e, infine, quelli legati alla formazione culturale.

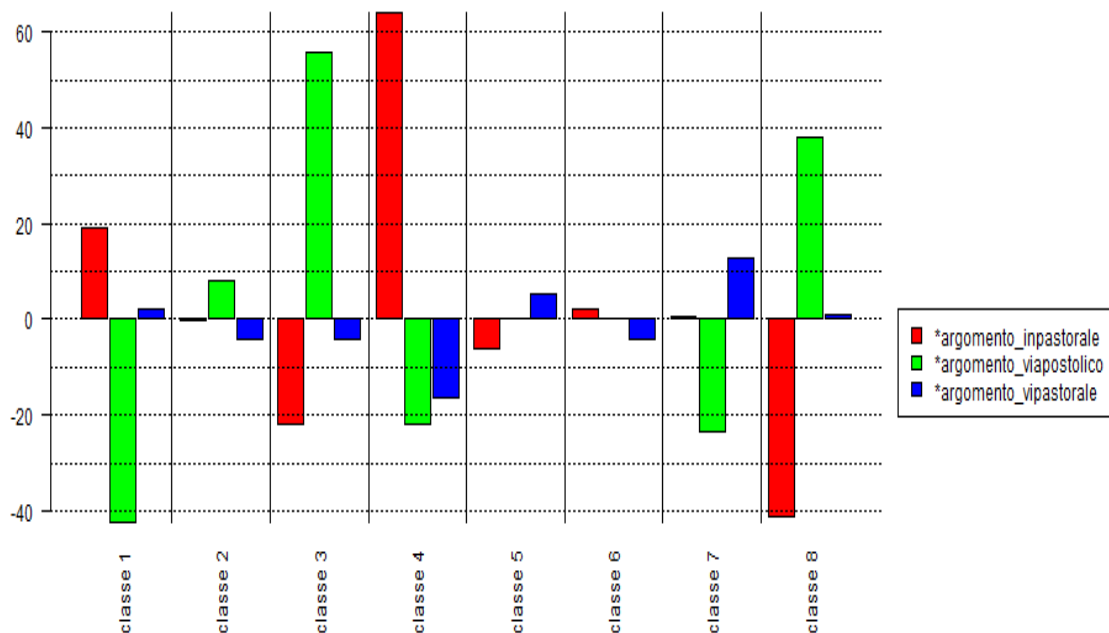


Figura 23. Grafico della distribuzione della variabile argomento per ciascuna classe semantica.

4.1.5.3 Le forme presenti con maggior frequenza nei discorsi

È sicuramente di grande interesse fare anche un confronto delle frequenze delle forme più utilizzate da ciascun papa rispetto agli altri pontefici. I grafici che seguono ne rappresentano affinità e differenze. Chiaramente le forme sono direttamente legate anche ai temi dei vari mondi lessicali spiegate precedentemente.

Rispetto agli altri papi, Paolo VI ha usato con maggior frequenza nei suoi discorsi il termine "laureati" (grafico di Figura 24), infatti è legato ai termini del *topic* della ricerca scientifica e dell'università. All'opposto troviamo, con una minor frequenza di utilizzo di questa forma, Papa Giovanni Paolo II e, leggermente sotto la media, le frequenze di utilizzo della parola laureati nei discorsi degli altri due papi (Ratzinger e Francesco).

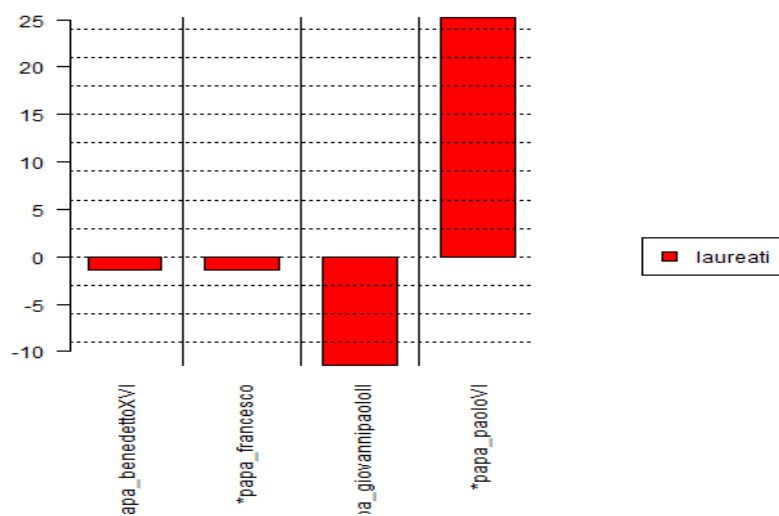


Figura 24. Distribuzione delle frequenze relative alla forma laureati.

Giovanni Paolo II invece ha usato la forma “uomo” con maggior frequenza (grafico di Figura 25) ed ha un utilizzo quasi esclusivo di questa parola nei suoi discorsi rispetto agli altri papi. Il termine “uomo” è anche la forma presente con maggior frequenza nella raccolta di tutti i testi che formano il *corpus*. Marcata invece la frequenza di sottoutilizzo nei discorsi degli altri tre papi.

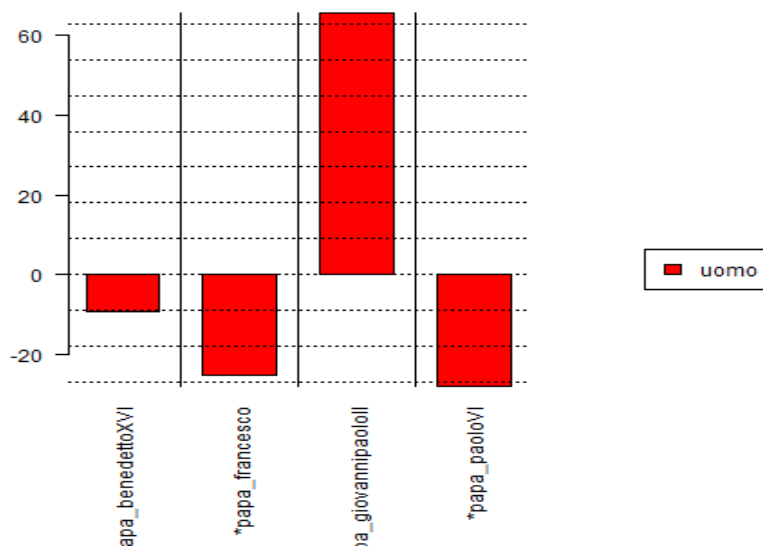


Figura 25. Distribuzione delle frequenze relative alla forma uomo.

La forma “ragione” è il termine utilizzato con una frequenza maggiore nei discorsi di Papa Ratzinger, che collega molte volte la ragione all’umanità e quindi alla fede cattolica. Frequenze di utilizzo sotto la media di questa forma per gli altri tre papi (Figura 26).

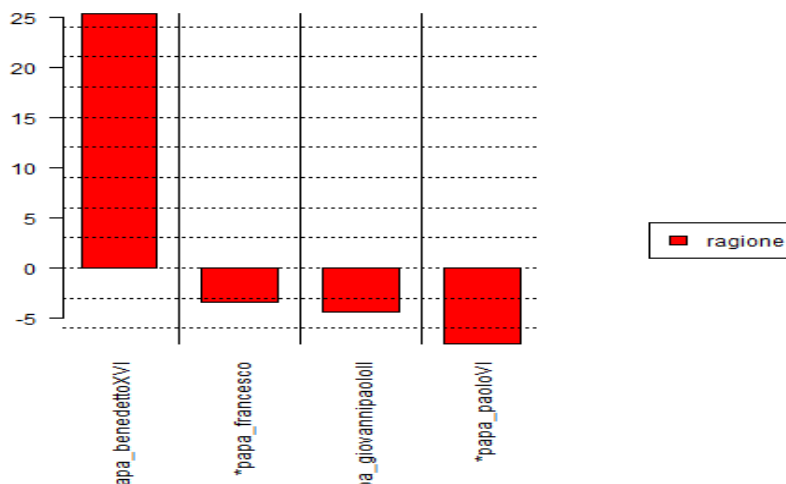


Figura 26. Distribuzione delle frequenze relative alla forma ragione.

Per il pontefice attuale, Francesco, (grafico di Figura 27) la forma che compare con maggiore frequenza nei suoi discorsi rispetto agli altri papi, è la parola “migranti”. Per Giovanni Paolo II lo stesso termine risulta essere, invece, una forma sottoutilizzata nei suoi scritti e di poco sotto alla media l’utilizzo da parte degli altri due pontefici (Paolo VI e Ratzinger).

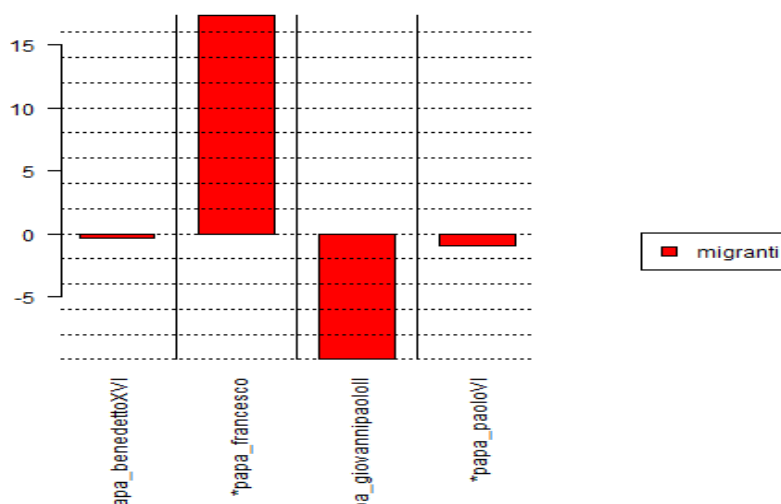


Figura 27. Distribuzione delle frequenze relative alla forma ragione.

I discorsi scritti da Francesco, infatti, si avvicinano molto al mondo lessicale i cui termini sono legati alla guerra, alla minaccia, alle violenze. Per il papa attuale questo tema è un argomento molto caro, soprattutto perché ha degli aspetti personali: i genitori (come descritto nel capitolo 2 della tesi), infatti, hanno vissuto in prima persona le fatiche e le problematicità di emigrare dall'Italia all'America del sud in cerca di lavoro agli inizi del Novecento. I suoi discorsi poi, rispecchiano e sono direttamente collegati alla situazione storica e politica attuale italiana, che non sempre riesce a dare risposte positive alla gestione dei flussi migratori di persone, che, scappando dai luoghi d'origine, vedono nell'Italia un Paese in cui poter migliorare la loro vita.

4.1.5.4 L'analisi delle corrispondenze

Al fine di osservare e analizzare le relazioni tra i *topics*, è stata applicata un'analisi delle corrispondenze: è quindi possibile individuare il posizionamento delle varie classi semantiche sui quattro assi del piano fattoriale (Figura 28).

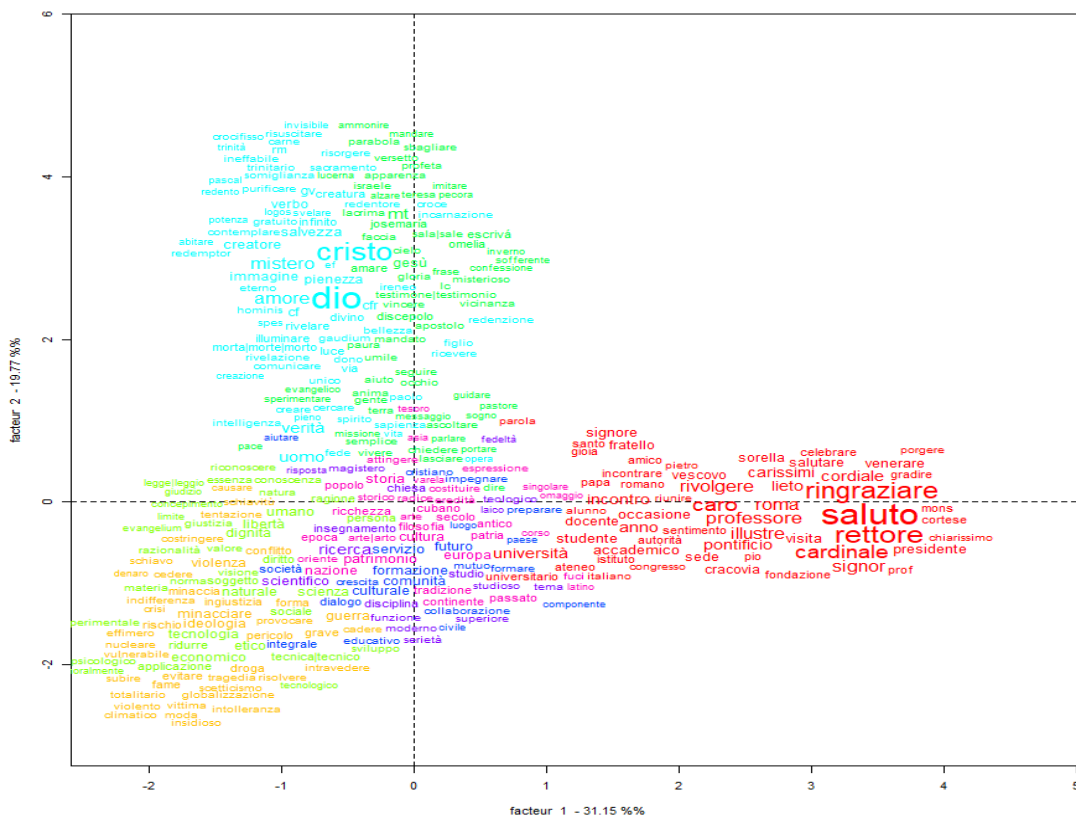


Figura 28. Grafico delle corrispondenze delle classi semantiche.

Nel piano fattoriale tanto più i *topics* saranno simili, dal punto di vista del profilo lessicale, tanto più avranno parole in comune e si posizioneranno vicini nel grafico.

Nello specifico, solo i termini evidenziati in rosso, relativi alla classe 1, assumono una maggiore varianza rispetto agli altri. Sono invece contrapposti nei quadranti ma con una condivisione di molte parole i temi legati agli aspetti religiosi del quadrante in alto a sinistra (divisi in due classi, quella puramente religiosa e quella legata agli aspetti filosofici) rispetto ai temi legati al sociale, all'umanità, agli aspetti relativi alla violenza e al pericolo, posizionati nel quadrante in basso a sinistra. Sempre nel quadrante in basso a sinistra ma spostati più in una posizione centrale i temi legati alla formazione culturale e agli aspetti scientifici. Sono dei mondi lessicali che condividono molte parole la cui frequenza è più alta in alcuni *topics* rispetto che in altri e che vanno a creare una macroarea omogena piuttosto che una serie di tematiche distaccate.

Per analizzare il posizionamento e le corrispondenze dei vari papi è stata, infine, applicata un'altra analisi delle corrispondenze prendendo in esame la variabile "papa" (Figura 29). Dal grafico si evince che Paolo VI si posiziona in maniera distante e isolata rispetto agli altri papi. Egli si colloca, infatti, nel quadrante in basso a destra. Questa zona a lui associata è definita dai temi della ricerca scientifica i cui termini sono molto legati ai suoi discorsi pontificali, dai quali traspare come l'unico papa che tratta di tali argomentazioni. Nella parte opposta rispetto a Paolo VI, troviamo, collocati nel quadrante in basso a sinistra, Papa Francesco e Papa Benedetto XVII. Non sono molto vicini ma trovandosi nello stesso quadrante condividono alcune argomentazioni, in particolar modo i temi legati alla classe semantica 4, i cui termini sono prettamente religiosi. Si distanziano invece per le parole legate al tema delle formalità nei discorsi (Benedetto XVI è molto formale rispetto a Francesco), riguardo il tema della religione connesso all'aspetto filosofico, per i termini legati alla minaccia della guerra e delle ideologie e infine per l'utilizzo delle parole collegate al tema della formazione culturale. L'unico papa a posizionarsi nel quadrante superiore è Giovanni Paolo II, più precisamente in quello di destra, con una distanza da Benedetto XVI che è di ugual misura alla distanza tra questo ultimo e Papa Francesco. Giovanni Paolo II si caratterizza e si distingue dagli altri papi, per i temi legati alla storia come patrimonio culturale italiano ed europeo e anche per quelli legati al sociale (umanità, dignità, scienza, libertà). Condivide, in parte, termini utilizzati anche da Benedetto XVI, legati ad argomentazioni religiose (mistero di Dio,

verità, amore). Giovanni Paolo II, come Benedetto XVI, non mostra dei picchi di sovra utilizzo di termini legati alle varie classi, quindi, essi si collocano in una parte centrale del grafico delle corrispondenze per la variabile “papa” e non agli estremi come Paolo VI e Papa Francesco che invece hanno un sovra utilizzo dei termini dei mondi lessicali a loro associati. Si può parlare di un utilizzo esclusivo per Papa Francesco nella classe 2 (temi sociali legati alla guerra, alla violenza e alla minaccia), e per Paolo VI nella classe 7 (termini legati alla ricerca scientifica).

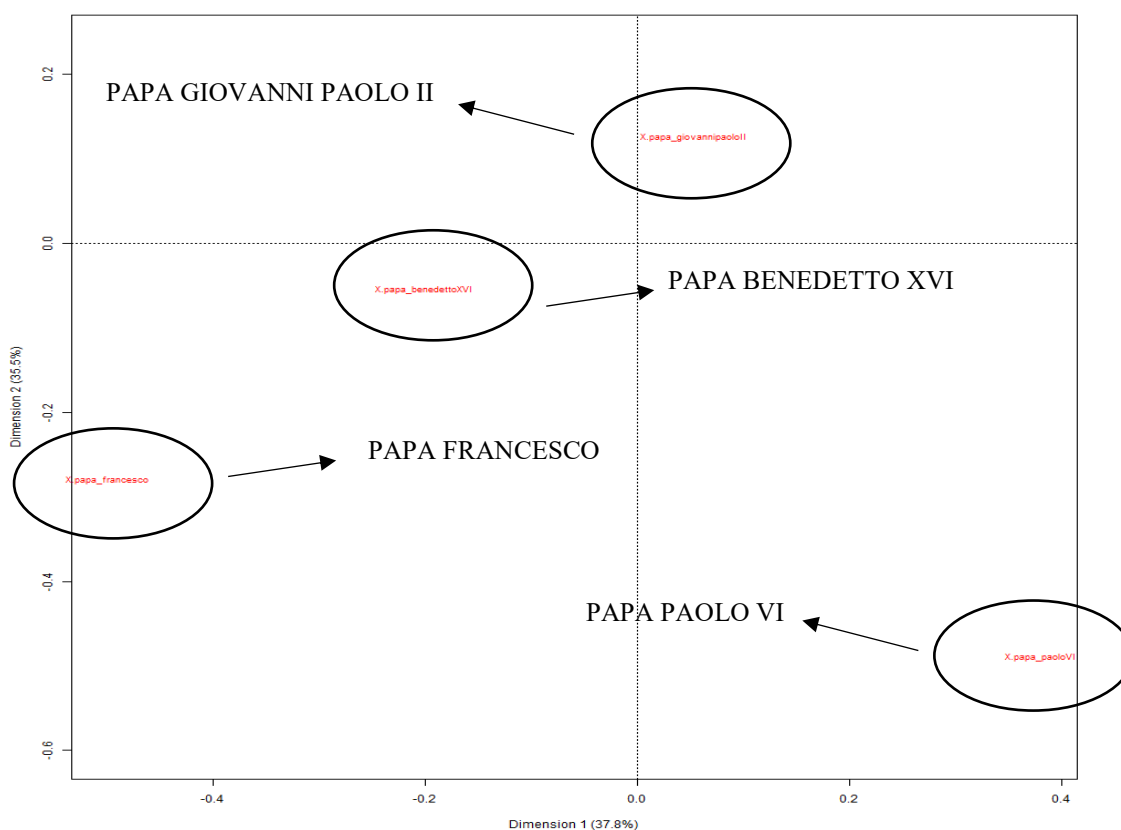


Figura 29. Grafico delle corrispondenze della variabile papa.

4.1.5.5 L’analisi delle similitudini

Per rappresentare i termini e i collegamenti utilizzati nei vari *topics* è stata applicata l’analisi delle similitudini (*cluster network analysis*) rappresentata dal grafico di Figura 30. Nello specifico le parole vengono visualizzate molto vicine quando sono utilizzate nei discorsi in maniera aggregata tra loro e più il collegamento tra i termini è marcato, più alta è la frequenza della rappresentazione grafica delle co-occorrenze. Nel grafico che

segue (Figura 30) non sono rappresentate tutte le forme del *corpus* ma soltanto quelle di frequenza maggiore, per dare una visualizzazione grafica più chiara e intuitiva.

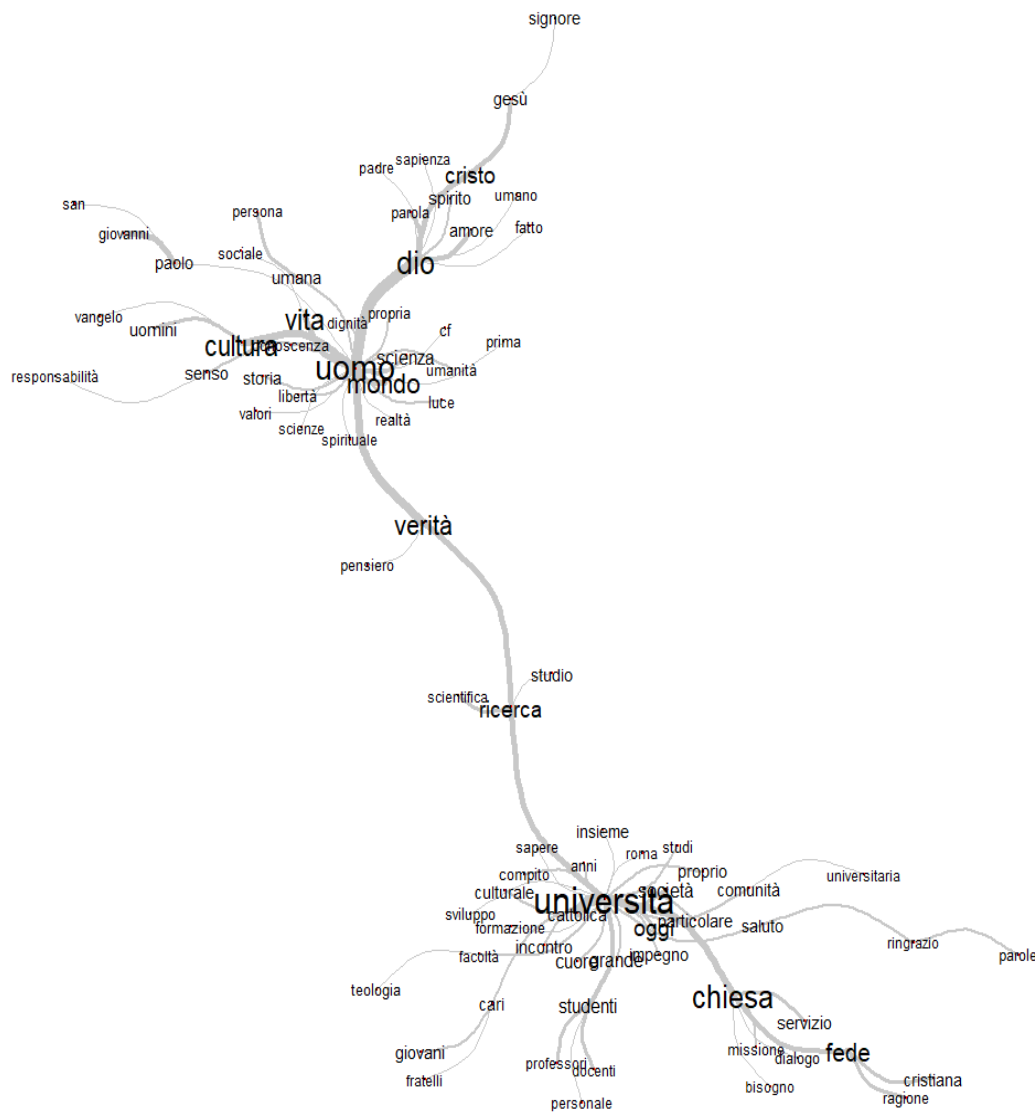


Figura 30. Rappresentazione grafica delle co-occorrenze.

I termini che risultano più connessi e che ricorrono maggiormente rispetto ad altri, nelle stesse porzioni di testo, sono da una parte uomo-vita, uomo-mondo, uomo-cultura e dall'altra università-ricerca, università-formazione, università-oggi, università-studenti. Sono presenti anche le connessioni università-chiesa, chiesa-fede ma anche uomo-Dio e Dio-Cristo. Di particolare importanza, anche per rispondere alla mia domanda di ricerca

iniziale, la coppia di parole verità-ricerca che rappresenta il collegamento principale tra i topics dell'umanità con le argomentazioni legate all'università. Dall'interpretazione del grafico che segue (Figura 31) si ricostruisce anche il ragionamento che ruota intorno ai temi principali dei discorsi, il tema dell'uomo, di Dio, degli aspetti culturali e della vita legati alla verità e alla ricerca. Verità e ricerca sono il canale di connessione con il mondo lessicale dell'università che ha un compito di formazione e di sviluppo nella società ed è legata alla chiesa e alla fede per quel che riguarda i discorsi dei papi riferiti alle università cattoliche.

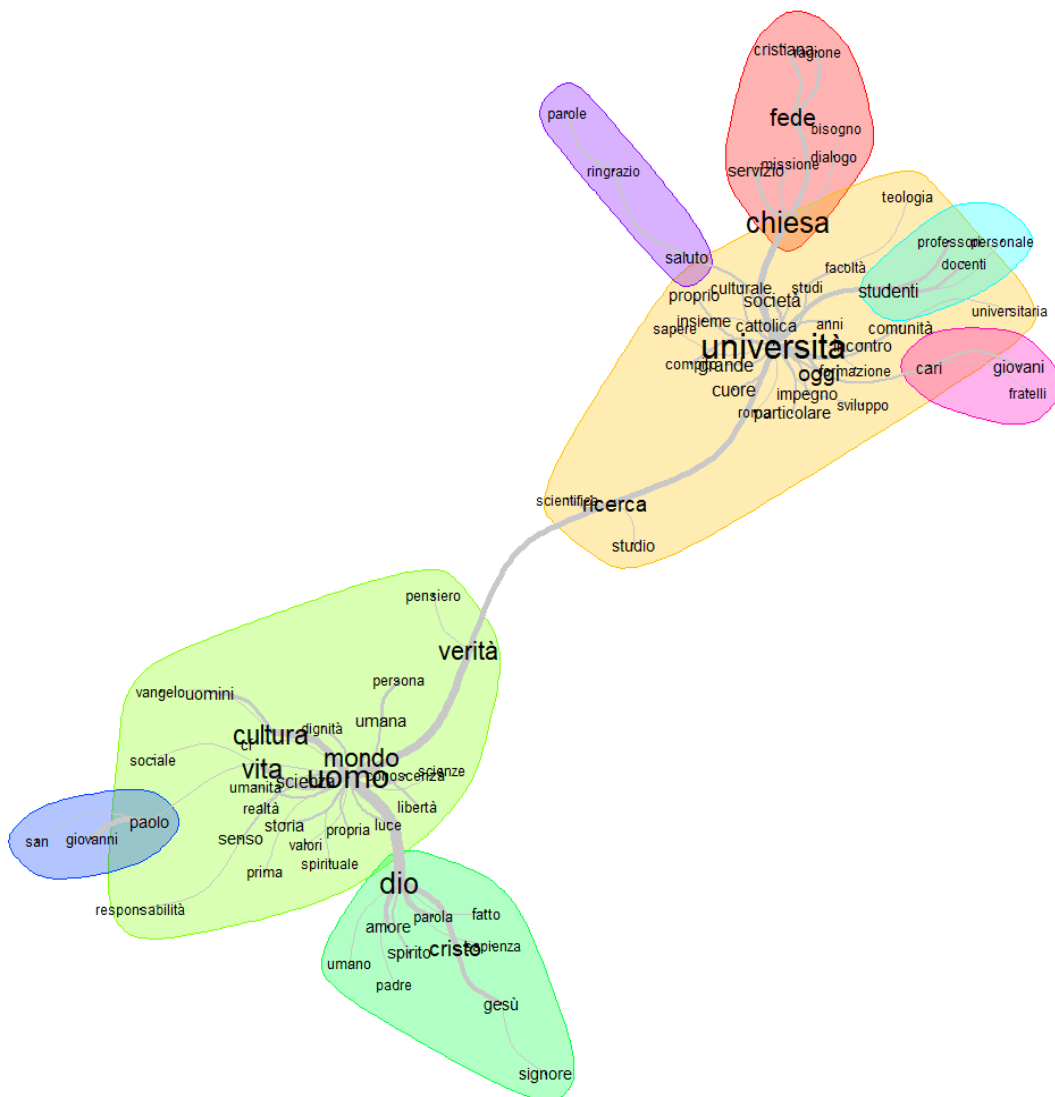


Figura 31. Rappresentazione grafica delle co-occorrenze a colori.

4.1.5.6 La proporzione e l'intensità delle classi semantiche negli anni

La proporzione e l'intensità delle classi semantiche del mio *corpus* sull'asse temporale è ben rappresentata nei grafici di Figura 32 e di Figura 34. Nello specifico nel grafico di Figura 32 sull'asse delle x sono rappresentati tutti gli anni dei discorsi dei vari papi, dal 1963 al 2021, e, sull'asse delle y, compare la percentuale dei vari segmenti di testo (la scala va da 0 a 1 dove 0 significa non rappresentata e 1 rappresentata per il 100%). Ogni anno può essere rappresentato come la somma di tutti i segmenti di testo estratti dai discorsi scritti in quell'anno, possiamo quindi riprodurre la proporzione di ogni classe semantica dalla somma dei segmenti di testo aggregate per classe in ogni anno. La larghezza delle barre è proporzionale al numero di segmenti di testo in un dato anno e la loro altezza rappresenta tutti i segmenti di testi in quell'anno. Si può infatti vedere la distribuzione dei topics (classi) in un dato momento. Dal punto di vista cronologico questa rappresentazione rivela l'andamento degli argomenti negli anni. I vari colori rappresentano le diverse classi semantiche, come calcolate dal *software* Iramuteq e rappresentate nel grafico di Figura 33 e quindi in questo ordine: classe 1 (rosso), classe 4 (verde menta), classe 5 (azzurro), classe 6 (blu), classe 7 viola, classe 8 (fucsia), classe 2 (senape) e classe 3 (verde).

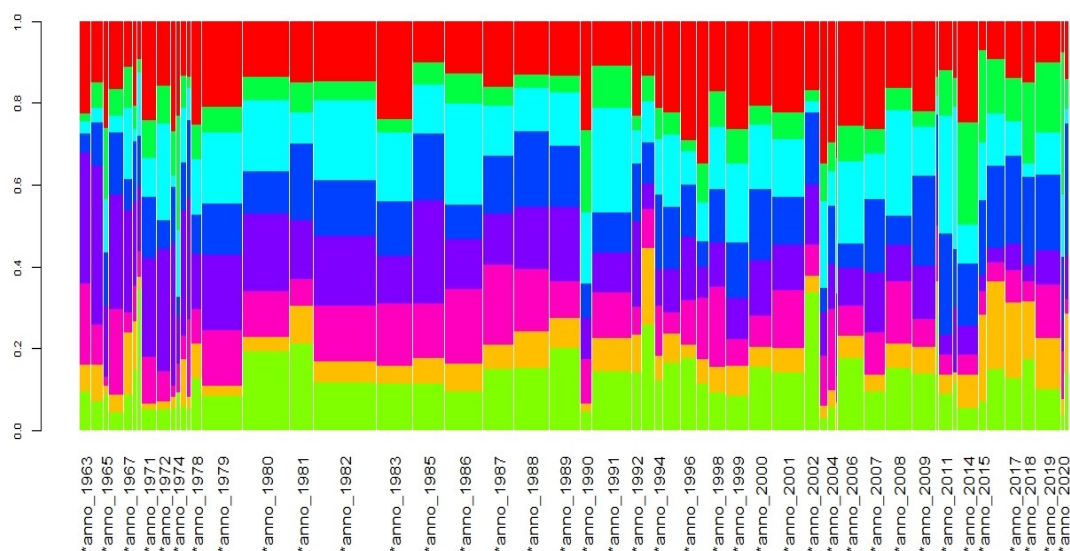


Figura 32. Proporzioni delle classi semantiche per anno.

Dal grafico si evince il papa che in termini di segmenti di testo è stato il più produttivo: le barre che hanno una larghezza maggiore sono quelle che si riferiscono agli anni Ottanta, al tempo del pontefice Giovanni Paolo II. Il meno produttivo invece è stato Paolo VI, le barre che rappresentano i segmenti di testo infatti sono molto strette.



Figura 33. L'ordine delle classi semantiche calcolate.

Nel grafico di Figura 34 sull'asse delle ascisse è sempre rappresentato l'anno, nelle ordinate invece le otto classi semantiche. In questo caso l'altezza delle linee di ogni classe è proporzionale alla dimensione della classe in termini di numero di segmenti di testo che contiene. La larghezza delle celle è proporzionale alla frequenza dei segmenti di testo per un dato anno. L'intensità del colore invece è proporzionale alla forza dell'associazione tra classe semantica e anno. Le celle bianche hanno un *p-value* superiore a 0,05 invece quelle di intensità più forte hanno valori di *p-value* minori a 0,001. Le sfumature di colore variano tra questi due estremi e questo grafico riesce così a figurare quello che si potrebbe chiamare "l'intensità dei temi nel tempo".

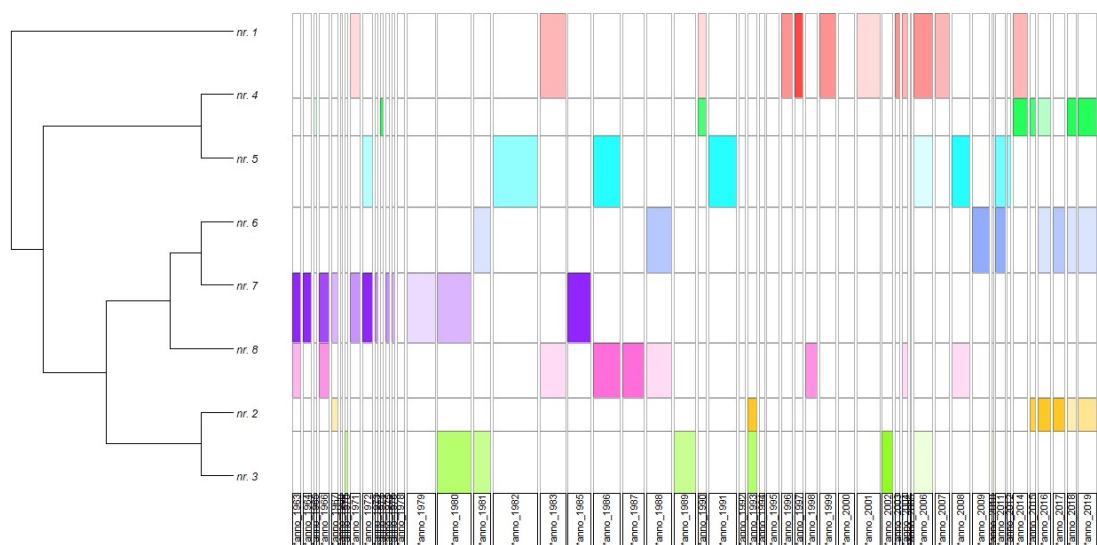


Figura 34. Intensità delle classi semantiche per anno.

Dal grafico posso anche dedurre la rappresentatività delle varie classi nel *corpus*, prendendo in esame la larghezza e la numerosità delle barre per ogni mondo lessicale (come anche già descritto dal dendrogramma ad inizio capitolo di Figura 15). Analizzando l'andamento delle classi semantiche nei vari anni (raffigurate nei grafici di Figura 32 e Figura 34) e ricollegandole ai vari papi sembra di poter dedurre che:

- Paolo VI (papato dal 1963 al 1978) ha sviluppato nei suoi discorsi termini legati per la maggiore ad un'unica classe semantica, quella del mondo lessicale 7: ricerca scientifica, studio, chiesa.
- Giovanni Paolo II (papato dal 1978 al 2005) ha utilizzato forme riferite ai *topics* 3, 5, 8 relative ai termini umanità, filosofia chiesa, storia europea e nazionale, cultura.
- Papa Benedetto XVI (papato dal 2005 al 2013) ha usato termini legati alle classi semantiche 1, 4, 5 senza picchi di sovrautilizzo: formalità, ambito religioso sia lato filosofico che relativo alla fede.

- Papa Francesco (riferimento del suo papato dal 2013 al 2021) palesa forme riferite al mondo lessicale delle classi 2, 4, 6 con sovra utilizzo dei termini legati a questi topics: guerra ideologie, religione fede, formazione culturale.

Il periodo temporale e il contesto storico in cui sono inseriti i vari papi influisce anche sulle tematiche che ricorrono nei loro discorsi.

Paolo VI si discosta dagli altri pontefici, sia come argomenti trattati nei discorsi e sia come unicità di temi: tratta nei suoi discorsi alle Università prettamente temi legati alla ricerca scientifica senza contestualizzare il periodo storico. Storicamente invece Paolo VI non viene ricordato per la sua popolarità nè rispetto ai suoi predecessori (è stato eletto dopo Papa Roncalli, il “Papa buono”, ricordato e amato da molti fedeli), nè rispetto ai successori (che argomenterò più avanti).

Giovanni Paolo II, eletto dopo Paolo VI, ha un approccio del tutto diverso da Paolo VI. Innanzitutto troviamo pluralità di argomentazioni trattate nella stesura dei discorsi, mentre riemerge, in maniera marcata, il tema della religione. Dal 1978, periodo storico caratterizzato da una forte secolarizzazione con il conseguente distacco dalla Chiesa di molte persone, la religione, e soprattutto il tema che riguarda l’aspetto della fede, è l’argomento principale nei discorsi dei vari pontefici. Giovanni Paolo II, a differenza del suo predecessore, dimostra fin da subito una visione più moderna della Chiesa, riuscendo a richiamare folle immense ad ogni suo incontro pubblico. La sua popolarità si deve anche all’utilizzo di un linguaggio che la società civile rimarca spesso come troppo democratico e populista. I discorsi di Giovanni Paolo II sono caratterizzati anche da temi legati alla storia e alla cultura nazionale ma, anche europea, essendo stato il primo pontefice non italiano dal 1523 e avendo effettuato, durante il suo pontificato molti viaggi all’estero. Con lo scopo di diffondere ancora di più la parola del Vangelo, Wojtyła, inaugura nel 1981 la televisione Vaticana e, nel 1995, il sito internet vatican.va rimarcando ancor di più questo suo aspetto popolare, democratico e di trasparenza che la Chiesa deve avere nei confronti dei suoi fedeli.

Nel 2005 viene eletto papa il tedesco Joseph Aloisius Ratzinger, Benedetto XVI. I suoi discorsi sono caratterizzati da un profilo molto formale e preciso nei confronti dei suoi interlocutori, caratteristica che lo contraddistingue rispetto agli altri pontefici presi in esame in questo lavoro. Questo aspetto di grande rigore, è indubbiamente collegato alla

situazione storico-politica in cui è cresciuto Benedetto XVI, il quale nel 1941, per obbligo del periodo, dovette arruolarsi in Germania nella “Gioventù hitleriana”. Ratzinger è un papa che coltiva l’amore per la scrittura (sono noti i suoi periodi di riflessione solitaria e le numerose importanti pubblicazioni) e per alcuni anni è anche docente universitario. Anche se è ben evidente questa sua diversità di essere meno popolare rispetto al suo predecessore, Papa Ratzinger è ricordato come primo pontefice ad utilizzare un suo profilo social su *Twitter*.

L’elezione di Papa Francesco nel 2013, riporta lo stile comunicativo in un’ottica più popolare che lo accomuna allo stile di Giovanni Paolo II. Papa Francesco si discosta molto nella formalità sulla stesura dei discorsi di Ratzinger e tutti ricordano il suo saluto in Vaticano dopo la sua proclamazione: un amichevole “buona sera a tutti”. La partenza dall’Italia all’America dei genitori in cerca di fortuna e il contesto geopolitico attuale influenzano i temi di molti dei suoi discorsi che si accostano in difesa degli emarginati, dei più deboli ma soprattutto degli immigrati. La popolarità di Bergoglio in pubblico non è la stessa che troviamo sui nuovi mezzi di comunicazione social dove non è solito comunicare molto e dove non riesce ad attirare numerosi *follower* come di persona nelle piazze.

CONCLUSIONI

La ricerca che ho sviluppato in questa tesi è stata un'analisi del contenuto dei discorsi dei vari papi che si sono succeduti nella storia della Chiesa, in particolare da Papa Paolo VI, eletto nel 1963, fino a Papa Francesco, eletto nel 2013 ed ancora in carica, focalizzando l'attenzione sui discorsi riferiti all'università e disponibili fino al dicembre 2021. I discorsi sono legati agli incontri del pontefice all'interno e all'esterno del Vaticano e riguardano, talvolta, i viaggi apostolici all'estero. La produzione media dei discorsi dei papi presa in considerazione è stata di almeno due discorsi per anno di papato, alcuni sono stati molto produttivi, altri meno. I discorsi sono sensibilmente diminuiti dal febbraio 2020 con la diffusione della pandemia causata dal virus SARS-Covid 19. Sono infatti state cancellate le riunioni pastorali e le varie uscite del papa dal Vaticano.

Gli strumenti *software* che sono stati utilizzati nel corso di questa ricerca hanno permesso di processare una grande quantità di dati in tempi veloci, fornendo una panoramica sui contenuti, sia di carattere generale, sia di carattere più specifico. L'aspetto interpretativo rimane delle mani del ricercatore. L'unica problematica che ho riscontrato, dovuta all'utilizzo del *software* Iramuteq, è stata nella costruzione del *corpus*: prima di ogni testo andava inserita la stringa identificativa con le variabili e le relative modalità che servivano al *software* per differenziare i vari discorsi. Nel mio caso avevo a disposizione ducentosessantasette testi, se ne avessi avuti migliaia forse sarebbe stato oneroso in termini di tempo perché è un lavoro da una parte sistematico ma dall'altra che deve essere eseguito in maniera molto precisa.

Il percorso di analisi del contenuto di questa ricerca è stato un percorso costituito da vari momenti: dalla selezione all'analisi dei testi fino alla costruzione e all'interpretazione delle categorie concettuali utili per valutare, spiegare e comprendere il fenomeno. I vari momenti, trattati in maniera incrociata e non in modo consecutivo, possono essere ricondotti ai principi fondamentali della sociologia comprendente di Weber. Si può quindi definire il processo di analisi come «una metafora dell'approccio comprendente e quindi di un percorso di ricerca nella sua completezza» (Sbalchiero, 2021).

Entrando nel merito dei contenuti, i risultati ottenuti dalle analisi sul *corpus* hanno fatto emergere delle differenze di argomenti trattati nei discorsi scritti dai vari papi sia sull'asse temporale, sia legate al contesto geopolitico.

Paolo VI ha focalizzato i suoi scritti su un solo argomento utilizzando termini esclusivamente legati alla classe semantica relativa alla ricerca scientifica. Egli non si è contraddistinto per una particolare caratteristica rispetto agli altri pontefici e, sul piano storico, non viene ricordato come un pontefice popolare, anche perché il suo papato si ritrova dopo papa Roncalli, molto apprezzato dal popolo e prima dell'amato Giovanni Paolo II. Neanche la poca produttività dei discorsi scritti negli anni di papato lo ha aiutato a definire un vero e proprio "idealtipo". Si può quindi concludere che Paolo VI non vada accumulato con nessun altro pontefice preso in considerazione.

I suoi successori dimostrano, invece, un'apertura e un cambiamento di rotta rispetto a questo approccio. Se il tema a carattere religioso, con riferimenti alla Chiesa a Dio e alla fede, è stato il filo conduttore degli scritti dei vari pontefici, temporalmente, il cambiamento è avvenuto alla fine degli anni Ottanta, con l'elezione di Giovanni Paolo II: in un periodo storico caratterizzato da una forte secolarizzazione riguardo ai temi religiosi (distacco delle persone dal credo religioso) e da una profonda crisi di identità della comunità da un lato e una crescita dei matrimoni civili dall'altro. Non a caso proprio con Giovanni Paolo II viene inaugurata nel 1981 la televisione Vaticana, allo scopo di una maggiore e più efficace diffusione della parola del Vangelo e, nel Natale 1995, sempre sotto il suo pontificato, viene creato e pubblicato il sito internet ufficiale del Vaticano raggiungibile ancor oggi alla pagina www.vatican.va. Giovanni Paolo II ha seguito argomentazioni legate alla cultura nazionale ed europea (grazie anche ai suoi molti viaggi) e al contesto sociale. Egli risulta essere stato un papa molto amato, in grado di radunare folle di persone nei suoi incontri pubblici, anche per il suo linguaggio semplice e poco formale, in grado di arrivare al cuore della gente comune.

Papa Benedetto XVI interrompe, nei suoi anni di papato, la linea comunicativa semplice e poco formale intrapresa da Giovanni Paolo II. I suoi discorsi, infatti, hanno un approccio più formale e la sua figura si discosta dalla semplicità del suo predecessore. La sua caratteristica è di essere un papa molto rigoroso e riservato che non lo paragona agli altri pontefici.

La linea comunicativa semplice e poco formale di Giovanni Paolo II viene poi ripresa e accentuata dall'attuale Vescovo di Roma. Papa Francesco, infatti, risulta essere molto legato ai *topics* sociali e spesso nei suoi discorsi tratta temi riferiti ai più deboli, agli

emarginati e agli immigrati. Non si accomuna a papa Ratzinger nell'aspetto formale e utilizza un linguaggio molto semplice e diretto.

Analizzando il grafico delle corrispondenze dei papi si evince quanto segue:

Paolo VI si discosta da tutti, Giovanni Paolo II si accosta a Benedetto XVI per la condivisione di termini legati all'aspetto filosofico della religione (Dio, mistero, uomo, verità) e Benedetto XVI si accomuna a Francesco per i temi trattati nei suoi discorsi prettamente religiosi (Gesù, Cristo, omelia).

Il papa legato alla formalità è, quindi, Benedetto XVI, mentre il meno formale per quel che concerne la stesura degli scritti è Papa Francesco (a lui si deve una formalità legata all'utilizzo di forme nei discorsi, quali "caro rettore, ringrazio, saluti, carissimi studenti"). Proprio questo aspetto relativo alla poca formalità di Papa Francesco deve far riflettere se il suo stile comunicativo è di fatto voluto, cercando di capire e spiegare a che tipo di pubblico vuole rivolgersi: Papa teologo o Papa pastore tra la gente?

Di particolare importanza è il risultato della frequenza dei termini utilizzati nei discorsi di un papa rispetto agli altri. I discorsi che ho scelto per questa mia ricerca sono riferiti alle università, ma la forma utilizzata con frequenza maggiore è "uomo". Questo significa che i pontefici mettono al centro del loro metodo comunicativo la persona, l'uomo e il suo ruolo fondamentale dal punto di vista della religione, ma, soprattutto, nel suo percorso di vita, e in relazione alla verità, alla scienza e alla formazione culturale. L'università riveste, invece, un suo ruolo fondamentale nella ricerca scientifica e nello studio mantenendo uno stretto legame con la Chiesa e la fede, sin particolare, per quel che riguarda i discorsi riferiti alle Università Cattoliche. Il papa più profondo e pacato per quanto concerne i discorsi, è Benedetto XVI, il quale ha usato con maggiore frequenza il termine ragione rispetto agli altri pontefici e non ha picchi di sovra utilizzo di termini in nessuna classe semantica analizzata. Lo stile comunicativo emerso dai discorsi che i papi hanno rivolto alle Università, dopo l'elezione di Giovanni Paolo II, ha, quindi, l'obiettivo di riavvicinare le persone alla Chiesa, con l'utilizzo di un linguaggio meno formale. Tale aspetto si è interrotto durante il papato di Ratzinger e poi è stato ripreso da Bergoglio. Si può quindi parlare di una sorta di deistituzionalizzazione della Chiesa iniziata con Giovanni Paolo II e portata avanti da Papa Francesco, con piazze piene di gente durante gli incontri pubblici del pontefice e meno fedeli, invece, nei luoghi di culto durante le celebrazioni dei riti religiosi.

Gli sviluppi di questa mia indagine potrebbero proseguire focalizzando l'attenzione su come potrebbero articolarsi, nel prossimo futuro, le posizioni dei papi: in linea con la situazione geopolitica, se si porterà avanti lo stile comunicativo di Giovanni Paolo II e Francesco oppure con un cambio di rotta se prevarrà un approccio simile a quello formale di papa Ratzinger.

Un altro sviluppo di ricerca potrebbe essere legato all'analisi del linguaggio utilizzato dai vari pontefici, per capire quale sia il più gradito alle persone. Sarebbe, inoltre, interessante verificare quali siano gli argomenti del pontefice attuale messi in evidenza dai media (tradizionali e non) e quali, invece, evidenziati dai papi precedenti.

FONTI

Bibliografia

AA.VV. (2014) *Enciclopedia dei papi*. Roma: Istituto della Enciclopedia italiana (contenuto disponibile online sul sito Treccani https://www.treccani.it/enciclopedia/elenco-opere/Enciclopedia_dei_Papi).

AMATURO E. (1993), *Messaggio simbolo comunicazione. Introduzione all'analisi del contenuto*, Carocci Editore.

BATINI C., SCANNAPIECO M, (2006). *Data Quality: Concepts, Methodologies And Techniques*, Springer DCSA.

BOIVIN, M. (1997). *La pédagogie prospective: nouveau paradigme*. Sainte-Foy, Presses de l'Université du Québec.

BOUDON R. (1980), *La logique du social. Introduction à l'analyse sociologique*, Hachette littérature.

BRYMAN A. (2016), *Social Research Methods*, Oxford University Press, Oxford.

CANNAVO L., FURDA L. (2007), *Ricerca sociale. Tecniche speciali di rilevazione, trattamento e analisi*, Carocci Editore.

CARDANO M. (2011), *La ricerca qualitativa*. il Mulino.

CARDANO, M (2011), *Ricerche sociali. Un'introduzione alla metodologia delle scienze sociali*, Carocci Editore.

- CHIOVARO F. BESSIERE G. (1996), *Urbi et orbi. I papi nella storia*, Electa Gallimard.
- CORBETTA P. (2015), *La ricerca sociale: metodologia e tecniche. Volume I - I paradigmi di riferimento*, Il Mulino.
- CORBETTA P. (2015), *La ricerca sociale: metodologia e tecniche. Volume II – Le tecniche quantitative*, Il Mulino.
- CORSI V. (2009), *La sociologia tra conoscenza e ricerca*, Franco Angeli Edizioni
- DE CONTI (2016), *Il concetto di paradigma*, Educare.it (rivista on line - ISSN: 2039-943X) - Vol. 16, n. 8 – agosto 2016.
- DE ROSE C. (2003), *Che cos'è la ricerca sociale*, Carrocci editore.
- DE ROSE C. (2018), *L'indagine campionaria e il sondaggio d'opinione. Metodi quantitativi della ricerca sociale*, Carrocci editore.
- FAGGIOLANI C. (2011), *L'identità percepita: applicare la Grounded Theory in biblioteca*, EUM-Edizioni Università di Macerata.
- HEMPEL C.G. (1952), *Fundamentals of concept formation in empirical science*.
- KENNETH D. BAILEY (2006), *Metodi della ricerca sociale. Vol. I. I principi fondamentali*, Il Mulino.
- LABOA J.M. (2007), *La storia dei papi. Tra il regno di Dio e le passioni terrene*, Jaka Book.
- LASSWELL H. D. (1927), *Propaganda Techique in the World War*, The MIT Press, 1971.

McENERGY T., WILSON A., (2001), *Corpus Linguistics*, Edinburgh University Press.

O'MALLEY J. (2011), *Storia dei papi*, Fazi.

PACE E., (2021), *Introduzione alla sociologia delle religioni*, nuova edizione, Carocci, Roma.

PANDOLFINI V. (2017), *Il sociologo e l'algoritmo. L'analisi dei dati testuali al tempo di Internet*, Franco Angeli, Milano.

RANKE (VON) L. (2013), *Storia dei papi*, Ghibli.

RICOLFI. L. (1997), *La ricerca qualitativa*, Carocci editore.

GILI G., (2001), *Il Problema della manipolazione: peccato originale dei media?* Franco Angeli, Milano.

REINERT M. (1998), *Mondes lexicaux et topoi dans l'approche Alceste. In S. Mellet, M. Vuillaume (eds.), Mots chiffrés et déchiffrés, Honoré Champion, Editeur, Paris.*

SBALCHIERO S., (2018), *Finding topics: a statistical model and a quali-quantitative method*, in Arjuna Tuzzi, *Tracing the Life-Course of Ideas in the Humanities and Social Sciences*, Springer, pp. 189-210.

SIMEONI M. (2020), *La Modernità di papa Francesco*, EDB.

SCHNAPPER D. (2005), *Comprensione e analisi tipologica*, in Borlandi M., Sciolla L., (eds), *La spiegazione sociologica. Metodi, tendenze, problemi*, Il Mulino, Bologna, pp. 169-184.

STATERA G (1984), *Contro la “nouvelle vague” antimetodologica», Sociologia e ricerca sociale*, Nel farsi della ricerca, numero speciale di Sociologia e ricerca sociale, 2004, XXV.

TUZZI A. (2003), *L'analisi del contenuto. Introduzione ai metodi e alle tecniche di ricerca*, Carocci editore, Roma.

Risorse web

<https://www.vatican.va> (consultato l'ultima volta il 22 settembre 2021).

<https://www.treccani.it/> (consultato l'ultima volta il 23 settembre 2021).

https://www.treccani.it/enciclopedia/elenco-opere/Enciclopedia_dei_Papi (consultato l'ultima volta il 25 settembre 2021).

https://www.repubblica.it/vaticano/2021/01/09/news/il_papa_inedito_il_silenzio_e_la_lingua_di_dio_-281865063/ (consultato l'ultima volta il 21 settembre 2021).

<https://www.vaticannews.va/it/papa/news/2021-09/papa-francesco-teologo-pastore-presentazione-libro-lev.html> (consultato ultima volta il 21 settembre 2021).

<https://www.comunicazione.va> (consultato l'ultima volta il 20 settembre 2021).

<https://www.comunicazione.va/it/chi-siamo/statuto.html> (consultato l'ultima volta il 22 settembre 2021).

<https://www.interris.it/copertina/gramaglia-papa-comunicatore/> (consultato l'ultima volta il 21 settembre 2021).

<https://www.lacomunicazione.it/voce/content-analysis/> (consultato l'ultima volta il 18 settembre 2021).

https://www.researchgate.net/publication/305566851_Messaggi_dal_Colle_I_discorsi_d_i_fine_anno_dei_presidenti_della_Repubblica (consultato l'ultima volta il 17 settembre 2021).

<https://francescacilento.com/analisi-del-contenuto/> (consultato l'ultima volta il 30 novembre 2021).

<http://www.iramuteq.org/> (consultato l'ultima volta il 20 gennaio 2022).

<https://www.youtube.com/watch?v=0S2FSXgcx-M> How to explore textual corpora with the Iramuteq *Software*? (consultato l'ultima volta il 10 marzo 2022).

<http://www.giat.org/?lang=it> (consultato l'ultima volta il 10 marzo 2022).

APPENDICE

I nomi di tutti i papi eletti (fonte <https://www.vatican.va/content/vatican/it.html>)

N°	NOME	INIZIO PONTIFICATO	FINE PONTIFICATO	NOME SECOLARE	PAESE DI NASCITA
1	Pietro		67		Bethsaida in Galilea
2	Lino	68	79		Tuscia
3	Anacleto o Cleto	80	92		Romano
4	Clemente	92	99		Romano
5	Evaristo	99	108		Greco
6	Alessandro I	109	116 o 119		Romano
7	Sisto I	117 o 119	126 o 128		Romano
8	Telesforo	127 o 128	137 o 138		Greco
9	Igino	138	142 o 149		Greco
10	Pio I	142 o 146	157 o 161		Aquileia
11	Aniceto	150 o 157	153 o 168		Emesa (Siria)
12	Sotero	162 o 168	170 o 177		Fondi
13	Eleuterio	171 o 177	185 o 193		Nicopoli (Epiro)
14	Vittore I	186 o 189	197 o 201		Africano
15	Zefirino	198	217 o 218		Romano
16	Callisto I	218	222		
17	Urbano I	222	230		Romano
18	Ponziano	21.VII.230	28.IX.235		Romano
19	Antero	21.XI.235	3.I.236		Greco
20	Fabiano	236	20.I.250		Romano
21	Cornelio	13.III.251	... VI.253		Romano

22	Lucio I	5.VII.253	5.III.254		Romano
23	Stefano I	12.III.254	2.VIII.257		Romano
24	Sisto II	30.VIII.257	6.VIII.258		Greco
25	Dionisio	22.VII.259	26.XII.268		patria ignota
26	Felice I	5.I.269	30.XII.274		Romano
27	Eutichiano	4.I.275	7.XII.283		Luni
28	Caio	17.XII.283	22.IV.296		Dalmata
29	Marcellino	30.VI.296	25.X.304		Romano
30	Marcello I	306	16.I.309		Romano
31	Eusebio	18.IV.309	17.VIII.309		Greco
32	Milziade o Melchiade	2.VII.311	10.I.314		Africano
33	Silvestro I	31.I.314	31.XII.335		Romano
34	Marco	18.I.336	7.X.336		Romano
35	Giulio I	6.II.337	12.IV.352		Romano
36	Liberio	17.V.352	24.IX.366		Romano
37	Damaso I	1.X.366	11.XII.384		Romano
38	Siricio	29.XII.384	26.XI.399		Romano
39	Anastasio I	27.XI.399	19.XII.401		Romano
40	Innocenzo I	22.XII.401	12.III.417		Albano
41	Zosimo	18.III.417	26.XII.418		Greco
42	Bonifacio I	29.XII.418	4.IX.422		Romano
43	Celestino I	10.IX.422	27.VII.432		Campania
44	Sisto III	31.VII.432	19.VIII.440		Romano
45	Leone I	29.IX.440	10.XI.461		Tuscia
46	Ilario	19.XI.461	29.II.468		Sardo
47	Simplicio	3.III.468	10.III.483		Tivoli
48	Felice III	13.III.483	25.II o 1.III.492		Romano
49	Gelasio I	1.III.492	21.XI.496		Africano

50	Anastasio II	24.XI.496	19.XI.498		Romano
51	Simmaco	22.XI.498	19.VII.514		Sardo
52	Ormisda	20.VII.514	6.VIII.523		
53	Giovanni I, martire	13.VIII.523	18.V.526		Tuscia
54	Felice IV	12.VII.526	20 o 22.IX.530		Sannio
55	Bonifacio II	22.IX.530	17.X.532		Romano
56	Giovanni II	31.XII.532	8.V.535	Mercurio	Romano
57	Agapito I	13.V.535	22.IV.536		Romano
58	Silverio, martire	8.VI.536	... 537		Frosinone
59	Vigilio	29.III.537	7.VI.555		Romano
60	Pelagio I	16.IV.556	4.III.561		Romano
61	Giovanni III	17.VII.561	13.VII.574	Catalino	Romano
62	Benedetto I	2.VI.575	30.VII.579		Romano
63	Pelagio II	26.XI.579	7.II.590		Romano
64	Gregorio I	3.IX.590	12.III.604		Romano
65	Sabiniano	13.IX.604	22.II.606		Blera nella Tuscia
66	Bonifacio III	19.II.607	10.XI.607		Romano
67	Bonifacio IV	25.VIII.608	8.V.615		Territorio dei Marsi
68	Deusdedit o Adeodato I	19.X.615	8.XI.618		Romano
69	Bonifacio V	23.XII.619	23.X.625		Napoli
70	Onorio I	27.X.625	12.X.638		Campania
71	Severino	X.638	2.VIII.640		Romano
72	Giovanni IV	24.XII.640	12.X.642		Dalmata
73	Teodoro I	24.XI.642	14.V.649		Gerusalem me

74	Martino I, Martire	5.VII.649	16.IX.655		Todi
75	Eugenio I	10.VIII.654	2.VI.657		Romano
76	Vitaliano	30.VII.657	27.I.672		Segni
77	Adeodato II	11.IV.672	16.VI.676		Romano
78	Dono	2.XI.676	11.IV.678		Romano
79	Agatone	27.VI.678	10.I.681		Siciliano
80	Leone II	17.VIII.682	3.VII.683		Siciliano
81	Benedetto II	26.VI.684	8.V.685		Romano
82	Giovanni V	23.VII.685	2.VIII.686		Siro
83	Conone	23.X.686	21.IX.687		patria ignota
84	Sergio I	15.XII.687	7.IX.701		Siro
85	Giovanni VI	30.X.701	11.I.705		Greco
86	Giovanni VII	1.III.705	18.X.707		Greco
87	Sisinnio	15.I.708	4.II.708		Siro
88	Costantino	25.III.708	9.IV.715		Siro
89	Gregorio II	19.V.715	11.II.731		Romano
90	Gregorio III	18.III.731	28.XI.741		Siro
91	Zaccaria	3.XII.741	15.III.752		Greco
92	Stefano II	26.III.752	26.IV.757		Romano
93	Paolo I	29.V.757	28.VI.767		Romano
94	Stefano III	7.VIII.768	24.I.772		Siciliano
95	Adriano I	1,9.II.772	25.XII.795		Romano
96	Leone III	27.XII.795	12.VI.816		Romano
97	Stefano IV	22.VI.816	24.I.817		Romano
98	Pasquale I	25.I.817	... II-V.824		Romano
99	Eugenio II	II-V.824	...VIII.827		Romano
100	Valentino	VIII.827	...IX.827		Romano
101	Gregorio IV	29.III.828	25.I.844		Romano
102	Sergio II	25.I.844	27.I.847		Romano

103	Leone IV	10.V.847	17.VII.855		Romano
104	Benedetto III	29.IX.855	17.IV.858		Romano
105	Niccolo I	24.IV.858	13.XI.867		Romano
106	Adriano II	14.XII.867	xx.XI o XII.872		Romano
107	Giovanni VIII	14.XII.872	16.XII.882		Romano
108	Marino I	XII.882	15.V.884		Gallese
109	Adriano III	17.V.884	xx.VIII o IX.885		Romano
110	Stefano V	xx.IX.885	14.IX.891		Romano
111	Formoso	6.X.891	4.IV.896		Vescovo di Porto
112	Bonifacio VI	11.IV.896	26.IV.896		Romano
113	Stefano VI	xx.VI.896	xx.VII o VIII.897		Romano
114	Romano	xx.VIII.897	xx.XI.897		Gallese
115	Teodoro II	xx.XII.897	xx.XII.897 o I.898		Romano
116	Giovanni IX	xx.XII.897	xx.I-V.900		Tivoli
117	Benedetto IV	I-V.900	xx.VII.903		Romano
118	Leone V	xx.VII.903	xx. IX.903		Ardea
119	Sergio III	29.I.904	14.IV.911		Romano
120	Anastasio III	xx.VI o IX.911	xx. VI o VIII o X.913		Romano
121	Landone	xx VII o XI.913	xx. III.914		Sabina
122	Giovanni X	xx III o IV.914	xx. V o VI.928		Tossignan o (Imola)
123	Leone VI	xx V o VI.928	xx. XII.928 o I.929		Romano
124	Stefano VII	xx.I.929	xx.II.931		Romano

125	Giovanni XI	xx.III.931	xx.I.936		Romano
126	Leone VII	xx.I.936	13.VII.939		Romano
127	Stefano VIII	14.VII.939	xx. X.942		Romano
128	Marino II	30.X o XI.942	...V.946		Romano
129	Agapito II	10.V.946	...XII.955		Romano
130	Giovanni XII	16.XII.955	14.V.964	Ottaviano	dei conti di Tuscolo
131	Leone VIII	4,6.XII.963	...III.965		Romano
132	Benedetto V	xx.V.964	4.VII.964 o 965		Romano
133	Giovanni XIII	1.X.965	6.IX.972		Romano
134	Benedetto VI	xx.XII.972, 19.I.973	xx.VII.974		Romano
135	Benedetto VII	xx.X.974	10.VII.983		Romano
136	Giovanni XIV	xx.XI o XII.983	20.VIII.984	Pietro	Pavia
137	Giovanni XV	xx.VIII.985	xx.III.996		Romano
138	Gregorio V	3.V.996	xx.II o III.999	Brunone dei duchi di Carinzia	Sassonia
139	Silvestro II	2.IV.999	12.V.1003	Gelberto	Aquitania
140	Giovanni XVII	16.V.1003	6.XI.1003	Siccone	Romano
141	Giovanni XVIII	25.XII.1003	xx.VI o VII.1009	Fasano	Romano
142	Sergio IV	31.VII.1009	12.V.1012	Pietro	Romano
143	Benedetto VIII	18.V.1012	9.IV.1024	Teofilatto dei conti di Tuscolo	

144	Giovanni XIX	19.IV.1024	1032	Romano dei conti di Tuscolo	
145	Benedetto IX	xx.VIII o IX.1032	xx.IX.1044	Teofilatto dei conti di Tuscolo	
146	Silvestro III	13 o 20.I.1045	xx.III.1045	Giovanni	Romano
147	Benedetto IX	10.III.1045	1.V.1045	Teofilatto dei conti di Tuscolo	
148	Gregorio VI	1.V.1045	20.XII.1046	Giovanni Graziano	Romano
149	Clemente II	24.XII.1046	9.X.1047	Suitgero dei signori di Morsleben von Horneburg	Sassonia
150	Benedetto IX	xx.X.1047	xx. VIII.1048	Teofilatto dei conti di Tuscolo	
151	Damaso II	17.VII.1048	9.VIII.1048	Poppone	Tirolo
152	Leone IX	2,12.II.1049	19.IV.1054	Brunone dei conti di Egisheim	Alsaziano
153	Vittore II	13.IV.1055	28.VII.1057	Gebeardo dei conti di Dollnstein-Hirschberg	Svevo
154	Stefano IX	2,3.VIII.1057	29.III.1058	Federico dei duchi di Lorena	

155	Niccolò II	xx.XII.1058, 24.I.1059	27.VII.1061	Gerardo	Borgogna
156	Alessandro II	30.IX, 1.X.1061	21.IV.1073	Anselmo	Baggio (Milano)
157	Gregorio VII	22.IV,30.VI.107 3	25.V.1085	Ildebrando	Tuscia
158	Vittore III	24.V.1086, 9.V.1087	16.IX.1087	Dauferio (Desiderio)	Benevento
159	Urbano II	12.III.1088	29.VII.1099	Oddone di Lagery	Châtillon- sur-Marne
160	Pasquale II	13,14.VIII.1099	21.I.1118	Raniero	Bleda o Galeata
161	Gelasio II	24.I,10.III.1118	28.I.1119	Giovanni Caetani	Gaeta
162	Callisto II	2,9.II.1119	13 o 14.XII.1124	Guido di Borgogna	
163	Onorio II	15,21.XII.1124	13 o 14.II.1130	Lamberto Scannabec chi	Fiagnano (Imola)
164	Innocenzo II	14,23.II.1130	24.IX.1143	Gregorio Papareschi	Romano
165	Celestino II	26.IX,3.X.1143	8.III.1144	Guido	Città di Castello
166	Lucio II	12.III.1144	15.II.1145	Gerardo	Bolognese
167	Eugenio III	15,18.II.1145	8.VII.1153	Bernardo	Pisa
168	Anastasio IV	12.VII.1153	3.XII.1154	Corrado	Romano
169	Adriano IV	4,5.XII.1154	1.IX.1159	Nicola Breakspear	Abbot's Lagnley (Hertfords hire)
170	Alessandro III	7,20.IX.1159	30.VIII.1181	Rolando Bandinelli	Siena

171	Lucio III	1. 6.IX.1181	25.IX.1185	Ubaldo Allucingoli	Lucchese
172	Urbano III	25.XI,1.XII.118 5	20.X.1187	Umberto Crivelli	Milanese
173	Gregorio VIII	21.25.X.1187	17.XII.1187	Alberto di Morra	Benevento
174	Clemente III	19,20.XII.1187	xx.III.1191	Paolo Scolari	Romano
175	Celestino III	10,14.IV.1191	8.I.1198	Giacinto Bobone	Romano
176	Innocenzo III	8.I,22.II.1198	16.VII.1216	Lotario dei conti di Segni	Gavignano (Roma)
177	Onorio III	18,24.VII.1216	18.III.1227	Cencio	Romano
178	Gregorio IX	19,21.III.1227	22.VIII.1241	Ugolino dei conti di Segni	Anagni
179	Celestino IV	25,28.X.1241	10.XI.1241	Goffredo da Castiglione	Milanese
180	Innocenzo IV	25,28.VI.1243	7.XII.1254	Sinibaldo Fieschi	Lavagna (Genova)
181	Alessandro IV	12,20.XII.1254	25.V.1261	Rinaldo dei signori di Ienne	Ienne (Roma)
182	Urbano IV	29.VIII,4.IX.126 1	2.X.1264	Giacomo Pantaléon	Troyes
183	Clemente IV	5,22.II.1265	29.XI.1268	Guido Foulques	Saint- Gilles (Francia meridional e)

184	Gregorio X	1.IX.1271,27.III .1272	10.I.1276	Tebaldo Visconti	Piacenza
185	Innocenzo V	21.I,22.II.1276	22.VI.1276	Pietro di Tarentaise	Savoia
186	Adriano V	11.VII.1276	18.VIII.1276	Ottobono Fieschi	Genovese
187	Giovanni XXI	16,20.IX.1276	20.V.1277	Pietro di Giuliano o Pietro Ispano	Lisbona
188	Niccolò III	25.XI, 26.XII.1277	22.VIII.1280	Giovanni Gaetano Orsini	Romano
189	Martino IV	22.II,23.III.1281	29.III.1285	Simone de Brie o di Brion o di Mainpincien	Francese
190	Onorio IV	2.IV, 20.V.1285	3.IV.1287	Giacomo Savelli	Romano
191	Niccolò IV	22.II.1288	4.IV.1292	Girolamo	Lisciano (Ascoli Piceno)
192	Celestino V	5.VII, 29.VIII.1294	13.XII.1294	Pietro del Morrone	Molise
193	Bonifacio VIII	24.XII.1294, 23.I.1295	11.X.1303	Benedetto Caetani	Anagni
194	Benedetto XI	22,27.X.1303	7.VII.1304	Niccolò di Boccasio	Treviso
195	Clemente V	5.VI, 14.XI.1305	20.IV.1314	Bertrando de Got	Villandrau t (Gironde)

196	Giovanni XXII	7.VIII,5.IX.1316	4.XII.1334	Giacomo Duèse	Cahors
197	Benedetto XII	20.XII.1334, 8.I.1335	25.IV.1342	Giacomo Fournier	Saverdun (Francia meridionale)
198	Clemente VI	7,19.V.1342	6.XII.1352	Pietro Roger	Maumont (Limosino)
199	Innocenzo VI	18,30.XII.1352	12.IX.1362	Stefano Aubert	Monts (Limosino)
200	Urbano V	28.IX,6.XI.1362	19.XII.1370	Guglielmo Grimoard	Grizac (Francia meridionale)
201	Gregorio XI	30.XII.1370, 3.I.1371	26.III.1378	Pietro Roger de Beaufort	Rosiers d'Egletons (Limosino)
202	Urbano VI	8,18.IV.1378	15.X.1389	Bartolomeo Prignano	Napoli
203	Bonifacio IX	2,9.XI.1389	1.X.1404	Pietro Tomacelli	Napoli
204	Innocenzo VII	17.X,11.XI.1404	6.XI.1406	Cosma Migliorati	Sulmona
205	Gregorio XII	30.XI,19.XII.1406	4.VII.1415	Angelo Correr	Veneziano
206	Martino V	11,21.XI.1417	20.II.1431	Oddone Colonna	Genazzano
207	Eugenio IV	3,11.III.1431	23.II.1447	Gabriele Condulmer	Veneziano

208	Niccolò V	6,19.III.1447	24.III.1455	Tommaso Parentucelli	Sarzana
209	Callisto III	8,20.IV.1455	6.VIII.1458	Alonso Borja	Torre del Canals presso Játiva (Valencia)
210	Pio II	19.VIII, 3.IX.1458	14.VIII.1464	Enea Silvio Piccolomini	Corsignano (Siena)
211	Paolo II	30.VIII, 16.IX.1464	26.VII.1471	Pietro Barbo	Veneziano
212	Sisto IV	1,9,25.VIII.1471	12.VIII.1484	Francesco della Rovere	Celle (Savona)
213	Innocenzo VIII	29.VIII, 12.IX.1484	25.VII.1492	Giovanni Battista Cibo	Genovese
214	Alessandro VI	11,26.VIII.1492	18.VIII.1503	Rodrigo de Borja	Játiva (Valencia)
215	Pio III	22.IX, 1,8.X.1503	18.X.1503	Francesco Todeschini-Piccolomini	Siena
216	Giulio II	1,26.XI.1503	21.II.1513	Giuliano della Rovere	Albisola (Savona)
217	Leone X	11,19.III.1513	1.XII.1521	Giovanni de' Medici	Fiorentino
218	Adriano VI	9.I,31.VIII.1522	14.IX.1523	Adriano Florensz	Utrecht
219	Clemente VII	19,26.XI.1523	25.IX.1534	Giulio de' Medici	Fiorentino
220	Paolo III	13.X, 3.XI.1534	10.XI.1549	Alessandro Farnese	Canino (Viterbo)

221	Giulio III	7,22.II.1550	23.III.1555	Giovanni Maria Ciocchi del Monte	Romano
222	Marcello II	9,10.IV.1555	1.V.1555	Marcello Cervini	Montefano
223	Paolo IV	23,26.V.1555	18.VIII.1559	Gian Pietro Carafa	Capriglia (Avellino)
224	Pio IV	26.XII.1559, 6.I.1560	9.XII.1565	Giovan Angelo Medici	Milanese
225	Pio V	7,17.I.1566	1.V.1572	Antonio (Michele) Ghisleri	Bosco (Alessandr ia)
226	Gregorio XIII	13,25.V.1572	10.IV.1585	Ugo Boncompag ni	Bolognese
227	Sisto V	24.IV, 1.V.1585	27.VIII.1590	Felice Peretti	Grottamm are (Ascoli Piceno)
228	Urbano VII	15.IX.1590	27.IX.1590	Giambattist a Castagna	Romano
229	Gregorio XIV	5,8.XII.1590	16.X.1591	Niccolò Sfondrati	Somma Lombarda
230	Innocenzo IX	29.X,3.XI.1591	30.XII.1591	Giovan Antonio Facchinetti	Bolognese
231	Clemente VIII	30.I,9.II.1592	3.III.1605	Ippolito Aldobrandin i	Fano
232	Leone XI	1,10.IV.1605	27.IV.1605	Alessandro de' Medici	Fiorentino

233	Paolo V	16,29.V.1605	28.I.1621	Camillo Borghese	Romano
234	Gregorio XV	9,14.II.1621	8.VII.1623	Alessandro Ludovisi	Bolognese
235	Urbano VIII	6.VIII, 29.IX.1623	29.VII.1644	Maffeo Barberini	Fiorentino
236	Innocenzo X	15.IX,4.X.1644	7.I.1655	Giovanni Battista Pamphilj	Romano
237	Alessandro VII	7,18.IV.1655	22.V.1667	Fabio Chigi	Siena
238	Clemente IX	20,26.VI.1667	9.XII.1669	Giulio Rospigliosi	Pistoia
239	Clemente X	29.IV,11.V.1670	22.VII.1676	Emilio Altieri	Romano
240	Innocenzo XI	21.IX, 4.X.1676	12.VIII.1689	Benedetto Odescalchi	Como
241	Alessandro VIII	6,16.X.1689	1.II.1691	Pietro Ottoboni	Veneziano
242	Innocenzo XII	12,15.VII.1691	27.IX.1700	Antonio Pignatelli	Spinazzola
243	Clemente XI	23,30.XI, 8.XII.1700	19.III.1721	Giovanni Francesco Albani	Urbino
244	Innocenzo XIII	8,18.V.1721	7.III.1724	Michelangel o Conti	Poli
245	Benedetto XIII	29.V, 4.VI.1724	21.II.1730	Pietro Francesco (Vincenzo Maria) Orsini	Gravina

246	Clemente XII	12,16.VII.1730	6.II.1740	Lorenzo Corsini	Fiorentino
247	Benedetto XIV	17,22.VIII.1740	3.V.1758	Prospero Lambertini	Bolognese
248	Clemente XIII	6,16.VII.1758	2.II.1769	Carlo Rezzonico	Veneziano
249	Clemente XIV	19,28.V, 4.VI.1769	22.IX.1774	Giovanni Vincenzo Antonio (Lorenzo) Ganganelli	Sant'Arcan gelo di Romagna
250	Pio VI	15,22.II.1775	29.VIII.1799	Giannangel o Braschi	Cesena
251	Pio VII	14,21.III.1800	20.VIII.1823	Barnaba (Gregorio) Chiaromonti	Cesena
252	Leone XII	28.IX, 5.X.1823	10.II.1829	Annibale della Genga	Monticelli di Genga (Fabriano)
253	Pio VIII	31.III, 5.IV.1829	30.XI.1830	Francesco Saverio Castiglioni	Cingoli
254	Gregorio XVI	2,6.II.1831	1.VI.1846	Bartolomeo Alberto (Mauro) Cappellari	Belluno
255	Pio IX	16,21.VI.1846	7.II.1878	Giovanni Maria Mastai Ferretti	Senigallia

256	Leone XIII	20.II, 3.III.1878	20.VII.1903	Vincenzo Gioacchino Pecci	Carpineto Romano
257	Pio X	4,9.VIII.1903	20.VIII.1914	Giuseppe Melchiorre Sarto	Riese (Treviso)
258	Benedetto XV	3,6.IX.1914	22.I.1922	Giacomo della Chiesa	
259	Pio XI	6,12.II.1922	10.II.1939	Achille Ratti	Desio (Milano)
260	Pio XII	2,12.III.1939	9.X.1958	Eugenio Pacelli	Romano
261	Giovanni XXIII	28.X, 4.XI.1958	3.VI.1963	Angelo Giuseppe Roncalli	Sotto il Monte (Bergamo)
262	Paolo VI	21,30.VI.1963	6.VIII.1978	Giovanni Battista Montini	Concesio (Brescia)
263	Giovanni Paolo I	26.VIII, 3.IX.1978	28.IX.1978	Albino Luciani	Forno di Canale (Belluno)
264	Giovanni Paolo II	16,22.X.1978	2.IV.2005	Karol Wojtyła	Wadowice (Kraków)
265	Benedetto XVI	19,24.IV.2005	28.II.2013	Joseph Ratzinger	Marktl am Inn (Baviera)
266	Francesco	13,19.III.2013		Jorge Mario Bergoglio	Buenos Aires (Argentina)

I protagonisti della *content analysis* (in ordine alfabetico).

Jean-Paul Benzécri (Oran, 28 febbraio 1932 – Villampuy, 24 novembre 2019) è stato un matematico e statistico francese, promotore della scuola francese dell'*Analyse des Données*. Famoso soprattutto per lo sviluppo dell'analisi della corrispondenza, una tecnica statistica per l'analisi delle tabelle di contingenza.

Bernard Reuben Berelson (Spokane, 1912 - Sleepy Hollow 1979) è stato uno scienziato comportamentale americano, importanti i suoi studi sulla comunicazione e sui mass media. È stato uno dei principali sostenitori dell'idea allargata delle "scienze comportamentali".

Herbert George Blumer (Saint Louis, 7 marzo 1900 – 13 aprile 1987) è stato un sociologo statunitense i cui principali interessi accademici furono incentrati sui metodi della ricerca sociale.

Klaus Krippendorff (Frankfurt am Main, 1932) è professore di Cybernetics, Language and Culture presso la Annenberg School for Communication, University of Pennsylvania, Philadelphia, USA. Dottorato in scienze della comunicazione presso l'Istituto di Comunicazione dell'Università dell'Illinois (Urbana-Champaign). Ha pubblicato testi sulla cibernetica e la teoria dei sistemi, la metodologia della ricerca sociale, la comunicazione umana, il discorso, la conversazione.

Paul Felix Lazarsfeld (Vienna, 13 febbraio 1901 – New York, 30 giugno 1976) è stato un sociologo statunitense di origini austriache. Professore ordinario di sociologia alla Columbia *University* è stato l'ideatore di molti aspetti della metodologia della ricerca sociale empirica.

Harold Dwight Lasswell (Donnellson, 13 febbraio 1902 – New York, 18 dicembre 1978) politologo statunitense, noto per i suoi studi dedicati alle scienze politiche e alle teorie della comunicazione. Membro della Scuola sociologica di Chicago, riteneva che la

propaganda fosse necessaria per la democrazia al fine di mantenere i cittadini. A lui si deve la nascita del nome del termine *content analysis*.

Pitirim Aleksandrovič Sorokin (Tur'ja, 21 gennaio 1889 – Winchester, 10 febbraio 1968) è stato un sociologo russo e statunitense, rivoluzionario e uomo politico. L'Università di Harvard lo invitò, nel 1930, a costituire il "Dipartimento di Sociologia" e nel 1949 fondò l'"*Harvard Research Center in Creative Altruism*" dedicandosi allo studio dei fenomeni sociali morali.

William I. Thomas (Contea di Russell, 13 agosto 1863 – Berkeley, 5 dicembre 1947) è stato un sociologo statunitense. Fu a capo della scuola di Chicago fino al 1918, si occupò della condizione degli immigrati in America, la sua opera più importante rimane "Il contadino polacco in Europa e in America" scritta con Florian Znaniecki.

Sergius Jakobson (Mosca, 9 maggio 1901 - Washington, 13 novembre 1979) è stato un russo-tedesco bibliotecario, storico e sovietologo americano. Bibliotecario presso la Biblioteca del Congresso degli Stati Uniti d'America uno dei massimi esperti di lingue dell'Europa orientale e storia dell'Europa. Fratello di del linguista Roman Ossipowitsch Jakobson.

Florian Znaniecki (Świątniki, 15 gennaio 1882 – Champaign, 23 marzo 1958) è stato un filosofo e sociologo polacco, emigrato negli Stati Uniti allo scoppio della Prima Guerra Mondiale fece parte della Scuola di Chicago definita la scuola dell'ecologia sociale urbana.

L'utilizzo del *software* Iramuteq

La costruzione del *corpus* per l'analisi dei dati con il *software*

La costruzione del *corpus* necessita la raccolta di tutti i testi in un file con estensione “.txt”. Ogni testo, riportato nel file, deve essere preceduto da una stringa ben definita. La stringa, per motivi di analisi del *software*, ha una precisa sintassi e identifica tutte le variabili: **** *var1_modalità *var2_modalità

Inizia con quattro asterischi, poi c'è uno spazio, un asterisco con la prima variabile e la relativa modalità preceduta da un *underscore*, poi uno spazio con un altro asterisco con la seconda variabile e la sua modalità preceduta da un *underscore* e così via per tutte le variabili che ho prefissato.

Riporto di seguito l'esempio della prima stringa del *corpus* della mia ricerca:

```
**** *papa_paoloVI *anno_1963 *mese_settembre1963 *destinatario_fuci *argomento_inpastorale
```

Il file di testo.txt così formato, prima di essere importato nel *software* Iramuteq, dovrà essere salvato in modalità “UTF8”.

Parametri della procedura di costruzione dei *cluster* semantici

Il numero di classi semantiche o mondi lessicali che si ottengono analizzando un *corpus* con il *software* dipende dai parametri che imposto (Figura A1).

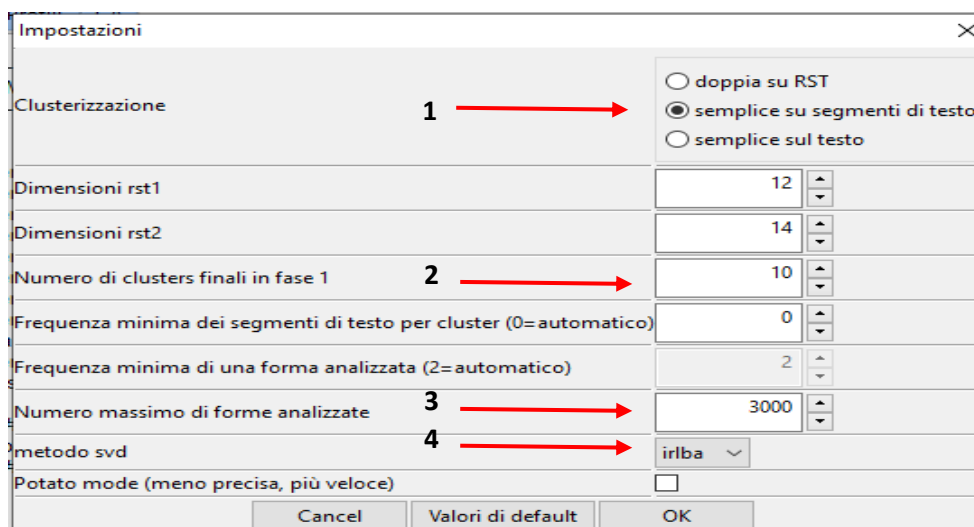


Figura A1. Stralcio immagine dei parametri del *software* Iramuteq

I parametri da tenere in considerazione sono in particolar modo quattro:

1. “Clusterizzazione”: procedura di costruzione dei cluster semantici: il tipo di classificazione che viene eseguita, scegliendo “semplice sui segmenti di testo” aggrega tutte le porzioni di testo. Se si scegliesse “semplice sul testo”, verrebbe fatta l’analisi aggregando tutti i testi. Quest’ultima andrebbe bene per testi molto brevi (come i *tweet*) dove non ha senso scomporli in porzioni.
2. Numero di *cluster* finali in fase 1: è il parametro che determina il numero massimo dei *cluster* che il *software* deve trovare e che determina il fine della procedura di iterazione. Impostato a dieci il *software*, nella costruzione dei *cluster*, controlla che la nona classe semantica sia sufficientemente distante dalla decima (in base al chi quadrato) e se ciò non fosse, la aggrega con un’altra. Procedo in questa maniera con tutte le classi trovate. Di solito, con un parametro impostato a dieci, vengono trovate quattro o cinque classi semantiche e l’impostazione del parametro pari a dieci è il valore di *default* oltre che quello di riferimento in letteratura. Aumentando il valore del parametro il *software* troverà più classi semantiche (più categorie ma con meno differenze e con qualche parola condivisa) e viceversa diminuendolo segnalerà meno *topics*.
3. Numero massimo di forme analizzate: il parametro di *default* è 3000. È però utile tenere in considerazione le forme e gli *hapax* del *corpus* da analizzare. Se ho 6.000 forme e 3.000 *hapax*, con un parametro impostato a 3.000, prendo tutte le parole che ho a disposizione. Se lo imposto a 1.500 invece vado a lavorare con parole con una soglia un po' più alta (uguale a due per la precisione, che significa che la parola è presente in almeno due segmenti di testo analizzato).
4. *Potato mode*: è il metodo utilizzato per il controllo della procedura di costruzione dei cluster semantici, meno preciso ma più veloce, per analisi di *corpus* di grandezza nell’ordine dei *big data*.