

POLITECNICO DI MILANO

Scuola di Ingegneria dell'Informazione

Corso di Laurea Magistrale in
Ingegneria Informatica



***Le abitudini degli accademici nella ricerca di letteratura scientifica:
evidenza empirica sul processo di discovery di articoli rilevanti***

Relatore: Prof. Chiara FRANCALANCI

Tesi di Laurea di:

Niccolò Zamperetti Matr. 814656

Anno Accademico 2014 - 2015

Sommario

Le abitudini degli accademici nella ricerca di letteratura scientifica:	1
1 Indice dei contenuti	3
1.1 Tabelle	3
1.2 Grafici.....	3
2 Introduzione	4
3 Stato dell'arte	7
3.1 Introduzione.....	7
3.2 Indicatori Bibliometrici	8
3.2.1 Impact Factor (IF), Immediacy Index e Citation Half-Life [1]	9
3.2.1.1 Svantaggi e critiche.....	10
3.2.2 Source Normalized Impact per Paper (SNIP) [21]	12
3.2.3 Y-Factor [22]	14
3.2.4 Eigen Factor e Article Influence Score [4]	16
3.2.5 SCImago Journal Rank (SJR) [24]	18
3.2.6 H-index [5].....	20
3.2.6.1 Vantaggi, svantaggi e critiche.....	21
3.2.6.2 Varianti	23
3.3 Search Engines & Online Archives.....	24
3.3.1 Google Scholar.....	25
3.3.2 Scopus	26
3.3.3 Web of Science (Web of Knowledge)	27
3.3.4 Analisi comparativa	27
3.4 Conclusioni.....	29
4 Metodologia	31
4.1 Introduzione.....	31
4.2 Il concetto di impatto	32
4.3 Impatto negli strumenti attuali	35
4.4 Fattibilità.....	37
4.5 Abitudini di ricerca	41
4.6 Ipotesi	47
4.7 Conclusione	51
5 Analisi dei dati	52
5.1 Introduzione.....	52
5.2 Analisi qualitativa globale	54
5.3 Analisi delle correlazioni	59
5.3.1 Metodo	59
5.3.2 Analisi	60
5.4 Test.....	63
5.4.1 Metodo	63
5.4.2 Analisi	68
5.4.2.1 Fonti di reperimento delle letture	68
5.4.2.2 Utilizzo di motori e archivi online per la scelta del topic di ricerca	75
5.5 Conclusione	81
6 Conclusioni	82
7 Bibliografia	84
8 Appendice	87

1 Indice dei contenuti

1.1 Tabelle

Tab. 1 Confronto G-index h-index [32].....	23
Tab. 2 Top ten Search Engine Global Rank [51]	28
Tab. 3 Least Visited Search Engines [51]	28
Tab. 4 Research Network Global Rank [51]	28
Tab. 5 Selezione di risultati rilevanti nelle correlazioni di Spearman	61
Tab. 6 Sommario dei risultati dei test di Mann-Whitney e χ^2 per il reperimento delle letture	71
Tab. 7 Sommario dei risultati dei test di Mann-Whitney e χ^2 per l'utilizzo degli attuali strumenti per la scelta del proprio topic di ricerca	77
Tab. 8 Risultati globali del questionario.....	87
Tab. 9 Coefficienti di correlazione di Spearman e relativi p-value	92
Tab. 10 Coefficienti di correlazione di Pearson e relativi p-value	98
Tab. 11 Coefficienti di correlazione di Kendall e relativi p-value	104
Tab. 12 T-Test rilevanti per l'analisi sulle fonti di letture	109
Tab. 13 Frequenze reali-attese e relativo test χ^2 d'indipendenza per l'analisi delle fonti di lettura	111
Tab. 14 T-Test rilevanti per l'analisi sull'utilizzo dei motori di ricerca e archivi online nella scelta del topic di ricerca	113
Tab. 15 Frequenze reali-attese e relativo test χ^2 d'indipendenza per l'analisi sull'utilizzo dei motori di ricerca nella scelta del proprio topic	115
Tab. 16 Test U di Mann-Whitney significativi più rilevanti per le fonti di reperimento delle letture	118
Tab. 17 Test U di Mann-Whitney significativi più rilevanti sull'utilizzo di tools nella scelta del proprio topic di ricerca.....	120

1.2 Grafici

Grafico 1 Distribuzione geografica	53
Grafico 2 Distribuzione nei Journals	54

2 Introduzione

La quantità di letteratura scientifica è aumentata notevolmente nel corso degli anni. Inoltre l'avvento del web ha contribuito ulteriormente a questa proliferazione, poiché i periodici, gli archivi e anche i singoli individui hanno cominciato a diffondere i propri lavori online. Con questa incredibile crescita di informazioni disponibili, il reperimento di pubblicazioni rilevanti, d'impatto e che quindi valgano la pena di essere lette è un problema impegnativo.

Parallelamente, hanno cominciato a diffondersi metriche che valutino l'impatto della produzione scientifica in modo più quantitativo, ovvero gli indicatori bibliometrici, e motori di ricerca indirizzati esclusivamente alla letteratura scientifica. Nonostante le moltitudini di critiche attratte dagli indicatori, al giorno d'oggi sembrano rimanere la misura più consolidata di "impatto" della produzione scientifica. Gli attuali motori di ricerca d'altro canto sono considerati generalmente limitati, non permettendo ad esempio una parametrizzazione della ricerca in base a questi stessi indicatori (sebbene siano i principali strumenti a calcolarli), oltre al fatto che le parole chiave cercate si estendono a molteplici campi scientifici, recuperando un vasto numero di pubblicazioni che non sono pertinenti con gli interessi dell'utente. Inoltre, la definizione di campo scientifico fornita dagli archivi online è solitamente molto ampia e lontana dalla nozione più pratica di campo di ricerca.

Questo lavoro parte dall'osservazione che affrontare un progetto per migliorare la ricerca online, in una direzione magari più mirata ai contenuti, è uno sforzo fattibile solo se le reali abitudini del mondo accademico dimostrano una necessità tangibile di meccanismi di ricerca migliori e la volontà di investire tempo e impegno nella scoperta di contenuti rilevanti. Il lavoro di tesi parte da questa osservazione e si pone l'obiettivo di verificare l'opportunità per il miglioramento dei meccanismi di ricerca, delineandone alcune linee guida operative. La prima domanda di ricerca di questo lavoro, quindi, è se gli accademici facciano effettivamente ricerca online e in che misura. La seconda domanda di ricerca è se utilizzino i meccanismi di ricerca avanzata degli attuali strumenti e in che misura li valutino efficaci. Questo fornirebbe indicazioni circa la loro predisposizione all'utilizzo di funzionalità di ricerca più sofisticate e che, potenzialmente, facilitino l'identificazione di letteratura rilevante. La terza domanda di ricerca si chiede come la ricerca venga operata nelle varie fasi del loro lavoro scientifico, dalla scelta dell'argomento sul quale fare ricerca,

all'identificazione della letteratura di riferimento e selezione delle citazioni da inserire nei propri articoli. Questo allo scopo di sondare l'effettiva necessità nel mondo accademico di uno strumento più complesso per reperire contenuti più rilevanti in breve tempo, oltre a capire le direzioni che si potrebbero prendere in tal senso.

Nella cercare di rispondere a queste domande di ricerca, si è posta particolare attenzione all'utilizzo degli indicatori bibliometrici di impatto in fase di search e alla loro implementazione all'interno degli attuali motori di ricerca. Come già osservato, dato l'obiettivo ultimo della ricerca di migliorare nell'identificazione di letteratura rilevante, capire se e come gli indicatori di impatto possano aiutare in questo processo rappresenta un passo quasi obbligato. Da questo punto di vista, ci si è resi conto che le nostre domande di ricerca possono anche aiutarci a capire se il fatto di effettuare la ricerca online possa in qualche modo influenzare negativamente alcuni comportamenti considerati viziosi in letteratura, specialmente in relazione alle conseguenze negative degli indicatori di impatto, quali:

- Una maggiore gregarietà e minor indipendenza di pensiero, ad esempio nel seguire assertivamente le direzioni degli autori di maggiore impatto nel proprio campo.
- Un maggiore opportunismo, ad esempio nella scelta delle citazioni da inserire in un proprio lavoro in modo condizionato al proprio outlet di pubblicazione.
- Una maggiore sensibilità agli indicatori bibliometrici, facendosi ad esempio influenzare eccessivamente da essi nelle varie fasi del proprio lavoro scientifico.

E' noto che l'automazione di qualunque attività tende a peggiorare processi che già abbiano dei problemi di prestazioni. Nel nostro caso, la ricerca di letteratura online potrebbe avere un effetto negativo sul processo di ricerca sopra delineato e, in particolare, potrebbe aumentare i comportamenti meno virtuosi, ovvero più opportunistici, degli accademici specialmente in relazione all'utilizzo degli indicatori di impatto. Di conseguenza, all'interno di questo lavoro di tesi si è formulata l'ipotesi che ci sia effettivamente un'influenza negativa dei motori di ricerca (online) sul processo di ricerca (accademica).

Per aiutare a rispondere a queste domande e validare le ipotesi fatte, questo lavoro presenta i risultati di un questionario empirico condotto con accademici nel campo

dei Management Information Systems, allo scopo di valutarne le abitudini di ricerca. I risultati da 326 intervistati svelano interessanti approfondimenti sulle abitudini di ricerca degli accademici e, in generale, indicano che nonostante il consenso sulla scarsa qualità degli attuali meccanismi di ricerca, solo una piccola minoranza di utenti sarebbe disposta a scambiare la semplicità di ricerca per la rilevanza dei risultati.

3 Stato dell'arte

3.1 Introduzione

L'avvento del web ha sicuramente cambiato il modo in cui gli accademici si avvicinano alla letteratura. Anzitutto, l'ammontare di documentazioni scientifiche disponibili online è cresciuto a dismisura nelle ultime decadi, con i vari archivi, periodici e gli stessi ricercatori che hanno contribuito alla diffusione digitale delle loro pubblicazioni.

Le pubblicazioni su periodici continuano a essere peer-reviewed, ma la crescente quantità di documenti ha portato alla necessità di avere indicatori più quantitativi della letteratura scientifica. Garfield ha inizialmente proposto il suo Impact Factor [1] proprio come ausilio alle biblioteche di riviste scientifiche, per avere un indicatore da monitorare nel processo di analisi di costi-benefici dei budgets di sottoscrizione, così da avere delle euristiche che permettessero di scegliere di abbonarsi a un periodico piuttosto che a un altro. Tuttavia, da allora, l'immediatezza di tale indicatore (e suoi simili) ha iniziato a interessare anche agenzie di finanziamento, comitati di promozione, università nazionali, politici etc., i quali hanno cominciato a usarli come metro di valutazione della qualità scientifica di riviste e individui. Questo ha avuto un impatto sia sull'editoria scientifica sia sui singoli ricercatori, che hanno cominciato a vedersi valutati sulla base di misure delle proprie "performance" scientifiche, le quali avrebbero influenzato sempre di più tutta la loro carriera. Le riviste e i periodici scientifici si sono resi conto di poter acquisire maggior visibilità e prestigio ottenendo dei buoni valori in tali indicatori, mentre gli accademici hanno capito di aver maggiori possibilità di carriera (o viceversa, di non averne molte in caso contrario).

Inoltre l'incremento della documentazione scientifica online non ha solo portato alla necessità di indicatori quantitativi, ma anche alla necessità di accedere ed effettuare ricerche in modo efficiente ed efficace in letteratura. A tale scopo sono nati motori di ricerca della letteratura scientifica. Questi strumenti a loro volta, pur colmando la lacuna sopra descritta, hanno avuto una notevole influenza sulle abitudini del mondo accademico, poiché il modo in cui presentano i risultati all'utente può portare alla formazione di opinioni molto diverse a seconda del loro ordinamento. Oltretutto tali motori contribuiscono a incrementare la percezione degli accademici che il proprio lavoro abbia una certa visibilità nel mondo scientifico.

Per questi motivi è bene ripercorrere la letteratura inerente gli indicatori bibliometrici e i motori di ricerca, in modo da avere un quadro generale dell'attuale panorama scientifico.

3.2 Indicatori Bibliometrici

Nella scienza il valore è definito dal grado con cui l'idea di qualcuno ha contribuito liberamente alla conoscenza e ha avuto un impatto sul pensiero altrui. Dal momento che gli autori usano le citazioni per indicare quali pubblicazioni abbiano influenzato il loro lavoro, l'impatto scientifico può essere misurato in funzione del numero di citazioni che una pubblicazione riceve. Cercando misure quantitative dell'impatto scientifico, negli ultimi anni sono proliferati numerosi indicatori basati su questo principio, relativi a riviste e periodici scientifici (Journals) o singoli autori.

Per quanto riguarda la valutazione dell'influenza dei Journals la maggior parte degli indicatori si può suddividere in due macro categorie:

- Basati sul numero di citazioni a un Journal: non tenendo in considerazione la fonte della citazione, sembrano più una misura del fattore di popolarità dello status di un Journal. Fa parte di questa categoria l'indice forse più famoso: l'Impact Factor [1], sebbene il suo intento originale fosse di aiutare le biblioteche di riviste scientifiche a effettuare una analisi di costi-benefici dei budgets di sottoscrizione, piuttosto che valutare la qualità di una rivista scientifica.
- Basati su algoritmi iterativi simili al PageRank di Google [2] [3]: nel processo iterativo, ad ogni passo un Journal trasferisce parte del suo 'prestigio' accumulato fino a quel momento ai Journals che cita a sua volta. Queste misure quindi tengono in considerazione la fonte di una citazione, e sembrano misurare il fattore di prestigio dello status di un Journal. Tra gli altri, fa parte di questa categoria l'EigenFactor [4].

Per la valutazione degli autori, l'indicatore più emergente (e controverso) è l'H-index [5], che tenta di unire in un'unica misura la reputazione e la produttività degli autori.

3.2.1 Impact Factor (IF), Immediacy Index e Citation Half-Life [1]

L'impact Factor è solo una delle 3 misure definite dall'Institute of Scientific Information (ISI) per valutare come un Journal riceva citazioni nel tempo. Le citazioni tendono a seguire una curva come quella nella Fig. 1 .

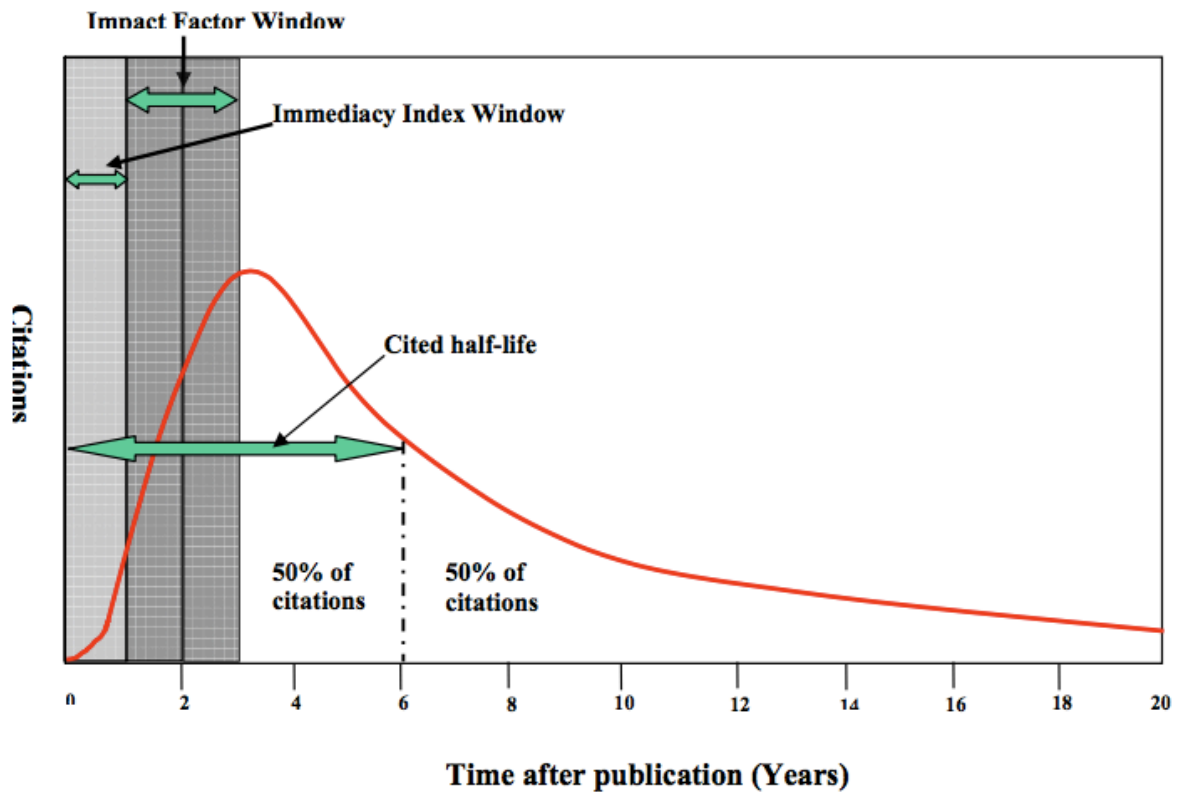


Fig. 1 Generalized Citation Curve [6]

- L'**Impact Factor** è una misura della dimensione relativa della curva nel 2° e 3° anno. È calcolato dividendo il numero di citazioni correnti che un Journal riceve papers ('citabili' o meno) pubblicati nei due anni precedenti, per il numero di papers ('citabili') pubblicati in quegli stessi anni. Ad es., l'IF per il 2014 è il numero di citazioni ottenute nel 2014 ad articoli del 2012-2013, divise per il n° di articoli pubblicati nel 2012-2013. Può essere pensata come il numero medio di citazioni che un articolo medio riceve per annum nei due anni successivi alla pubblicazione. Solo gli elementi classificati come 'articoli', 'reviews' (revisioni) e 'proceeding papers' sono considerati 'citabili'. Quindi solo questi elementi sono contati nel denominatore, mentre le citazioni a tutti i papers (anche editoriali, news, lettere agli editori, etc.) sono contate nel numeratore.

- L'**immediacy index** è calcolato dividendo le citazioni che un journal riceve nell'anno corrente , per il numero di articoli che ha pubblicato in quell'anno. Dà una misura della asimmetria della curva, ovvero di quanto il picco ricada vicino all'origine. Può essere considerata una misura di quanto velocemente gli oggetti vengano pubblicati dal momento delle pubblicazione.
- La **citation half-life** è il numero di anni che il numero di citazioni attuale impiega per decrescere del 50%. E' una misura del tasso di declino della curva, ovvero di quanto a lungo gli articoli nel journal continuano a essere citati dopo le pubblicazione.

3.2.1.1 Svantaggi e critiche

L'impact factor è stato protagonista di innumerevoli critiche nel corso degli anni, che hanno fatto emergere numerosi difetti e debolezze:

- *Auto-citazioni*: il calcolo dell'IF considera anche le auto-citazioni, ovvero le citazioni provenienti da articoli dello stesso Journal, che può quindi cercare di manipolarlo inducendo gli autori a inserire citazioni a propri articoli passati [7] [8].
- *Elementi 'non-citabili' e problema del Numeratore/Denominatore* [6]: Come già accennato sopra vi è una certa asimmetria nel calcolo del rapporto che definisce l'IF. Gli elementi marcati non-'citabili' (come editoriali, news, lettere agli editori ...) pur apportando un basso contributo scientifico contengono spesso citazioni (contate nel numeratore) ad articoli dello stesso Journal che li pubblica, ma non essendo coinvolti nel denominatore contribuiscono notevolmente ad aumentarne l'IF. Di conseguenza i journal possono manipolare il proprio IF attraverso una pubblicazione indiscriminata di questi elementi.
- *Articoli teorici*: Risulta che gli articoli teorici ricevano il doppio delle citazioni rispetto ad articoli empirici [9], e questo potrebbe essere usato per manipolare l'indicatore con una preferenza forzata a pubblicare articoli di carattere teorico.
- *Parzialità per la lingua inglese*: la preferenza del Science Citation Index per Journal in lingua inglese contribuisce a un basso IF per i pochi Journal non inglesi inclusi nel database [10], dal momento che la maggior parte delle

citazioni a papers non inglesi provengono da altri papers non in lingua inglese [11] [12].

- *Reviews*: Generalmente le reviews ricevono il doppio delle citazioni rispetto ad articoli originali [13] [14]. Di conseguenza gli editori possono cercare di migliorare l'IF prediligendo la pubblicazione di reviews rispetto ad altri articoli. Infatti i Journal che pubblicano soprattutto reviews o materiale di supporto tendono a essere in posizioni migliori in classifica.
- *Campo scientifico*: l'IF esibisce una significativa variazione a seconda del campo scientifico considerato [6]. Di conseguenza, comunità scientifiche più ristrette genereranno un minor numero di citazioni [15], oltre che a presentare abitudini di citazione del tutto specifiche di quel campo. Per questi motivi non sono paragonabili gli IF di journal appartenenti a diversi campi scientifici.
- *Outliers*: singoli articoli che ottengono un gran numero di citazioni possono portare a variazioni considerevoli nell'IF di un Journal [16]. Di conseguenza, spesso l'IF è influenzato da pochi articoli molto citati e non risulta quindi rappresentativo di un articolo 'medio' di quel Journal.
- *Hot topic*: gli autori tendono a preferire argomenti di ricerca "hot", poiché contribuiscono ad aumentare l'IF [17].
- *Riduzione di interdisciplinarietà*: emerge che i journals più alti in classifica secondo l'IF coprono una gamma meno estesa di discipline rispetto a quelli più in basso, quindi sembra dannoso per l'interdisciplinarietà [18].
- *Studi negativi o di conferma*: questo tipo di articoli attira generalmente molta meno attenzione dell'articolo originale dal quale derivano. Di conseguenza gli editori sono spesso riluttanti nell'accettazione di articoli di questo genere [19] [20].
- *Autori importanti e gruppi numerosi*: gli autori cercano di pubblicare su Journal importanti, ma è vero anche il contrario: i Journal cercano di pubblicare materiale di autori rinomati nel loro campo in modo da trarne vantaggio per l'attenzione che attireranno in futuro, e spesso vengono preferiti i loro articoli a prescindere dal contenuto [20]. Analogamente si prediligono spesso articoli con grande numero di coautori per aumentare le probabilità che venga citato da ciascuno dei suoi coautori in futuro [20].

3.2.2 Source Normalized Impact per Paper (SNIP) [21]

Si tratta di un interessante variazione dell'IF proposta da un gruppo di ricerca della Leiden University per cercare di compensare il problema legato alla varianza rispetto al campo scientifico (subject field). Tuttavia a differenza dell'IF, insiste sul database di Scopus ed è liberamente accessibile online.

Papers: vengono considerati solo articoli, conference proceedings e revisioni come papers a tutti gli effetti.

- + Citazioni gratuite a documenti non citabili e auto-citazioni degli editoriali non sono quindi presi in considerazione.
- Le auto-citazioni da papers normali vengono comunque contate.
- Non vi sono poi normalizzazioni legate alle revisioni, quindi lo SNIP risulterà leggermente più alto per Journal con tante revisioni, come per l'IF.

Citation potential (potenziale di citazione): indica quanto frequentemente i papers in un campo citano altri paper. E' definito come il numero medio di riferimenti citati in un paper nel campo. Se un journal j contiene m articoli $[a_{1j}, \dots, a_{mj}]$, e r_{ij} è il numero di riferimenti citati in a_{ij} , il citation potential R_{ij} nel campo del journal risulta:

$$R_j = \frac{\sum_{i=1}^m r_{ij}}{m}$$

Subject field (campo) del Journal: Il campo di un journal in un certo anno è definito come l'insieme di papers pubblicati in quell'anno che citino almeno un paper di quel journal nei dieci anni precedenti.

- + La delimitazione di un campo non dipende quindi da qualche categorizzazione definita a priori, ma interamente basata sulle relazioni delle citazioni.
- + Un campo può essere definito precisamente anche quando journal generali o multidisciplinari che coprono diversi campi scientifici giocano un ruolo importante in esso (Nature, Science,...).
- + Corregge le differenze nelle abitudini di citazione tra campi diversi, specialmente la frequenza a cui un autore cita altri papers e la rapidità della maturazione dell'impatto delle citazioni.

- + Corregge le differenze nella copertura del database tra campi: minor copertura del database porta a citation potentials più alti, che tendono a portare a uno SNIP più alto rispetto all'IF.
- + Non corregge solo le differenze tra categorie di journals, ma anche tra journal che coprono diversi argomenti, approcci o ricerche all'interno della stessa categoria di journal.
- Papers appartenenti formalmente a un campo, ma che non citano un journal, non sono inclusi per definizione nel campo di quel journal e questo potrebbe causare una parzialità (nonostante la finestra di 10 anni è usata per ridurre questo effetto).

Database citation potential (DCP): Per considerare la misura in cui un database copra un campo, non conta il numero totale di riferimenti citati nei paper di un campo, ma il numero di riferimenti citati che siano pubblicati in journals elaborati dal database. Se r_{ij}^{DB} è il numero di riferimenti citati contenuti in un paper i e pubblicati in fonti elaborate dal database, il citation potential dei journals sorgenti processati dal R_j^{DB} (DCP) è:

$$R_j^{DB} = \frac{\sum_{i=1}^m r_{ij}^{DB}}{m}$$

Oppure, se f_j denota per il campo del journal J la frazione di riferimenti citati che siano pubblicati in fonti elaborate dal database:

$$R_j^{DB} = f_j \cdot R_j$$

Raw impact per paper (RIP): il numero medio di citazioni per papers pubblicati in un journal (non è altro che l'IF). Se C_j denota il numero di citazioni in un particolare anno a papers pubblicati in J nei 3 anni precedenti, e A_j il numero di articoli pubblicati in quegli stessi anni, il RIP di j è:

$$RIP_j = \frac{C_j}{A_j}$$

Relative DCP (RDCP): come fattore di normalizzazione viene usato il journal mediano nel database, in termini di DCP del suo campo. Il DCP di ogni journal è diviso per il DCP di tale mediano. Se $M^{DB} = \text{mediano in } \{ R_j^{DB} \}$ è il DCP mediano tra i campi di tutti journals sorgente nel database:

$$RDCP_j = \frac{R_j^{DB}}{M^{DB}}$$

SNIP: Infine, lo SNIP per un journal J può essere definito:

$$SNIP_j = \frac{RIP_j}{RDCP_j}$$

Altri risultati riportati dai propositori dell'indicatore:

- Journals che coprono argomenti emergenti tendono ad avere un maggiore citation potential di journals che pubblicano papers su argomenti più 'classici' [21].
- Journals che pubblicano lettere o brevi comunicazioni tendono a mostrare un maggiore citation potential di journals 'normali'. Apparentemente, gli autori che citano journal di lettere sono più concentrati sulla letterature recente (1-3 anni) [21].

3.2.3 Y-Factor [22]

A cavallo tra le due categorie sopracitate di indicatori relativi ai journals, questo cerca di combinare il canonico IF con un valore calcolato seguendo l'algoritmo di PageRank [2], nella speranza di unire in un unico indicatore entrambi i fattori di popolarità e di prestigio che costituiscono lo status di un journal. Nonostante rispetto ai due descritti in seguito non abbia trovato una applicazione pratica in diffusi tools online, fornisce una buona introduzione agli indicatori basati su tale concetto. Anzitutto si definiscono un insieme di journals V , un insieme di archi direzionati $E \subseteq V^2$ e una funzione $W(v_i, v_j) \rightarrow N^+$ che mappi ciascun arco tra due journals v_i and v_j a una frequenza di citazione intera e positiva.

PageRank:

Nell'ambito di pagine web, l'equazione che governa il processo iterativo di trasferimento dei valori di PageRank da una pagina web a un'altra è la seguente:

$$PR(v_i) = \frac{(1 - \lambda)}{N} + \lambda \sum_j PR(v_j) \times \frac{1}{O(v_j)}$$

- Pagine v_j si collegano alla pagina ricevente v_i e ognuna trasferisce una proporzione del proprio PageRank, denotato da $PR(v_j)$, a v_i .

- I valori di PR vengono trasferiti equamente sugli archi uscenti, e quindi sono normalizzati dal numero di archi uscenti da v_j , denotato da $O(v_j)$.
- $\lambda \in [0,1]$ rappresenta l'attenuazione dei valori di prestigio trasferiti.
- $\frac{(1-\lambda)}{N}$ rappresenta la minima quantità di prestigio assegnata a ogni pagina.

PageRank pesato:

Nella Rete di Citazioni dei Journal, tuttavia, non tutti gli archi hanno ugual peso: alcuni journals sono connessi da più citazioni di altri. Per questo definiamo una proporzione di propagazione $w(v_j, v_i)$ tra i journals v_i e v_j normalizzando i pesi dei collegamenti uscenti da un particolare v_j come segue:

$$w(v_j, v_i) = \frac{W(v_j, v_i)}{\sum_k W(v_j, v_k)}$$

Per ogni particolare journal v_j , tutti gli archi (v_j, v_i) ora sommano a uno e possono essere usati per determinare la frazione di PageRank da trasferire. Quindi otteniamo ora l'equazione dei PageRank pesati per il journal v_i come segue:

$$PR_w(v_i) = \frac{(1-\lambda)}{N} + \lambda \sum_j PR_w(v_j) \times w(v_j, v_i)$$

Di fatto, la equa distribuzione dei valori di PR data dal fattore $1/O(v_j)$, è stata sostituita dalla proporzione di propagazione $w(v_j, v_i)$ permettendo dunque di poterlo calcolare anche per la Rete di Citazione dei Journal. Infine, definiamo il Y-Factor come:

$$Y(v_j) = IF(v_j) \times PR_w(v_j)$$

Journals con un alto punteggio di Y-Factor, avranno a loro volta un alto punteggio in IF, PR pesato o entrambi.

Altri risultati riportati dai propositori dell'indicatore:

- Si osserva che journals con il più alto IF sono largamente posizionati nell'area della medicina, con una grande presenza di journals di revisioni. Questo conferma la caratterizzazione dell'IF come una metrica orientata alla popolarità, dal momento che questo tipo di journal pubblica tipicamente materiale di background che molto è probabile venga citato spesso [22].

- Globalmente, la lista secondo il PageRank pesato mostra una maggiore variazione di discipline accademiche, e molti di quelli in cima alla classifica sono generalmente considerati journals di alto prestigio [22].
- Nonostante le forti discrepanze nella top ten delle due metriche, si rivela una significativa correlazione generale, confermata da un coefficiente di Spearman di $\rho=0.61$, $p<0.01$ tra le due [22].

3.2.4 Eigen Factor e Article Influence Score [4]

Questo è probabilmente il più popolare, attualmente, tra gli indicatori che seguono un procedimento affine al PageRank. Insiste sul database del Thomson Reuters Journal Citation Report (JCR). Definendo la matrice di cross-citazione \mathbf{Z} , sarà:

Z_{ij} = citazioni del journal j nell'anno N ad articoli pubblicati nel journal i nel periodo $[N-5, N-1]$ (cinque anni precedenti).

- + Nel farlo, tutti gli elementi sulla diagonale vengono posti a 0. Questo passaggio omette le auto-citazioni, così che si possano evitare i journals che tendono ad abusare di questa pratica. D'altro canto rimuoverle del tutto può svantaggiare journals che coprono piccole nicchie.
- L'utilizzo di una finestra di 5 anni da un lato riduce la variabilità statistica, ma dall'altro svantaggia i journals più nuovi (con età minore di 5 anni).

Viene poi normalizzata Z per le somme delle colonne (i.e. per il numero totale di citazioni uscenti da ogni journal) così da ottenere la matrice \mathbf{H} stocastica per colonne:

$$H_{ij} = \frac{Z_{ij}}{\sum_k Z_{kj}}$$

Vettore degli articoli a: In tale vettore, l'elemento a_i è il numero di articoli pubblicati dal journal i nella finestra di 5 anni, diviso per il numero totale di articoli pubblicati da tutti i journals negli stessi 5 anni. Di conseguenza gli elementi di a sommano a 1, e l' i -esimo elemento rappresenta la frazione di tutti gli articoli pubblicati proveniente dal journal i .

Dangling nodes: sono i journals che non citano nessun altro journal. Ne consegue, che ogni colonna di \mathbf{H} che contiene tutti 0 è un dangling node. Definiamo quindi la matrice \mathbf{H}' :

$H' = H$ modificata, ottenuta sostituendo tutti i dangling nodes col vettore a .

- Corrisponde a un cammino casuale sulla rete delle citazioni con i dangling nodes sostituiti da nodi totalmente connessi.

$A = a \cdot e^T$ è una matrice con colonne identiche, ciascuna uguale al vettore degli articoli a .

Si definisce quindi la matrice stocastica P come segue:

$$P = \alpha H' + (1 - \alpha)A$$

Corrisponde al processo di Markov che con probabilità α segue un cammino casuale sulla Rete di Citazioni dei Journals, mentre con probabilità $(1 - \alpha)$ si “teletrasporta” a un journal casuale, in proporzione al numero di articoli pubblicato da ogni journal.

- Il teletrasporto avviene con probabilità proporzionale al numero di articoli pubblicati dal journal così da evitare di sopravvalutare l’influenza di piccoli journals e di sottovalutare quella di journals più grandi.
- Aggiungere la probabilità di teletrasporto $(1 - \alpha)$ assicura che la matrice sia irriducibile e aperiodica, garantendo dunque che abbia un unico autovettore dominante secondo il teorema di Perron-Frobenius.
- Inoltre, senza tale probabilità di teletrasporto, le classifiche potrebbero risultare inaffidabili e volatili quando alcuni componenti sono connessi in modo estremamente sparso.
- Similmente, il teletrasporto previene che il sistema rimanga intrappolato in piccole aggregazioni di nodi quasi dangling. Se piccoli gruppi di journals vengono occasionalmente citati da fuori, ma raramente citano all’esterno di tale gruppo, il processo di Markov identificato da H' potrebbe rimanere intrappolato in tale zona per lungo tempo, di fatto sopravvalutando i journals in esso.

Influence vector π^* : l’autovettore dominante di P . Questo fornisce i pesi che verranno usati nella ponderazione dei valori di citazione; sotto l’interpretazione di processo stocastico il vettore π^* corrisponde alla frazione di tempo passata in stato stazionario in ogni journal rappresentato in P .

Infine, il punteggio di EigenFactor EF_i per un journal i è definito come la percentuale di citazioni pesate totali che i riceve da tutte le sorgenti, ovvero

$$EF = 100 \frac{H\pi^*}{\sum_i [H\pi^*]_i}$$

(Notare l'utilizzo della matrice **H** e non della versione modificata **H'**)

Mentre il punteggio di Article Influence AI_i per il journal i è una misura dell'influenza di citazione per articolo del journal i , ed è definito:

$$AI_i = 0.01 \frac{EF_i}{a_i}$$

- L'EF dà una misura dell'influenza totale che il journal fornisce, mentre l'IF fornisce una misura dell'influenza per articolo per un certo journal. Di conseguenza l'AI è stato definito per dare un indicatore comparabile all'IF.
- Il fatto che l'EF a differenza dell'IF non presenti un denominatore lo rende sensibile al numero di elementi citabili. In altre parole journals con un basso numero di articoli è probabile che abbiano un minor EF [23].

3.2.5 SCImago Journal Rank (SJR) [24]

Ulteriore processo iterativo simil-PageRank che ha acquisito una certa popolarità negli ultimi anni. È stato sviluppato da SCImago un gruppo di ricerca dell'università di Granada, ed è liberamente accessibile online. A differenza dell'EF, tuttavia, insiste sul database di Scopus invece che su quello dell'ISI, quindi i due indici presentano differenze nella copertura globale di journals. Si definiscono:

SJR_i - Scimago Journal Rank del journal i .

C_{ji} - Citazioni dal journal j al journal i .

C_j - Numero di riferimenti del journal j .

d - Costante, di solito 0.85.

e - Costante, di solito 0.10.

N - Numero di Journals

Art_j - Numero di Articoli del journal j

Dangling Nodes – come per l'EF, journals dell'universo che non hanno riferimenti a nessun altro journal dell'universo.

$$\begin{aligned}
SRJ_i = & \underbrace{\frac{(1-d-e)}{N}}_1 + e \underbrace{\frac{Art_i}{\sum_{j=1}^N Art_j}}_2 + d \underbrace{\sum_{j=1}^N \frac{C_{ji} + SJR_j}{C_j} \cdot \frac{1 - (\sum_{k \in \{Dangling-nodes\}} SJR_k)}{\sum_{h=1}^N \sum_{k=1}^N \frac{C_{kh} + SJR_k}{C_k}}}_3 \\
& + d \underbrace{\left[\sum_{k \in \{Dangling-nodes\}} SJR_k \right]}_4 \cdot \frac{Art_i}{\sum_{j=1}^N Art_j} \\
SJRQ_i = & \frac{SJR_i}{Art_i}
\end{aligned}$$

Analizzandone i vari addendi:

1. $\frac{(1-d-e)}{N}$: è il minimo prestigio assegnato a tutti.
2. $e \frac{Art_i}{\sum_{j=1}^N Art_j}$: Prestigio che ottiene un journal sulla base del numero di articoli pubblicati nella finestra di 3 anni.

1 e 2 sono costanti in tutte le iterazioni e insieme formano il minimo prestigio che ottiene ciascuna rivista.

3. $d \sum_{j=1}^N \underbrace{\frac{C_{ji} + SJR_j}{C_j}}_a \cdot \underbrace{\frac{1 - (\sum_{k \in \{Dangling-nodes\}} SJR_k)}{\sum_{h=1}^N \sum_{k=1}^N \frac{C_{kh} + SJR_k}{C_k}}}_b$: il prestigio ottenuto dai journals che

citano quello considerato.

- a. $\frac{C_{ji} + SJR_j}{C_j}$: la percentuale di prestigio che un journal X trasferisce a Y è costante in tutte le iterazioni, e dipende dal (n° di connessioni da X a Y) / (n° di connessioni di X), oltre al valore di JR della iterazione precedente.
- b. Tale correttore va aggiunto a causa di un problema: il numero di riferimenti di un journal X a tutti gli altri journal dell'universo considerato è più piccolo del numero totale di riferimenti, il che implica che parte del suo prestigio non viene distribuito. Di conseguenza il sistema non convergerebbe in assenza di un correttore che risolva il problema. Questo fattore è usato per distribuire tra i journals menzionati nell'universo il prestigio corrispondente alle citazioni che escono da quell'universo. Prestigio che è distribuito proporzionalmente a quello ottenuto dalle citazioni uscenti.

4. $d \left[\underbrace{\sum_{k \in \{Dangling-nodes\}} SJR_k}_a \right] \cdot \underbrace{\frac{Art_i}{\sum_{j=1}^N Art_j}}_b$: il prestigio che ottiene un journal di dangling nodes.
- a. $\sum_{k \in \{Dangling-nodes\}} SJR_k$: la quantità di prestigio da distribuire è la somma del prestigio di tutti i dangling nodes nella iterazione precedente.
- b. $\frac{Art_i}{\sum_{j=1}^N Art_j}$: il prestigio che un nodo concreto riceve è direttamente proporzionale al numero di articoli che ha pubblicato.

Il processo iterativo che viene eseguito è il seguente:

1. *Assegnamento dei valori iniziali di SJR*: poiché ogni passo del processo iterativo coinvolge il valore dell'iterazione precedente è necessario assegnare un valore iniziale predefinito ad ogni journal. Tuttavia, dato che il processo converge, tali valori non influenzano il risultato finale ma solo il numero di iterazioni necessarie a raggiungerlo.
2. *Processo iterativo di calcolo*: viene ripetuto il calcolo dello SJR per ogni journal finché la differenza dei valori tra due iterazioni successive è minore di un limite prefissato.
3. *Calcolo dello SJRQ*: come per l'EF, lo SJR è un indicatore di prestigio globale di un journal. Di conseguenza dividerlo per il numero di articoli pubblicati da tale journal nei 3 anni fornisce un indicatore paragonabile all'IF. Inoltre permette di valutare il prestigio di un journal senza dover tenere in conto fattori come la frequenza di un journal, il numero di articoli pubblicati, etc.

3.2.6 H-index [5]

La valutazione scientifica di un individuo è sempre stato un problema. Poiché il lavoro di uno scienziato è definito dall'insieme delle sue pubblicazioni, sembra ragionevole usarlo per determinarne il suo valore. Tuttavia il semplice numero di pubblicazioni misura solamente la quantità e non l'impatto, mentre il n° totale di citazioni ricevute misura solamente l'impatto e può essere gonfiato da poche pubblicazioni estremamente citate senza premiare la produttività. A tale scopo nel 2005 Hirsch [5] propose l'H-index come metrica che possa coinvolgere entrambi gli aspetti delineati. È definito come segue: uno scienziato ha un indice h se h delle sue pubblicazioni N hanno ricevuto almeno h citazioni ciascuna, mentre N-h pubblicazioni

hanno ottenuto meno di h citazioni ciascuna. A titolo di esempio, quindi, un h-index pari a 10 implica che lo scienziato ha pubblicato almeno 10 articoli e ciascuno di essi ha avuto almeno 10 citazioni.

Nonostante lo scopo originale dell'indicatore fosse la valutazione dell'impatto scientifico individuale, si può applicare a un qualunque gruppo di individui e, volendo anche a un intero journal. Difatti una applicazione ai journal è stata proposta [25]. In tale contesto l'indice non dovrebbe essere calcolato in tutta la finestra di vita dell'individuo (in questo caso il journal) ma per un periodo definito così da seguire il comportamento degli altri indici in tale contesto. Tale applicazione ha portato a portato ai seguenti risultati:

- Ridimensionata la posizione nella classifica relativa all'h-index di journals di revisioni, poiché pubblicano in media un minor numero di papers.
- Anch'esso mostra una notevole dipendenza dal campo scientifico, a causa delle diverse abitudini di ciascuna disciplina. Di conseguenza anche per esso non andrebbero paragonati journals di diversi campi.

3.2.6.1 Vantaggi, svantaggi e critiche

Rispetto ai precedenti indicatori presenti per la valutazione individuale ha senza dubbio dei vantaggi:

- Unisce in un'unica misura i fattori di quantità (produttività) e qualità (impatto) di un ricercatore.
- È una misura cumulativa dell'output di ricerca, poiché al crescere del n° di pubblicazioni può solo crescere.
- A differenza del numero di totale di citazioni non è sensibile a una o poche pubblicazioni estremamente citate.

Tale indice ha tuttavia raccolto anche innumerevoli critiche negli ultimi anni. Le più rilevanti sono le seguenti:

- *Auto-citazioni*: il calcolo dell'H-index include le auto-citazioni, ed essendo proporzionale anche al numero di citazioni può indurre cambiamenti nei comportamenti di pubblicazione degli scienziati [26] [27].

- *Assenza di semantica*: papers che includano errori estremamente gravi e grossolani potrebbero venire largamente citati come critica, ma avendo di fatto un impatto positivo sull'indice di quello specifico autore [28]. In realtà questo è un problema che sussiste per tutti gli indicatori attualmente esistenti che siano basati sul numero di citazioni, e si basa sulla difficoltà di poter discernere in modo automatico e affidabile quali siano effettivamente citazioni e quali critiche.
- *Monotonia*: poiché l'h-index può solo crescere non permette di distinguere tra scienziati attivi o scienziati con lunghi periodi di inattività (o addirittura ritirati, o peggio ancora deceduti). A sua volta questo mette in posizione di svantaggio i giovani ricercatori rispetto ai più anziani, per il semplice fatto che abbiano scritto un minor numero di pubblicazioni a causa della loro età [29].
- *Dipendenza dal campo scientifico*: come per l'IF, schemi e comportamenti di citazione tipici di specifiche discipline influiscono notevolmente sull'h-index. Tali comportamenti possono dipendere dal numero medio di riferimenti in un articolo di quel campo, numero medio di papers prodotti, la dimensione (n° di scienziati) del campo o l'attrattività dell'area di ricerca. Ci si può quindi aspettare un H-index più alto in certe discipline piuttosto che in altre [29]. Di conseguenza non può essere usato per paragonare autori appartenenti a diversi campi scientifici [30].
- *Insensibilità a papers altamente citati*: Papers altamente citati sono importanti per la determinazione dell'h-index, ma una volta selezionati nei top h, è indifferente il numero di citazioni che ricevono. Nonostante sia una proprietà cercata proprio per evitare le sopravvalutazioni derivanti del semplice numero di citazioni totali, finisce con l'essere troppo limitante [31]. Anche qui, autori giovani ma molto influenti avranno pochi papers molto citati, ponendoli in posizione di svantaggio.
- *Coautori e ordinamento*: l'h-index non tiene conto del numero di autori di un articolo, né del loro ordinamento, mentre hanno una specifica rilevanza nella determinazione della quota con la quale ciascuno di essi ha contribuito alla produzione dello stesso.

3.2.6.2 Varianti

Le critiche volte all'h-index hanno portato fioritura di molteplici varianti proposte negli ultimi anni. Tra le tante, le principali:

- G-index [32]: uno dei miglioramenti più interessanti dell'h-index. È così definito: posti gli articoli in ordine decrescente di numero di citazioni, l'indice è il più alto valore (in classifica) tale per cui i top g articoli abbiano, tutti insieme (sommati), almeno g^2 citazioni. È provato che tale valore sia unico e sia sempre maggiore o uguale all'h-index [33]. Rispetto all'h-index i principali vantaggi sono dati dal fatto che tenga conto del peso degli articoli più citati e che il numero di documenti non sia un limite dell'indicatore [32]. Si riporta di seguito una tabella (Tab. 1) che chiarisce ulteriormente il concetto del G-index, ponendo a confronto due ipotetici autori dal contributo scientifico evidentemente diverso, ma per i quali i rispettivi h-index risulterebbero (ingiustamente) identici.

Paper	Citation Author A	Rank ² A	Sum Citations A	Citation Author B	Rank ² B	Sum Citations B
1	30	1	30	10	1	10
2	20	4	50	6	4	16
3	15	9	65	5	9	21
4	6	16	71	4	16	25
5	4	25	75	2	25	27
6	2	36	77	1	36	28
7	1	49	78	0	49	28
8	1	64	79	0	64	28
9	1	81	80	0	81	28
10	0	100	80	0	100	28
	h-index A=4		g-index A=9	h-index B=4	0	g-index A=5

Tab. 1 Confronto G-index h-index [32]

- m quotient [5]: poiché l'h-index dipende dalla lunghezza della carriera di un ricercatore, già lo stesso Hirsch aveva proposto di dividere l'indicatore per il numero di anni a partire dalla prima pubblicazione, definendolo m quotient.
- a-index [34]: è definito come il numero medio di citazioni dei papers nell'Hirsch core, ovvero nei top h articoli di un autore.
- m-index [35]: variazione dell'a-index. Dato che la distribuzione delle citazioni è asimmetrica, la mediana e non la media simmetrica dovrebbe essere presa come misura di tendenza centrale tra gli articoli nell'Hirsch core.

- R-index [36]: gli autori hanno notato che l'a-index punisce gli scienziati migliori (secondo l'h-index) poiché prevede una divisione per h. Quindi suggeriscono piuttosto di prendere la radice quadrata della somma delle citazioni degli articoli nell'Hirsch core per il calcolo dell'indice. Tuttavia essendo una misura dell'intensità di citazione all'interno dell'Hirsch core risulta molto sensibile a singoli papers molto citati [35].
- AR-index [36]: a sua volta un adattamento dell'r-index. È definito come la radice della somma del numero medio di citazioni all'anno degli articoli nell'Hirsch core. Fa quindi uso anche dell'età di una pubblicazione, ed è una misura che può aumentare o diminuire nel tempo.
- Hm-index [37]: recente proposta per tenere conto della presenza di multipli co-autori, basata su un conteggio frazionalizzato in base al numero di coautori.

3.3 Search Engines & Online Archives

A seguito dell'introduzione del World Wide Web, la documentazione scientifica disponibile online è cresciuta velocemente. Con tale quantità di dati in continua crescita, il recupero delle informazioni è diventato un aspetto rilevante nell'ambiente accademico. Sono quindi nati diversi motori di ricerca e archivi online, con lo scopo di cercare le informazioni desiderate in uno o più database e presentarle ordinate secondo criteri di rilevanza rispetto allo specifico input inserito dall'utente. Gli indicatori bibliometrici prima definiti giocano un ruolo importante in questo processo poiché forniscono delle misure sul livello di precisione con cui le informazioni vengono presentate all'utente.

Esistono molteplici motori di ricerca 'verticali', ovvero legati a un determinato campo scientifico (e.g. PubMed per la medicina, ACM, IEEE etc. per il campo ingegneristico...) per il quale forniscono materiale specifico, ma vista la necessità di maggiore interdisciplinarietà nel mondo accademico si sono resi disponibili motori di ricerca di carattere generale in grado di effettuare ricerche attraverso tutti i diversi archivi online verticali (e più in generale nel web). Fino al 2004 il principale motore di ricerca di questo tipo è stato Web of Science (WoS, ora Web of Knowledge), ma negli ultimi anni si sono aggiunti due ulteriori protagonisti, ovvero Google Scholar e Scopus. Diversi lavori hanno quindi cercato di identificare i vantaggi e gli svantaggi di ciascuno di essi.

3.3.1 Google Scholar

Il motore indicizza il testo completo delle pubblicazioni attraverso diversi formati di documenti e discipline. La rilevanza che determina l'ordinamento dei risultati dipende dalla classifica degli autori, dal numero di riferimenti collegati alla pubblicazione e la loro rilevanza nelle altre letterature accademiche, e dalla posizione in classifica del journal in cui appare la pubblicazione.

Vantaggi:

- *Documenti sul web*: poiché molti autori disseminano su altri canali (e.g. sul loro sito web, social network accademici etc.) i loro lavori prima dell'effettiva pubblicazione in un Journal, i ricercatori possono trovare informazioni più attuali di quelle presenti in database commerciali [38].
- *Numero di risultati*: Simili ricerche per parola chiave restituiscono spesso meno risultati in Scopus e WoS (meno di cento) rispetto a Google Scholar (diverse centinaia) [39].
- *Tipi di documenti*: Scholar e Scopus permettono di identificare un considerevole numero di citazioni in tipi di documenti non coperti da WoS [39] [40].
- *Copertura Geografica e linguistica*: ha una notevole copertura in tal senso [41].

Svantaggi:

- *Materiale inappropriato*: Scholar include nei risultati note amministrative, informazioni bibliotecarie, appunti di studenti, etc., che non sono considerabili esattamente letteratura scientifica [38].
- *Assenza di trasparenza*: Google non fornisce informazioni sulla sua copertura. Non è chiaro a quali archivi degli editori abbia accesso per la ricerca e quali no [39] [40].
- *Ridondanza*: capita spesso che ottenga dati ridondanti e li presenti come risultati differenti. In alcuni casi la stessa pubblicazione può essere presentata più di una volta, e questo ha un impatto anche sul calcolo degli indicatori bibliometrici legati alla ricerca [42] [41].

- *Errori booleani*: spesso vi sono inspiegabili inconsistenze relative al numero di risultati restituiti quando vengono applicati operatori logici (e.g. 'the OR a' produce meno risultati di 'the') [43] [41] [44].
- *Errata data di pubblicazione*: sono stati evidenziati errori nella rilevazione della data di pubblicazione dei documenti, almeno per l'American Physical Society [42], spesso l'ISSN o i numeri di pagina vengono interpretati come anno [41].
- *Iniziali autore*: la ricerca di un autore con solo un'iniziale per il nome, produce i risultati di tutti gli autori con lo stesso cognome e quella specifica lettera in una qualunque delle iniziali [42]. Non c'è come in WoS o Scopus una funzionalità che permetta di distinguere autori con lo stesso nome e iniziale [45]. Anche questo ha una influenza sul calcolo dell'h-index degli autori.
- *Autori errati o fittizi*: spesso Scholar commette errori nell'estrapolare gli autori di un articolo, finendo per confondere due autori dai nomi simili, o creare addirittura autori fittizi. Noti esempi di autori fittizi sono 'F Password' e 'M Profile'. Questo genere di errore finisce per assegnare citazioni ad autori che non lo meritano (o che addirittura non esistono proprio), sottraendole a chi ne avrebbe diritto [29].

3.3.2 Scopus

È un database di abstract (estratti, sommari) e indicizzazione con collegamenti ad articoli completi, prodotto da Elsevier Co. Dispone di funzionalità intelligenti che permettono di tracciare, analizzare e visualizzare le ricerche.

Vantaggi:

- *Abstract*: a differenza di Scholar e WoS, una percentuale maggiore di records di Scopus include l'abstract dell'articolo [46].
- *Copertura*: Scopus sostiene di avere una copertura mondiale, con più della metà di contenuti provenienti da Europa, America Latina e la regione Atlantica-Pacifica, mentre WoS copre principalmente Nord America ed Europa Occidentale [47]. Fornisce anche una copertura decisamente migliore di WoS nell'ambito della letteratura sulla human-computer interaction [48].

Svantaggi:

- *Discipline umanistiche*: non copre fonti di discipline artistiche e umanistiche, e ha una copertura modesta per quanto riguarda le scienze sociali [46].

3.3.3 Web of Science (Web of Knowledge)

È un servizio online (a sottoscrizione) di indicizzazione delle citazioni scientifiche gestito dalla Thomson Reuters, che fornisce una funzionalità di ricerca globale. Dà accesso a molteplici databases relativi a ricerca interdisciplinare, e permette una esplorazione in profondità di sotto-campi specializzati all'interno di una disciplina accademico-scientifica. Fornisce strumenti per cercare, tracciare, misurare e collaborare nelle scienze, scienze sociali, arte e discipline umanistiche.

Vantaggi:

- **Periodo coperto**: un vantaggio riguarda la copertura che si espande fino ad articoli del 1945, mentre per Scopus solo fino al 1966 [49]. In generale risulta migliore la copertura di materiale più vecchio rispetto a Scopus [50].

Svantaggi:

- **Parole simili**: rispetto a Scopus e Scholar, WoS non effettua la ricerca anche per comuni variazioni delle parole passate in input [50].
- **Costi**: WoS e Scopus sono prodotti relativamente costosi, quindi potrebbe risultare troppo oneroso accedere ad entrambi, mentre Scholar è gratuito [47].

3.3.4 Analisi comparativa

Il lavoro precedente di Penagos [51] ha evidenziato le diverse funzionalità offerte da un grande numero di motori di ricerca e social network accademici. In generale sembra che i motori di ricerca verticali siano di qualità mediamente più alta (in termini di funzionalità offerte), mentre tra quelli di carattere generale, Scopus è quello che offre funzionalità più sofisticate in assoluto. Ha anche effettuato una analisi delle visite ricevute da essi usando il tool Alexa.com, ponendoli in una classifica globale in base a tale parametro. Per quanto Alexa.com sembra misurare il numero di visite al dominio principale (e.g. google.com per scholar.google.com) sono emersi alcuni rilievi interessanti:

Engine	Global Rank
--------	-------------

Google Scholar	1
Academic Search	43
Microsoft Academic Search	43
PubMed	279
VET-Bib	775
Astrophysics Data System	1165
sciencedirect	1183
CiteSeerX	2140
Behavioral and Brain Sciences	2156
IEEE Explore	2313

Tab. 2 Top ten Search Engine Global Rank [51]

MathGuide	1117263
ATLA Religion Database	1202426
Indian Citation Index	1337760
Safetylit	2303038
GoPubMed	3086143
MetaCrawler	4788672
Current Index to Statistics	5455192
Anthropological Index Online	6972868
Psycline	7206980
Food Science and Technology Abstracts	12768417

Tab. 3 Least Visited Search Engines [51]

Research Network	Global Rank
LinkID	12
Academia.edu	884
ResearchGate	1676
ArXiv	7512
Social Science Research Network	10815
CiteULike	13554
Mendely	20201
MyScienceWork	133328
viXra	450238

Tab. 4 Research Network Global Rank [51]

- I motori di ricerca in fondo alla classifica presentano funzionalità molto limitate rispetto alla top 10 della Tab. 2, quindi sembra che i ricercatori siano attratti maggiormente dai motori più potenti [51].

- L'unico motore di ricerca generale a comparire nella top 10 è Scholar, nonostante sia quello più criticato e con le debolezze maggiori.
- In particolare sorprende lo scarso utilizzo di Scopus, che presenta funzionalità molto più articolate di Scholar o WoS. Il fatto stesso che Scopus sia inferiore a molti motori di ricerca verticali fa pensare che le abitudini dei ricercatori siano specifiche della loro disciplina, o perché sono soddisfatti della loro letteratura di settore, o perché i motori interdisciplinari presentano meccaniche di ricerca di bassa qualità.
- In ogni caso, la classifica della maggior parte dei motori di ricerca risulta estremamente bassa a livello globale. A titolo di esempio, si è vista la posizione in classifica dei siti di alcune università. Risulta che l'Imperial College (<http://www3.imperial.ac.uk>) è classificato 19,881°, la London School of Economics (www.lse.ac.uk) 10,538° (sopra Scopus) e il Politecnico di Milano (www.polimi.it) 18,244°. In generale, sembra che gli accademici usino i motori di ricerca accademici meno di quanto gli studenti di una singola grande università cerchino sul sito del loro ateneo. Ovviamente gli studenti superano in numero gli accademici, ma resta un fatto sorprendente e mostra che per gli accademici la ricerca è più distribuita e/o meno prioritaria di quanto non sia per gli studenti la ricerca di informazioni. Oltretutto gli studenti si suppone facciano ricerche di letteratura durante il loro lavoro di tesi, e quindi contribuiscano in qualche misura all'utilizzo di tali motori.
- I social network accademici attirano una notevole quantità di visite, con Academia e ResearchGate più alte in classifica di molti motori di ricerca [51]. Sembrerebbe quindi che i ricercatori siano interessati a instaurare reti di cooperazione virtuali con i colleghi, anche se va tenuta in conto la possibilità (non remota) che buona parte delle visite siano riversate da risultati di Scholar a elementi di tali social network, piuttosto che da effettivi utenti registrati.

3.4 Conclusioni

L'analisi della letteratura ha portato a considerazioni molto interessanti. Come ci si aspetterebbe gli indicatori bibliometrici, in quanto metriche di valutazione personale, hanno ricevuto non poche critiche dal mondo scientifico. Si sono evidenziate le debolezze e cercate alternative che li migliorassero. Nonostante questo, tuttavia,

indicatori come l'IF e l'H-index risultano ancora affermati e il loro utilizzo è abbastanza consolidato.

Anche i motori di ricerca hanno raccolto molte critiche, ma di fatto restano uno strumento necessario alla consultazione della letteratura. Stupisce anzi che secondo i dati di Alexa.com, all'infuori di Scholar (poiché bisogna tenere conto del fatto che le visite probabilmente vengano contate verso il dominio principale Google), gli accademici sembrano fare molta meno ricerca in letteratura di quello che ci aspetterebbe, in particolar modo su motori di carattere generale che favoriscono l'interdisciplinarietà.

Tutte queste considerazioni, e più in generale il rilevamento delle effettive abitudini di ricerca degli accademici, necessitano di essere assodate e, nel capitolo seguente, viene descritta la metodologia adoperata a tale scopo.

4 Metodologia

4.1 Introduzione

Questo lavoro si inserisce nel progetto finanziato dalla Association of Information Technology, in collaborazione con la London School of Economics, il Birkbeck College University of London, la University of Oxford e il Politecnico di Milano, indirizzato a capire se un motore di ricerca accademica basato maggiormente sul concetto di “impatto” delle pubblicazioni possa essere utile e di interesse dei ricercatori, e una sua eventuale prototipazione se i risultati fossero positivi. In particolare il lavoro di questa tesi si concentra sulla prima parte di tale progetto, ovvero sull’analisi delle abitudini del mondo accademico per determinare le direzioni da prendere in questo senso.

Con la crescente massa di documentazione scientifica disponibile online, uno strumento di questo genere sembrerebbe essere necessario, per permettere di recuperare articoli più rilevanti in meno tempo. Infatti dalla nascita del web stiamo assistendo a un sovraccarico informativo anche nel mondo della letteratura scientifica [52] [53], e questo rende sempre più difficile e lungo il reperimento di informazioni interessanti [54].

Di conseguenza nel mondo scientifico si sono sviluppate negli anni metriche di “impatto” più quantitative, basate su indicatori bibliometrici. L’utilizzo di indicatori di questo genere per misurare la rilevanza di un articolo va incontro tuttavia a due problemi principali:

- 1) Questo genere di indicatori sono stati protagonisti di ampi criticismi in letteratura, che verranno ricapitolati nella sezione 4.2.
- 2) A prescindere dalle critiche, tali indicatori non vengono sfruttati appieno negli attuali motori di ricerca letteraria, non rendendo possibile una estesa parametrizzazione della ricerca su di essi. Problema che verrà esposto in maggiore dettaglio nella sezione 4.3.

Questo farà emergere dei dubbi sulla fattibilità di uno strumento di questo genere (sezione 4.4) e alla conseguente formulazione delle domande di ricerca di questo lavoro (sezione 4.5), le cui risposte potrebbero delineare dei comportamenti viziosi

nel mondo accademico, sui quali si sospetta abbia un'incidenza l'utilizzo degli strumenti esistenti (sezione 4.6).

4.2 Il concetto di impatto

Attualmente il concetto di "impatto" si è concretizzato sostanzialmente solo a livello bibliometrico, con la produzione di numerosi indicatori che si basano, in fin dei conti, solo sul numero di citazioni, per quanto articolate o meno possano essere le operazioni eseguite su di esso. Inoltre questi indicatori si concentrano sull'impatto di un autore o di una rivista (Journal), ignorando il contenuto. Dal punto di vista di un lettore, invece, sembrerebbe più interessante una metrica d'impatto basata sui contenuti e contestualizzata. La letteratura ha infatti raccolto negli anni una nutrita critica a questo genere di indicatori (soprattutto Impact Factor e H-index), ripercorsa nel capitolo di Stato dell'arte. In breve, le considerazioni più rilevanti inerenti l'Impact Factor riguardano:

- *Manipolazione*: lo specifico calcolo che definisce l'indicatore si presta a innumerevoli operazioni manipolatorie, poiché caratterizzato da imperfezioni teoriche.
 - Include completamente le autocitazioni del journal. Questo può indurre un comportamento vizioso che vede l'inserimento ingiustificato di citazioni nei propri articoli al solo scopo di aumentare l'indicatore [7] [8].
 - È incoerente nella selezione degli elementi presi in considerazione nel numeratore e nel denominatore [6]. Gli elementi marcati come "non citabili" (editoriali, news, lettere agli editori...) non vengono considerati nel denominatore, ma le citazioni provenienti da essi vengono conteggiate nel calcolo del numeratore. È palese come questo possa indurre una indiscriminata pubblicazione di editoriali, lettere etc. che includano volutamente un ampio numero di citazioni a vecchi articoli dello stesso journal (proprio grazie al fatto che le autocitazioni vengano incluse nel calcolo), dando un contributo consistente all'incremento del proprio IF.
 - Gli editori possono essere portati a prediligere la pubblicazione di determinati tipi di articoli, a prescindere dalla qualità dei loro contenuti, perché attirano in generale un maggior numero di citazioni. A titolo di

esempio potrebbero preferire la pubblicazione di revisioni (rispetto ad articoli originali) [13] [14], articoli che affrontino problemi teorici (piuttosto che pratici) [9] e articoli che trattino di argomenti “hot” (che rendono l’IF sensibile a mode e montature) [17]. Analogamente possono essere riluttanti nell’accettazione di articoli che in generale destano minore attenzione, come studi negativi o di conferma (rispetto all’originale da cui derivano) [19] [20].

- i Journal cercano di pubblicare materiale di autori rinomati nel loro campo in modo da trarne vantaggio per l’attenzione che attireranno in futuro, e spesso vengono preferiti i loro articoli a prescindere dal contenuto. Analogamente si prediligono spesso articoli con grande numero di coautori per aumentare le probabilità che venga citato di essi in futuro [20].
- *Parzialità per la lingua inglese*: c’è una preferenza nell’inclusione di Journal in lingua inglese [10] che porta inevitabilmente a bassi IF per quelli non anglofoni, dato che la maggior parte delle citazioni ad essi arriva da altri journal nella loro lingua [11] [12].
- *Campo scientifico*: l’IF esibisce una significativa variazione a seconda del campo scientifico considerato [6]. Comunità scientifiche più ristrette genereranno un minor numero di citazioni, oltre che a presentare abitudini di citazione del tutto specifiche di quel campo [15]. Per questi motivi non sono paragonabili gli IF di journal appartenenti a diversi campi scientifici.
- *Outliers e asimmetria*: singoli articoli che ottengono un gran numero di citazioni possono portare a variazioni considerevoli nell’IF di un Journal. Di conseguenza, spesso l’IF è influenzato da pochi articoli molto citati [16] e non risulta quindi rappresentativo di un articolo ‘medio’ di quel Journal. Non dovrebbe quindi essere usato per determinare il valore di singoli articoli o autori.
- *Riduzione di interdisciplinarietà*: emerge che i journals più alti in classifica secondo l’IF coprono una gamma meno estesa di discipline rispetto a quelli più in basso, quindi sembra essere dannoso per l’interdisciplinarietà [18].

L'h-index, nato per valutare l'impatto scientifico degli individui, dal canto suo sembrava apportare un contributo migliore, perché si prefiggeva di unire in un'unica misura gli aspetti quantitativi e qualitativi della produzione scientifica individuale. Anch'esso ha tuttavia ha presentato diverse problematiche:

- *Auto-citazioni*: anche per esso, non vi sono fattori di correzione delle autocitazioni, che potrebbero quindi indurre a pratiche viziose indirizzate esclusivamente ad accrescere il proprio indice [26] [27].
- *Assenza di semantica*: citazioni in critica di un paper hanno comunque un impatto positivo sull'indice di quello specifico autore, a causa dell'assenza di semantica nel conteggio delle citazioni [28]. In realtà questo è un problema che sussiste per tutti gli indicatori bibliometrici attualmente esistenti, ed è dovuto alla effettiva difficoltà di poter discernere in modo automatico e affidabile tra i diversi tipi di citazione.
- *Monotonia*: l'h-index può solo crescere, rendendo impossibile distinguere gli scienziati ritirati, con lunghi periodi di inattività o deceduti [29].
- *Dipendenza dal campo scientifico*: come l'IF, comportamenti di citazione e specificità tipiche di un determinato campo scientifico rendono inconsistente qualunque paragone tra autori appartenenti a diverse discipline [29] [30].
- *Insensibilità alle code*: l'h-index è volutamente molto robusto sia rispetto ai papers poco citati, sia rispetto a quelli molto citati [31]. Il problema è che lo è anche troppo. La presenza di papers altamente citati è necessaria alla formazione di un buon indice, ma non danno alcun contributo una volta selezionati tra i top h. Questo pone ancora una volta in svantaggio gli autori più giovani, poiché anche se fossero estremamente affermati, avrebbero articoli molto citati ma in numero ridotto semplicemente a causa della loro età.
- *Coautori e ordinamento*: l'h-index non tiene conto del numero di autori di un articolo, né del loro ordinamento, mentre hanno una specifica rilevanza nella determinazione della quota con la quale ciascuno di essi ha contribuito alla produzione dello stesso.

Il concetto di "impatto" sembra quindi un argomento di interesse degli accademici. Infatti nell'applicazione bibliometrica non ha solo portato a una vasta critica, ma anche a un grande numero di proposte alternative che cercassero di migliorare i

difetti evidenziati. Alcune di esse hanno visto applicazione, ma in numero relativamente ristretto rispetto alla grande quantità di proposte. Quindi l'utilizzo di IF e H-index risulta nonostante tutto ancora consolidato nella comunità scientifica, che sembra quindi poco propensa a un cambiamento in tal senso.

Dopotutto, quindi, un nuovo potenziale strumento di ricerca potrebbe ancora avere senso che basi i propri criteri di rilevanza e impatto sugli attuali indicatori bibliometrici, magari prediligendo le alternative più promettenti tra quelle proposte per ridurre il numero di difetti.

4.3 Impatto negli strumenti attuali

La necessità di una parametrizzazione della ricerca sulla base di un concetto di "impatto" non vede una concretizzazione negli attuali motori di ricerca letteraria. Questi prevedono un modello gerarchico di ricerca che è rigidamente suddiviso in discipline. Altri parametri di ricerca come citazioni o parole chiave sono ortogonali a questo modello, ma non vengono ricalcolate in base alla specifica ricerca di contenuto da parte di un autore. In Scopus ad esempio il concetto di campo scientifico è molto ampio (e.g. 'ingegneria', 'economia',...) e prevede che ogni Journal sia associato a uno o più campi, e gli autori a loro volta siano associati a determinati campi come conseguenza della loro associazione con i Journal in cui hanno delle pubblicazioni. Questo approccio gerarchico presenta diverse limitazioni (valide per tutti i motori di ricerca):

- L'impatto di un autore non può essere usato come criterio di ricerca.
- L'impatto di un autore/Journal non può essere ristretto a un determinato campo. Non è possibile ad esempio restringerne il calcolo in una specifica ricerca dell'utente (e.g. autori/journal maggiormente d'impatto in <sostenibilità e IT>, o negli ultimi 5 anni).
- Le citazioni in un certo campo non possono essere usate come criterio di ricerca.
- Le parole chiave non hanno una relazione con il campo scientifico. Non è possibile quindi ridefinire una propria idea di campo scientifico attraverso espressioni logiche che coinvolgano parole chiave, per poi ricalcolare gli indicatori in base a tale ricerca.

Permettere una ricerca più flessibile potrebbe aiutare a soddisfare il requisito di un utente in cerca di contenuti, che usi tutti gli indicatori, in particolare l'impatto, per evidenziare quelli più rilevanti. Questo supporterebbe un modello di ricerca più interdisciplinare (e cooperativo) rispetto a quello gerarchico corrente.

Tra tutti quelli esistenti, Scopus sembra essere il migliore, poiché è quello che maggiormente si avvicina a questi concetti, disponendo di funzionalità avanzate molto più sofisticate della concorrenza, ma che non colmano comunque esaustivamente le lacune evidenziate. Google Scholar tuttavia sembra quello più utilizzato in virtù della sua copertura apparentemente più ampia e del suo minor tasso di eliminazione dei duplicati che, come conseguenza, porta a stime più alte degli indicatori bibliometrici (sebbene questo sarebbe da considerare un difetto). Nonostante questo, stando alla letteratura sembra essere il motore di minore qualità rispetto a Scopus e WoS, presentando notevoli difetti tecnici:

- *Materiale inappropriato*: include materiale che non è considerabile propriamente letteratura scientifica (e.g. informazioni bibliotecarie, appunti di studenti, ...) [38].
- *Limitate funzionalità*: rispetto ai concorrenti dispone di meno possibilità per filtrare, ordinare etc. i risultati.
- *Assenza di trasparenza*: Google rende esplicite informazioni inerenti la sua copertura (a quali archivi editoriali acceda) [39] [40].
- *Ridondanza*: sembra essere quello con una minore eliminazione dei duplicati, che porta a stime in eccesso degli indicatori bibliometrici [42] [41].
- *Errori booleani*: presenta inspiegabili incongruenze nel numero di risultati presentati quando vengono applicati operatori logici alla ricerca [43] [41] [44].
- *Errata data di pubblicazione*: spesso interpreta ISSN, n° di pagina o altro come date di pubblicazione [41] [42].
- *Autori errati o fittizi*: spesso Scholar commette errori nell'estrapolare gli autori di un articolo [42] [45], finendo per confondere due autori dai nomi simili, o creare addirittura autori fittizi ('F Password' e 'M Profile'). Questo genere di errore finisce per assegnare citazioni ad autori che non lo meritano (o che addirittura non esistono proprio), sottraendole a chi ne avrebbe diritto [29].

Ancora una volta, la notevole proliferazione di studi sugli strumenti attualmente esistenti farebbe pensare che la qualità della ricerca sia di interesse degli accademici. Tuttavia la maggior parte delle critiche a riguardo si concentra sulle inconsistenze che tali strumenti presentano nella registrazione delle citazioni e nel conseguente calcolo degli indicatori bibliometrici. Sembrerebbe quindi un ulteriore segnale della scarsa propensione generale a voler cambiare gli attuali indicatori di impatto basati sul numero di citazioni, piuttosto che sul contenuto.

4.4 Fattibilità

La prototipazione, ma già solo la formulazione, di un eventuale strumento che migliori la ricerca letteraria basandosi maggiormente sui contenuti è un problema estremamente complesso. Questo genere di problema potrebbe rientrare nell'area della "ricerca semantica", ovvero il tentativo di migliorare la precisione della ricerca comprendendo le intenzioni dell'utente e il significato contestuale dei termini che compaiono nello spazio di ricerca, al fine di generare risultati più rilevanti. In generale un motore di ricerca semantico dovrebbe disporre di certe caratteristiche:

- Trend attuali: dovrebbe essere in grado di restituire risultati rilevanti in base a news e trend attuali nel momento della ricerca (e.g. "chi è il nuovo presidente del consiglio" dovrebbe restituire quello attuale).
- Posizione della ricerca: dovrebbe fornire risultati rilevanti in base al luogo in cui è avviata la ricerca (e.g. "qual è la temperatura" ricercato da Milano dovrebbe restituire risultati sulle temperature a Milano).
- Intento della ricerca: dovrebbe fornire risultati appropriati basandosi sulle intenzioni della ricerca piuttosto che sulle specifiche parole usate nell'interrogazione (e.g. "fa caldo e ho sete, cosa potrei prendere?" dovrebbe restituire qualunque genere di bevanda fredda ed escludere qualunque bevanda calda).
- Variazioni: dovrebbe fornire risultati per tutte le variazioni semantiche delle parole coinvolte nella ricerca, come plurali, tempi verbali, coniugazioni, etc.
- Sinonimi: dovrebbe fornire risultati per tutti i sinonimi delle parole usate, se hanno significati simili in quel contesto (e.g. "la più grande montagna" e "la più alta montagna" dovrebbero dare risultati simili visto che 'grande' e 'alta'

hanno significati simili in quel contesto, sebbene non vogliano dire la stessa cosa in generale).

- Interrogazioni generalizzate e specializzate: dovrebbe essere in grado di rilevare relazioni tra interrogazioni generalizzate e specializzate (e.g. un articolo sul “diabete” e uno sulla “salute” generale potrebbero entrambi combaciare in una ricerca sulla “salute”, anche se tale termine non compare esplicitamente in quello sul diabete).
- Rilevazione del concetto: dovrebbe rilevare il concetto più ampio legato a una ricerca (e.g. “problemi di traffico a Milano” dovrebbe restituire anche risultati legati a “malfunzionamento semafori” perché è una potenziale causa di “problemi di traffico”).
- Linguaggio naturale: dovrebbe essere in grado di interpretare interrogazioni poste in linguaggio naturale.
- Gruppi di parole: dovrebbe essere in grado di capire il diverso significato che acquisiscono determinate parole quando sono combinate in modo diverso.

Piuttosto che utilizzare algoritmi come il PageRank usano tecniche basate sulla semantica per stimare la rilevanza di un risultato. Per farlo è necessario passare attraverso una fase di disambiguazione, nella quale si cerca di determinare il significato di ogni parola, che a sua volta influenzerà il significato delle parole vicine fino al raggiungimento di una situazione di massima plausibilità e coerenza della frase. Questo necessita spesso un'ontologia: ovvero una specifica formale della concettualizzazione di un dominio di conoscenza condiviso (i.e. un vocabolario controllato che descrive oggetti e le relazioni tra di essi). Un'ontologia a sua volta può essere rappresentata da una struttura dati costituita da una rete semantica, ovvero una rete in cui ogni concetto lessicale corrisponde a un nodo (definito solitamente synset) che è a sua volta collegato ad altri nodi attraverso relazioni con struttura gerarchica e ereditaria, così che ogni concetto sia arricchito dalle caratteristiche e dal significato dei nodi vicini. Ciascun nodo aggrega un insieme di sinonimi che rappresentino lo stesso concetto lessicale e può contenere singole parole, parole composte (e.g. no-stop, abat-jour,...) o collocazioni (e.g. carta di credito).

Nel caso della letteratura scientifica si potrebbe ad esempio migliorare il concetto di “campo scientifico” in modo che sia più flessibile e parametrabile dall'utente, o si

potrebbe implementare un sistema di ricerca avanzata che esegua interrogazioni semantiche basate sul corpo degli articoli.

Dall'elenco delle caratteristiche attese si rende tuttavia palese che tali modelli richiedono un notevole dispendio di risorse, e prima di implementarli sarebbe quindi opportuno valutare se ne valga la pena, ovvero se fosse qualcosa che andrebbe incontro a un effettivo bisogno non corrisposto e un conseguente utilizzo.

Scopus stesso, che fornisce funzionalità molto complesse e in qualche modo vicine a questi concetti, sembrerebbe avere uno scarso utilizzo, mentre Google Scholar, strumento molto più limitato, sembra avere più successo [51]. Inoltre, sempre secondo i dati di Alexa.com [51], sembra che Scopus e molti altri motori ricevano globalmente meno visite persino di alcuni siti istituzionali, come quello della London School of Economics. In generale sembrerebbe quindi che il mondo accademico faccia poca ricerca online.

È necessario dunque esplorare in modo più approfondito le abitudini di ricerca degli accademici, per prendere delle decisioni sulle direzioni da seguire in un approccio di ricerca maggiormente basato sui contenuti. Nel corso del lavoro sono nati diversi sospetti per i quali vale la pena cercare risposte:

- I motori di ricerca non sono tutti uguali.
 - Tra quelli di carattere generale Scopus sembra il motore di miglior qualità, in virtù delle sue funzionalità più sofisticate rispetto alla concorrenza. Ma viene effettivamente usato di più? E come viene realmente valutata la sua efficacia? I dati di Alexa.com [51] sembrano indicare uno scarso utilizzo di questo strumento. Scholar riscontra effettivamente un maggior successo?
 - I motori di ricerca generali offrono già funzionalità di ricerca avanzata (a prescindere dalla diversa qualità nei vari concorrenti). Ma queste vengono realmente usate? E vengono ritenute efficaci?
 - I motori di ricerca verticali d'altro canto risultano bivalenti. Essendo specifici di un certo dominio funzionano già implicitamente da filtro dei risultati, quindi ci si aspetterebbe che venissero usati di più. Tuttavia sempre per la loro specificità inducono una minore interdisciplinarietà. I dati di Alexa.com [51] sembrano suggerire che anch'essi vengono usati

poco. Ma questo è vero anche nella pratica? E in tal caso, sarebbe a causa della loro ridotta propensione all'interdisciplinarietà?

- Sempre secondo i dati di Alexa.com [51] sembra stia emergendo un nuovo metodo di ricerca “social”, dato il numero di viste che attraggono i social network accademici. Un approccio che ponga una maggiore attenzione ai contenuti basandosi sulle relazioni sociali con altri ricercatori è realmente nelle abitudini del mondo accademico?
- Scholar rispetto alla concorrenza ha il vantaggio di abilitare un maggiore accesso non solo ad articoli accademici ma anche a documentazione di carattere pratico (brevetti, white papers delle compagnie, ...). Ma gli accademici sfruttano la cosa? Può essere effettivamente una delle motivazioni che portano a prediligere l'utilizzo di Scholar?
- Si dà per scontato che i motori di ricerca servano a trovare articoli all'interno di un topic, però potrebbero essere anche utilizzati anche al contrario, ovvero per estrapolare quali siano i topic più “hot” dagli articoli più recenti. Dopotutto l'esplorazione degli “hot topic” può essere un modo per estrarre contenuto e indirizzare in modo più preciso la ricerca (e.g. recommendation systems). È una tematica che potrebbe interessare il mondo accademico? Gradirebbero uno strumento del genere?
- Gli indicatori bibliometrici sembrano una consolidata misura di “impatto” nel mondo accademico. Ma influenzano realmente e in che modo il lavoro scientifico?

Nella seguente sezione “Abitudini di ricerca” è riportata l'esplorazione più dettagliata delle domande di ricerca che nascono da queste considerazioni.

Le risposte a queste stesse domande, implicitamente, possono delineare dei comportamenti di ricerca più viziosi. Ci si è dunque chiesti se l'utilizzo degli strumenti attuali possa in qualche modo avere un'influenza su tali comportamenti. Questo ha quindi portato alla formulazione di alcune ipotesi a riguardo, riportate nel capitolo “Ipotesi”.

4.5 Abitudini di ricerca

Si è pensato che il metodo più appropriato per esplorare estensivamente le abitudini del mondo accademico fosse attraverso un questionario online. Poiché non si disponeva delle risorse e del tempo necessari a effettuare più rounds per migliorare la qualità delle domande, si è cercato di costruire quest'ultime nel modo più appropriato possibile in un unico tentativo. Così si è giunti alla formulazione delle seguenti domande di ricerca, che cercheranno risposta nel questionario:

1. **Utilizzo degli strumenti esistenti:** per capire se un nuovo strumento di ricerca più avanzato possa essere interessante per gli accademici, bisogna prima capire come utilizzino e come valutino quelli già esistenti. Qualche funzionalità avanzata di buona qualità esiste (Scopus), ma se non venisse realmente usata o non fosse considerata interessante non avrebbe senso proporre di nuove ancor più complesse. A maggior ragione se, per contro, venissero usati molto di più e valutati efficaci strumenti più semplici come Scholar. Inoltre potrebbe valere la pena anche sapere se vengano usati di più strumenti verticali o di carattere generale, per capire se un nuovo strumento di ricerca debba seguire una o l'altra direzione. L'approccio che si ritiene più appropriato sarebbe quello di un motore di ricerca di carattere generale poiché favorisce l'interdisciplinarietà, ma d'altro canto un motore verticale funziona già implicitamente da filtro migliorando la qualità dei contenuti, nel caso si volesse rimanere circoscritti al proprio campo scientifico.

D1. Usano i motori e archivi online per scegliere il proprio topic di ricerca?

- i. Google Scholar?
- ii. Motore generico di Google?
- iii. Scopus?
- iv. Web of Science?
- v. Motori di ricerca verticali del proprio campo?

D2. Usano i motori di ricerca per trovare letture relative al proprio topic?

- i. Google Scholar?
- ii. Motore generico di Google?

- iii. Scopus?
- iv. Web of Science?
- v. Motori di ricerca verticali del proprio campo?

D3. Usano altri metodi più “tradizionali” per trovare letture?

- i. Seguono i riferimenti di papers già letti?
- ii. Ricevono la versione cartacea di Journals rilevanti nel proprio campo?
- iii. Attraverso il lavoro di editori per alcuni Journals del proprio campo?

D4. Come valutano la ricerca avanzata dei motori che usano per trovare letture?

- i. Google Scholar?
- ii. Scopus?
- iii. Web of Science?
- iv. Motori di ricerca verticali del proprio campo?

D5. Usano funzionalità più avanzate?

- i. Le funzionalità analitiche dei risultati di Scopus?
- ii. Il servizio di alerting di Scopus?

2. **Approccio “social”**: per ottenere risorse con una maggior attenzione ai contenuti si può seguire anche metodiche “social” (e.g. *crowdsourcing*, basato sul popolare concetto di “*wisdom of the crowd*”). Ovvero si può fare affidamento a relazioni sociali con altri ricercatori per ottenere contenuti più rilevanti. Dopotutto ci si tende a fidare dei pareri dei colleghi per i quali si prova rispetto e stima, anche nella valutazione della documentazione scientifica, quindi potrebbe essere verosimile come approccio che faciliti l’identificazione degli articoli più rilevanti. L’utilizzo di social network accademici come ResearchGate o Academia è sembrato più di successo di quanto ci si aspettasse, ma non è l’unica metodologia che presuppone relazioni sociali: ci sono anche le partecipazioni alle conferenze, i rapporti con

i membri del proprio gruppo di ricerca etc. Sarebbe quindi interessante capire se queste metodiche siano effettivamente nelle abitudini dei ricercatori.

D6. Usano metodiche “social” per trovare letture?

- i. Usano ResearchGate?
- ii. Usano Academia?
- iii. Vanno alle conferenze del proprio campo?
- iv. Tengono presentazioni alle conferenze del proprio campo?

D7. Usano i social network accademici per scegliere il proprio topic di ricerca?

- i. ResearchGate?
- ii. Academia?

D8. Come valutano i social network accademici per il reperimento di letture?

- i. ResearchGate?
- ii. Academia?

D9. Usano i servizi di alerting offerti dai social network accademici?

- i. Di ResearchGate?
- ii. Di Academia?

D10. Usano metodiche “social” per disseminare il proprio lavoro?

- i. Vanno personalmente alle conferenze dove sono pubblicati i loro lavori?
- ii. Contattano altri ricercatori su ResearchGate?
- iii. Contattano altri ricercatori su Academia?
- iv. Creano reti di contatti attraverso progetti di ricerca comuni?
- v. Circolano i propri papers per ottenere riscontri anticipati?
- vi. Circolano i propri papers per coinvolgere altri ricercatori?

D11. La conoscenza di un collega è un driver nel determinare se un documento valga la pena di essere letto?

- i. Se lo si conosce personalmente?
- ii. Se è tra le proprie cerchie in un social network come ResearchGate o Academia?
- iii. Se è un autore affermato e di impatto nel proprio campo?

D12. Tendono a seguire le direzioni dei colleghi per scegliere un topic di ricerca?

- i. Degli autori affermati e di maggiore impatto nel proprio campo?
- ii. Di un membro peer/senior nel proprio gruppo di ricerca?

D13. Tendono ad abbandonare un topic di ricerca sotto le direzioni di un membro peer/senior del proprio gruppo di lavoro?

D14. Scelgono il proprio topic di ricerca a seguito di cooperazioni con dei colleghi?

- i. All'interno del proprio campo scientifico?
- ii. All'esterno del proprio campo scientifico?

3. **Documentazione pratica:** il contenuto è importante quando si hanno dei problemi pratici da risolvere. Di solito, infatti, è quando si incorre in un problema pratico che si ha maggiore bisogno di recuperare contenuti rilevanti in breve tempo, perché si vorrebbe cercare di risolverlo il prima possibile. D'altro canto gli articoli teorici sembrano attirare il doppio delle citazioni, quindi potrebbero essere quelli più letti. Inoltre una peculiarità di Scholar è proprio quella di recuperare una maggiore quantità di documentazione di carattere pratico (e.g. brevetti, white papers aziendali, ...), e potrebbe essere una possibile motivazione per prediligere il suo utilizzo. Quindi vale la pena esplorare questa eventualità, perché potrebbe indicare se valga la pena o meno recuperare documentazione di questo genere in un eventuale nuovo strumento. Dunque quali sono realmente gli interessi degli accademici in tal senso?

D15. Quali sono le caratteristiche delle loro letture?

- i. Discutono di problemi teorici?
- ii. Discutono di problemi pratici?

iii. Descrivono casi di studio?

D16. Cercano anche altra documentazione oltre ai papers?

i. White papers aziendali?

ii. Brevetti?

iii. Slides?

iv. Video?

D17. Il fatto che una pubblicazione sia di un certo tipo è un driver nella determinare se un documento valga la pena di essere letto?

i. Se affronta un problema teorico?

ii. Se affronta un problema pratico?

iii. Se porta a risultati utili e applicabili?

4. **Hot topic:** l'esplorazione degli "hot topic" dal set di articoli più recenti potrebbe essere una applicazione interessante di estrazione di contenuti dalle pubblicazioni. Spesso infatti potrebbe essere un processo complesso e lungo estrapolare quali siano i trend attuali nel mondo scientifico, in particolare se si cerca di farlo analizzando i risultati dei motori di ricerca. Quindi una funzionalità che identifichi e raccolga automaticamente i topic più frequenti nella recente letteratura potrebbe essere interessante per i ricercatori, in quanto, velocizzando tale processo, riuscirebbero a indirizzare più precisamente le loro ricerche future. Che idee hanno gli accademici a riguardo?

D18. Come definirebbero il fatto che un topic si possa considerare "hot"?

i. Se molti ricercatori ci stanno lavorando?

ii. Se molte compagnie ci stanno investendo?

iii. Se è incluso in molte call per ricerche finanziate?

iv. Se è il focus di molte pubblicazioni?

v. Se i papers che lo trattano hanno una maggiore probabilità di essere citati?

D19. Tendono a fare ricerca su topic che siano "hot"?

D20. Sono propensi ad abbandonare un topic perché non è più “hot”?

D21. Troverebbero utile uno strumento che raccolga automaticamente tutti i topic più frequenti in un campo scientifico?

5. **Citazioni e impatto bibliometrico:** attualmente le più consolidate metriche d’impatto del lavoro scientifico sono gli indicatori bibliometrici. Il mondo accademico ha dedicato una grandissima attenzione a questo tema, e contemporaneamente ha dimostrato una scarsa propensione a cambiare verso una differente metrica di “impatto”, poiché la maggior parte delle alternative proposte in letteratura ha riscosso un basso successo. Potrebbe quindi valere la pena mantenere questi indicatori (o quantomeno le loro varianti più promettenti) per determinare l’impatto di un documento. Anche perché i motori di ricerca attuali, a parte a calcolarle globalmente, non le utilizzano per determinare la rilevanza dei risultati e non permettono all’utente di parametrare la ricerca in base ad essi. Vista la loro persistenza nel mondo accademico vale quindi la pena esplorare come queste metriche influiscano sul lavoro dei ricercatori e sulle loro abitudini di citazione.

D22. Tendono a fare ricerca su topic con maggiore probabilità di citazione?

D23. Gli attuali indicatori più consolidati sono dei driver nel determinare se un documento vale la pena di essere letto?

i. Se ha un alto numero di citazioni?

ii. Se sono pubblicate su un Journal con un alto IF?

iii. Se uno degli autori ha un alto h-index?

D24. Con che criteri scelgono le citazioni da includere nei propri lavori?

i. In base alla pertinenza con il proprio topic di ricerca?

ii. Per supportare specifiche affermazioni del proprio lavoro?

iii. In base alla conoscenza personale di uno degli autori?

iv. In base all’IF del Journal in cui è pubblicato l’articolo citato?

v. Se almeno uno degli autori dell’articolo citato ha un alto h-index?

vi. In base al Journal sul quale intendono pubblicare il lavoro?

vii. Si sforzano di includere papers altamente citati?

viii. Citano autori che più probabilmente li ri-citeranno a loro volta?

6. **Domande di contesto:** le domande precedenti rispondono a specifiche esigenze. Tuttavia, per mascherare le motivazioni sottostanti nella speranza di ottenere risposte più attendibili, si è pensato convenisse aggregarle tematicamente al momento della presentazione agli accademici, anche quando servono a rispondere a esigenze diverse (e.g. aggregare tutte le domande che riguardano i driver nel determinare se un documento valga la pena di essere letto, o tutte quelle inerenti la scelta del proprio topic di ricerca). Di conseguenza si sono aggiunte alcune domande di contorno che servissero a completare ciascun contesto, ovvero:

D25. Altre motivazioni di scelta del topic di ricerca?

- i. Come conseguenza di contratti con delle compagnie?
- ii. Come conseguenza di progetti peer-reviewed?
- iii. Secondo i propri interessi personali di ricerca?

D26. Altre motivazioni di abbandono di un topic di ricerca?

- i. Perché il progetto collegato è terminato?
- ii. Per mancanza di fondi?
- iii. Perché pianificano di cambiare gruppo di ricerca?
- iv. Perché pianificano di cambiare istituto?
- v. Per perdita di interesse intellettuale?

D27. La pertinenza col proprio topic di ricerca è un driver nel determinare se un documento valga la pena di essere letto?

4.6 Ipotesi

Le risposte alle domande descritte finora potrebbero implicitamente delineare dei comportamenti viziosi nelle abitudini degli accademici. In particolare potrebbero emergere tre categorie di comportamenti:

1. Un comportamento più gregario e meno indipendente. Il lavoro scientifico dovrebbe essere libero in principio. Ciascuno dovrebbe osservare lo scenario accademico nel suo complesso ed avere spirito critico nel valutare il lavoro dei

colleghi in modo imparziale, al fine di prendere le proprie decisioni. Tuttavia potrebbero anche avvenire condizionamenti da dei colleghi, che finiscono per influire sulla propria attività scientifica. Nella scelta del proprio argomento di ricerca, ad esempio, si potrebbe essere portati a seguire le direzioni di un autore perché è considerato affermato e d'impatto nel proprio campo, sperando di seguire la scia del suo successo. Oppure si potrebbero seguire le sue direzioni semplicemente perché si asseconda assertivamente l'opinione condivisa e lo si reputa un'autorità, a prescindere dall'effettivo valore delle sue pubblicazioni. In entrambi i casi, il risultato è che si lascia influenzare eccessivamente il proprio lavoro. Un altro esempio di comportamento gregario potrebbe essere dato dall'eccessivo condizionamento subito dai colleghi all'interno del proprio gruppo di ricerca. Si potrebbe finire per scegliere o abbandonare un argomento di ricerca solo in base a direzioni provenienti da essi, piuttosto che da motivazioni e ragionamenti mutuati autonomamente secondo i propri interessi personali.

2. Un comportamento più opportunistico. Ovvero la tendenza ad assumere comportamenti mirati a un proprio tornaconto in termini di incremento dei propri indicatori. Dopotutto i ricercatori sono consapevoli che verranno valutati in base alle loro performance dalle università, dalle aziende, dai governi etc. Di conseguenza potrebbero instaurarsi nelle loro abitudini atteggiamenti volti ad incrementare le possibilità di essere valutati positivamente da queste istituzioni. Come si è già descritto, non è un segreto che esista la possibilità che i Journals prediligano pubblicazioni di autori affermati e, viceversa, gli autori cerchino di pubblicare i propri lavori su Journal importanti (e.g. con un alto IF), entrambi volti allo scopo di aumentare il proprio prestigio attraverso l'incremento dei propri indicatori bibliometrici. Inoltre, per lo stesso motivo, gli stessi Journal potrebbero avanzare determinate pretese nelle citazioni da includere nei propri lavori. Ne consegue che gli autori potrebbero condizionare la scelta delle citazioni da includere nel proprio lavoro in base al Journal sul quale intendono pubblicarlo, allo scopo di incrementare le possibilità di accettazione dallo stesso. Un altro esempio di comportamento opportunistico molto più semplice potrebbe essere una maggiore inclinazione a scegliere topic di ricerca che si pensa godano di una maggiore probabilità di citazione, e l'immediato beneficio di che si spera di trarne è evidente.

3. Una maggiore sensibilità agli indicatori bibliometrici. Così come si viene valutati dalle istituzioni attraverso di essi, i ricercatori potrebbero cominciare a loro volta a usarli come metri di valutazione del lavoro altrui. Questi stessi indicatori potrebbero quindi cominciare ad esercitare una sempre maggiore influenza in tutte le varie fasi del lavoro scientifico. Ad esempio, nel determinare il valore di una lettura, potrebbero essere condizionati dal fatto che sia pubblicata in un Journal con un alto IF, o che uno degli autori abbia un alto h-index. Oppure ancora potrebbero farsi condizionare da essi anche nella scelta delle citazioni da includere nei propri lavori: così come si pensa che pubblicazioni con migliori indicatori abbiano un maggiore valore, si potrebbe pensare che anche il proprio lavoro acquisirà valore includendo citazioni ad articoli del genere. Questo a sua volta si potrebbe anche ricondurre a un comportamento opportunistico del punto precedente, in quanto anche non attribuendo un particolare valore a queste metriche, si potrebbe in ogni caso decidere di inserire citazioni con buoni valori nella speranza di aumentare i propri. Sia nel caso che li si riconosca come metriche valide o meno, il risultato potrebbe essere quindi lo stesso: subire la loro influenza durante il proprio processo di lavoro scientifico.

In generale è nato il sospetto che la crescente importanza degli indicatori bibliometrici nella carriera dei ricercatori e il parallelo utilizzo dei motori di ricerca che contribuiscono a metterli in evidenza, possano aumentare l'incidenza di tali comportamenti. I motori di ricerca esercitano un notevole potere sul lavoro scientifico, in quanto già solo l'ordine in cui presentano i risultati può contribuire alla formazione di un'opinione piuttosto che l'esatto opposto nei suoi utilizzatori. Inoltre attraverso di essi si potrebbe enfatizzare la percezione della necessità di acquisire sempre maggiore visibilità e prestigio per migliorare (o mantenere) la propria carriera, in quanto è anche attraverso tali strumenti che il mondo accede agli indicatori con cui valutare l'impatto di un ricercatore. Ci si è dunque chiesti se l'utilizzo degli attuali strumenti di ricerca possa in qualche modo cronicizzare ed enfatizzare questo genere di comportamenti. L'ipotesi a cui si è giunti, è che allo stato attuale un'influenza ce l'abbiano, e in particolare di carattere peggiorativo (i.e. l'utilizzo di tali strumenti contribuisce ad incrementare l'incidenza di questi comportamenti).

Tutto questo, nello specifico, ha portato alla formulazione di tre ipotesi di ricerca:

H1 - Gli utilizzatori assidui degli strumenti attuali tendono ad avere un comportamento più gregario, quindi meno indipendente. Siccome in tali motori il peso maggiore nell'attribuire rilevanza ai contenuti è dato dal numero totale di citazioni, i primi risultati tendono a essere stabili e quasi sempre gli stessi. Questo dà poco spazio ad articoli nuovi o che trattano argomenti di nicchia, mentre contribuisce a mantenere in evidenza i documenti già popolari. Quindi sembra verosimile che coloro che li usano frequentemente possano essere più influenzati dagli autori di maggiore impatto del loro campo, dato che i loro lavori saranno quelli che visualizzano più spesso. Questo tratto di personalità non è tuttavia legato ai soli autori affermati nel proprio campo, ma si potrebbe anche estendere ad altri contesti. Si pensa più in generale che possano tendere a presentare un comportamento maggiormente condizionato dai colleghi, anche all'interno del proprio gruppo di ricerca. Si potrebbe rilevare questa tendenza osservando l'andamento delle risposte a domande come 2.D12.i, 2.D12.ii, 2.D13, 4.D18.i .

H2 - Gli utilizzatori assidui degli strumenti attuali tendono ad avere un comportamento più opportunistico. Sempre perché il fattore principale nel determinare la rilevanza dei risultati è dato dal numero totale di citazioni, si pensa si possa instaurare nei loro utenti una maggiore convinzione della necessità di dover attirare a loro volta un gran numero di citazioni, così da poter essere più in evidenza nel mondo scientifico. Questo potrebbe dar adito a comportamenti viziosi mirati a migliorare la propria visibilità. Inoltre è attraverso questi stessi strumenti che vengono consultati gli indicatori bibliometrici dei ricercatori per valutarne l'impatto, quindi è proprio interesse cercare di ottenere buone performance all'interno di tali strumenti e nulla vieta di farlo anche attraverso alcune delle azioni manipolatorie sopra descritte. Si potrebbe rilevare questa tendenza osservando l'andamento delle risposte a domande come 4.D18.v, 4.D20 (anche in negativo, ovvero se quelli più virtuosi sono meno inclini a farlo), 5.D22, 5.D24.vi, 5.D24.viii.

H3 - Gli utilizzatori assidui degli strumenti attuali tendono ad essere più sensibili agli indicatori bibliometrici. Come si è già detto, sono questi strumenti che rendono disponibili pubblicamente gli indicatori relativi a ciascun ricercatore. Di conseguenza si è portati a pensare che questo possa indurre a guardare con maggiore attenzione tali indicatori, poiché è attraverso di essi che si viene

valutati da chi li consulta, e quindi si potrebbe finire per usarli a propria volta come metriche di valutazione del lavoro altrui. Sensibilità che, come anticipato, si riverserebbe in tutte le fasi della produzione scientifica. Si potrebbe rilevare questa tendenza osservando l'andamento delle risposte a domande come 5.D23.i, 5.D24.iv, 5.D24.v , 5.D24.vii.

4.7 Conclusione

In questo capitolo si è esposto il problema della fattibilità e delle direzioni che dovrebbe seguire un potenziale strumento che permetta una ricerca migliore nella letteratura scientifica, possibilmente con una maggiore attenzione ai contenuti. A tale scopo si è però presentata la necessità di rispondere determinati interrogativi sulle abitudini del mondo accademico, che hanno portato alla formulazione delle domande di ricerca di questo lavoro.

Parallelamente si è evidenziato che le risposte a queste stesse domande potrebbero delineare dei comportamenti di carattere vizioso nei ricercatori, e l'ipotesi di questa tesi è che l'utilizzo degli attuali strumenti di ricerca eserciti un'incidenza negativa su di essi.

Per rispondere a queste domande e validare queste ipotesi si è pensato di creare un questionario da somministrare a un gruppo di accademici, la cui analisi è riportata nel capitolo seguente.

5 Analisi dei dati

5.1 Introduzione

Inizialmente il progetto sopracitato si riproponeva di organizzare uno studio Delphi per determinare le direzioni da prendere nell'eventuale sviluppo di uno strumento che basi la ricerca in letteratura maggiormente sui contenuti. Tuttavia le considerazioni emerse nell'analisi della letteratura hanno portato alla consapevolezza della necessità di uno studio estensivo sulle abitudini di ricerca del mondo accademico. Non sarebbe stato possibile ottenere un output del genere da un Delphi, che coinvolge un ristretto numero di esperti che cercano di convergere (sperabilmente) su un topic in una serie di round. Di conseguenza si è presentata la necessità di gestire un questionario esteso a un grande numero di accademici. Il design del questionario non è riportato in quanto completamente visionabile dalla Tab. 8 in Appendice, in cui sono riportati i risultati globali ottenuti. La gestione del questionario è avvenuta sulla piattaforma SurveyFace [55], in virtù del fatto che rispetto ai concorrenti non presenti quasi alcuna limitazione nella versione gratuita.

Per la selezione degli accademici si è scelto l'ambito dei Management Information Systems (MIS) e più in generale della Computer Science, per l'attinenza al campo di ricerca dei gruppi coinvolti, e anche nella speranza che presentassero una maggiore affinità con gli strumenti tecnologici sui quali sarebbero stati interrogati. In particolare sono stati selezionati tutti gli accademici che abbiano pubblicato almeno un articolo tra il 2013 e il 2015 nei top Journals nell'ambito dei MIS. I Journals selezionati a tale scopo sono:

- MIS Quarterly (MISQ)
- Journal of MIS (JMIS)
- Journal of Information Technology (JIT)
- Journal of the Association for Information Systems (JAIS)
- Journal of Strategic Information Systems (JSIS)
- Information Systems Research (ISR)
- Information Systems Journal (ISJ)
- European Journal of Information Systems (EJIS)

- Management Science (MS)
- IEEE Transaction on Software Engineering (IEEE)
- Communications of the ACM (CACM)

Questo ha portato alla raccolta di 3453 nominativi.

Il questionario ha ottenuto complessivamente 596 risposte tra il 20 marzo 2015 e il 22 aprile 2015, ma sono state analizzate esclusivamente le 326 risposte complete, in modo da poterne effettuare analisi consistenti le cui inferenze si potessero applicare all'intero campione considerato. Di seguito il Grafico 1 riporta la distribuzione geografica di coloro che hanno risposto, mentre il Grafico 2 ne mostra la distribuzione nei diversi Journals.

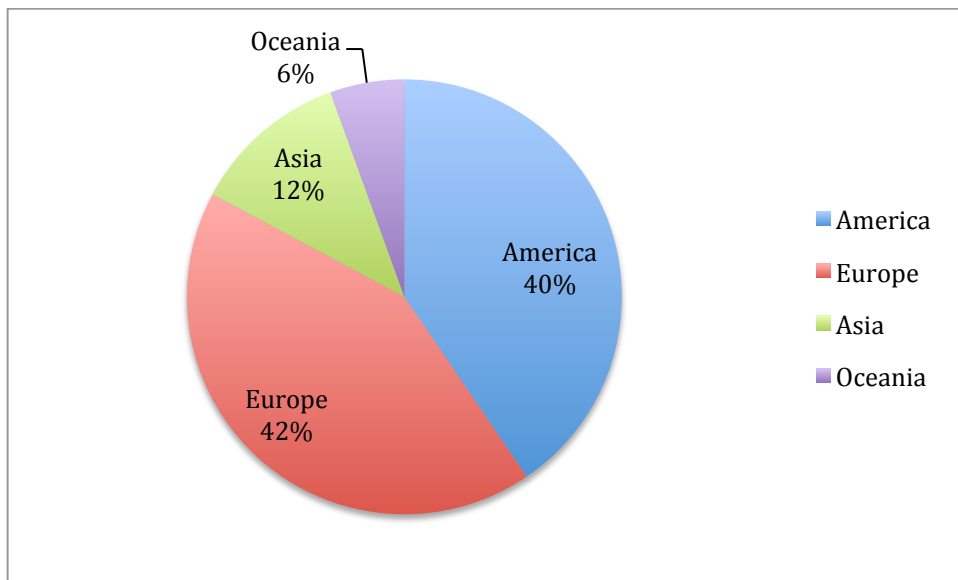


Grafico 1 Distribuzione geografica

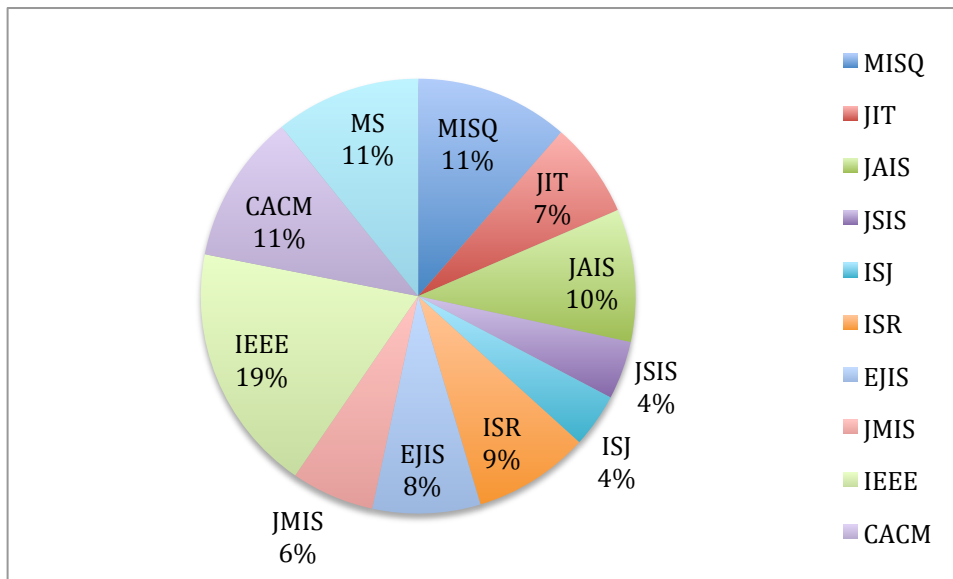


Grafico 2 Distribuzione nei Journals

5.2 Analisi qualitativa globale

Il sommario generale usato per l'analisi qualitativa globale, e le singole risposte necessarie a effettuare analisi più elaborate sono stati scaricati direttamente dalla piattaforma usata per la gestione del questionario, SurveyFace.

Di seguito si evidenziano alcuni dei risultati più rilevanti, mentre la Tab. 8 in Appendice riporta i risultati completi (in percentuale, con le frequenze tra parentesi).

- Il fatto che un topic sia “hot” se ha una maggiore probabilità di essere citato ha ricevuto un consenso globale paragonabile alle altre definizioni proposte. Comunque si può dire che globalmente vi è consenso più o meno uniformemente distribuito tra le diverse opzioni proposte, fornendo informazioni per la domanda 4.D18. L'unica che ha riscosso meno consenso è la definizione che prevede una maggiore opportunità di cooperazioni con altri ricercatori.
- Il principale motivo di scelta del proprio topic di ricerca risulta in preponderante maggioranza l'interesse personale, che è un comportamento sano. A seguire, la scelta si basa sulle opportunità di cooperazione con ricercatori all'*interno* del proprio campo scientifico (2.D14.i). Inoltre viene discretamente seguita l'onda di un “hot topic”, con il ~46% che afferma di fare ricerca frequentemente o più su tali argomenti (4.D19), di più persino delle opportunità di cooperazione con ricercatori *fuori* dal proprio campo scientifico (2.D14.ii), denotando una bassa predisposizione all'interdisciplinarietà. Le altre opzioni

meno seguite hanno ottenuto distribuzioni più o meno uniformi tra loro con: ~30% fanno ricerca su topic con maggiore probabilità di citazione “frequentemente” o più (5.D22), ~29% scelgono frequentemente o più in base a suggerimenti di membri peer/senior del proprio gruppo (2.D12.ii) e ~26% seguono frequentemente o più le direzioni degli autori di maggiore impatto nel proprio campo (2.D12.i).

- Il cambiamento del proprio argomento di ricerca risulta abbastanza lento, in media ogni 4,33 anni.
- Il principale motivo di mutamento del proprio topic di ricerca è la perdita di interesse intellettuale, ancora un comportamento sano, a seguire la fine di un progetto collegato ad esso. Dei restanti con minor consenso, il fatto che un topic non sia più “hot” (4.D20) è comunque abbastanza sentito con ~32% di “strongly/weakly agree”, mentre il seguire le direzioni di un membro peer/senior del proprio gruppo (2.D13) lo è meno (~14% “strongly/weakly agree”).
- Riguardo l'utilizzo di archivi online e motori di ricerca *per scegliere il proprio topic di ricerca* (1.D1) ha avuto esiti sorprendenti: il ~60% sostiene di non usarli (o poco) a tale scopo (coloro che hanno inserito ‘Strongly’ o ‘weakly agree’) mentre un ~40% sostiene di basare su di essi tale scelta. L'incredibile influenza che questi strumenti sembrano avere sull'orientamento della ricerca scientifica è preoccupante, in particolar modo visti i problemi che ciascuno di essi comporta. Oltretutto il più usato a tale scopo sembra essere Google Scholar, con ~57% che sostiene di usarlo (‘strongly’ o ‘weakly’), nonostante sia quello maggiormente criticato in letteratura. Come se non bastasse, subito dopo Scholar il più usato sembra il motore di ricerca generico di Google, che non è nemmeno intrinsecamente legato al mondo accademico, il quale oltretutto ha ricevuto un ‘agreement’ maggiore degli stessi motori di ricerca verticali (~45% contro ~33%, sommando le risposte ‘strongly’ e ‘weakly agree’).

Sembra quasi che il mondo si divida in due: quelli che usano gli strumenti di casa Google per scegliere gli argomenti di ricerca, e quelli che usano altre vie. I restanti motori di ricerca e i social network accademici (2.D7) presentano un notevole distacco nell'utilizzo rispetto a Google e Scholar (nei social network,

ResearchGate sembra tuttavia più affermato di Academia). La funzionalità avanzata di analisi dei risultati di Scopus è prossima all'inutilizzo (1.D5.i). Questo scenario conferma quanto poco il mondo accademico sembri attratto da funzionalità sofisticate come quelle offerte da Scopus, mentre predilige in modo predominante la semplicità di ricerca di uno strumento limitato come Scholar.

- La maggior parte (~76% di “yes o “may be yes”) degli accademici sembra interessata a un ipotetico strumento che raccolga automaticamente i topic di ricerca più frequenti nei papers scientifici (4.D21).
- Per quanto riguarda le fonti che vengono usate *per trovare letture inerenti al proprio topic di ricerca* (1.D2), emergono ancora gli strumenti forniti da Google, con ~64% per Scholar e ~53% per Google che affermano di usarli nel breve termine (i.e. ‘ogni giorno’ o ‘ogni settimana’). Comunque ~61% nel breve termine segue anche i riferimenti nella bibliografia di papers interessanti che ha letto (1.D3.i). Seguire i riferimenti è un comportamento compatibile con gli indicatori bibliometrici iterativi (e.g. EigenFactor), poiché di fatto seguono un cammino lungo il grafo delle citazioni, e questo potrebbe indirizzare una maggior giustificazione all'utilizzo di questo genere di indicatori rispetto a quelli basati in generale sul numero di citazioni (e.g. Impact Factor). Anche qui, i social network accademici (2.D6.i, 2.D6.ii) e gli altri motori di ricerca vedono un distacco notevole nell'utilizzo rispetto a Scholar e Google. Sembrano sfruttare abbastanza le conferenze del proprio campo per reperire letture (2.D6.iii, 2.D6.iv), mentre sembrano non usare molto metodiche più tradizionali come l'essere reviewer di qualche Journal del proprio campo (1.D3.iii) o ricevere la versione cartacea di un Journal (1.D3.ii).
- Scholar è percepito come incredibilmente *efficace nella ricerca di letture inerenti al proprio topic* (1.D4), con ~31% che lo valuta ‘eccellente’ e il ~48% che lo valuta ‘buono’. Il che è sorprendente vista la grande quantità di critiche ricevute, è quindi probabile che la maggior parte degli utenti non sia consapevole delle debolezze che cela tale strumento. Dopo Scholar, sono percepiti abbastanza bene anche i motori di ricerca verticali, con ~10% di ‘eccellente’ e ~25% di ‘buono’, sebbene ~40% non li usi proprio. Tutti gli altri, per quel poco che vengono usati, sono valutati per la maggioranza come

'buoni' o meno. Confrontando i social network accademici (2.D8), ResearchGate sembra più usato di Academia (~58% non lo usa mai contro ~75%) e con una reputazione leggermente migliore. In generale quindi si conferma quanto il mondo accademico si reputi complessivamente soddisfatto delle semplici ma limitate funzionalità di Scholar, rispetto alla complessità di strumenti come Scopus.

- La larga maggioranza degli accademici cerca quasi esclusivamente papers (~85% di 'always'), degli altri tipi di documenti al massimo si arriva a ~21% di 'frequently' (notare che non si sta nemmeno parlando dell' 'always') per i 'Company White Paper' e ancor meno per gli altri. Quindi sembra in generale che al mondo accademico interessi pochissimo documentazione diversa dai canonici articoli (3.D16).
- Non sembra esserci un maggior interesse per papers teorici o pratici, si distribuiscono in modo abbastanza uniforme tra di essi (3.D15). Solo i casi di studio sembrano ricevere leggermente meno attenzione.
- I servizi di alerting su nuove ricerche/documenti che potrebbero essere di interesse degli accademici non sono usati quasi per nulla, sia quello di Scopus (1.D5.ii) che dei social network accademici (2.D9), ma tra tutti ResearchGate sembra quello che riscuote leggermente più successo.
- Tra i fattori che determinano se una pubblicazione vale la pena di essere letta quello dominante su tutti è la pertinenza al proprio topic di ricerca (~97% di 'strongly/weakly agree', dei quali ~80% 'strongly'), facendo emergere ancora una bassa predisposizione all'interdisciplinarietà. A seguire vi è il fatto che tratti di argomenti applicabili e utili (~46% 'strongly', ~40% 'weakly'), e che trattino di argomenti teorici o pratici, senza una particolare polarizzazione tra le due (3.D17). Il fatto che un autore abbia un alto h-index (5.D23.iii) *non* è visto come un valido motivo (~72% 'strongly/weakly disagree'), facendo emergere la bassa percezione che si ha di tale indicatore, il che è concorde con le numerose critiche rilevate in letteratura. D'altro canto l'impact factor della rivista che lo ha pubblicato ha una notevole influenza (~67% di 'strongly/weakly agree', di cui ~23% 'strongly') (5.D23.ii), facendo emergere una considerazione notevolmente più alta di quello che ci si aspettava di tale

indicatore, così come per il numero totale di citazioni (5.D23.i) che ha ottenuto frequenze analoghe. Dei restanti hanno ottenuto buone frequenze anche il fatto che l'articolo sia scritto da un autore noto e affermato nel proprio campo (2.D11.iii) o che lo si conosca personalmente (2.D11.i), mentre l'appartenenza a una propria cerchia sui social network accademici (2.D11.ii) lo è molto meno.

- Tra i motivi che guidano la scelta delle citazioni da inserire in una propria pubblicazione (5.D24), quelli dominanti sono la necessità di supportate specifiche affermazioni (~73% di 'always'), e la pertinenza al proprio topic di ricerca (~62% di 'always'), che sono dopotutto due comportamenti sani di citazione. Tuttavia una porzione non irrilevante risulta influenzata dal journal sul quale dev'essere pubblicato l'articolo (~57% lo fa frequentemente o più, dei quali ~13% sempre), quindi sembra che le pressioni editoriali dei journals si facciano sentire. Le motivazioni basate sugli indicatori bibliometrici non vengono seguite altrettanto spesso, ma non poco quanto ci si aspetterebbe e comunque facendo emergere ancora una bassa percezione dell'h-index degli autori citati (~92% raramente o meno, dei quali ~67% mai), e una buona percezione dell'impact factor della rivista (~40% lo fa frequentemente o più). Anche lo sforzo di inserire papers altamente citati viene fatto più spesso di quanto ci si aspetterebbe (~36% lo fa frequentemente o più).
- I propri lavori vengono disseminati (2.D10) soprattutto partecipando a conferenze in cui è pubblicato il proprio lavoro (~82% frequentemente o più, di cui ~30% sempre). A seguire vengono svolte attività che coinvolgano contatti con altri ricercatori (creazione di rete di contatti da cooperazioni in progetti di ricerca comuni, ~56% frequentemente o più, e circolazione anticipata della pubblicazione per ricevere riscontri, ~40% frequentemente o più, o per coinvolgere altri ricercatori, ~25% frequentemente o più), anche se questi non sembrano avvenire attraverso social network accademici che risultano ancora scarsamente usati (ResearchGate un po' più di Academia).

L'analisi qualitativa ha fatto emergere diverse considerazioni sorprendenti, ma prima fra tutte, conferma l'incredibile utilizzo di Scholar e Google nonostante la nutrita critica in letteratura a tali strumenti: non solo hanno visto un vasto utilizzo nel procurarsi le letture, ma anche nell'effettiva scelta del proprio topic di ricerca. Questo

preoccupa perché la bassa qualità di tale strumento per la ricerca in letteratura potrebbe influenzare negativamente il lavoro scientifico di molti accademici, se non usato con accortezza e spirito critico. È anche confermato uno scarso utilizzo di tutti gli altri motori di ricerca di carattere generale.

L'analisi smentisce invece l'utilizzo dei Research Network, che non vedono un utilizzo elevato come quello che farebbe evincere l'analisi su Alexa.com. Questo, in una certa misura, conferma il sospetto che la grande affluenza di visite registrate sui Research Network sia incanalata principalmente da Scholar e Google, usati invece di frequente, i quali spesso riportano nei risultati documenti contenuti in tali siti.

5.3 Analisi delle correlazioni

5.3.1 Metodo

Il primo passo successivo all'analisi qualitativa, è stato cercare l'esistenza di correlazioni tra le risposte date alle diverse domande. Per svolgere tale analisi le singole risposte sono state ricodificate secondo una codifica 'naturale' (e.g. 1 per 'strongly disagree', 2 per 'weakly disagree', etc. e analogamente per le risposte che rappresentano una frequenza). I dati così ottenuti sono stati quindi importati nel software statistico open-source R [56], nel quale si è svolta l'analisi delle correlazioni e relativi test che ne determinassero la significatività. Sono stati analizzati i 3 coefficienti di correlazione più comuni (Pearson, Spearman e Kendall) per valutare la concordanza delle osservazioni che si possono fare ponendoli a confronto. Definendo i tre coefficienti:

Coefficiente di correlazione per ranghi di Spearman

È una misura statistica non parametrica della correlazione tra due variabili, senza alcuna ipotesi all'infuori della misura ordinale dei dati. Può essere definito come:

$$\rho_s = 1 - \frac{6 \sum_i D_i^2}{N(N^2 - 1)}$$

Dove $D_i^2 = r_i - s_i$ con r_i e s_i i ranghi della prima e della seconda variabile nella i -esima osservazione, e N è il numero complessivo di osservazioni. Di fatto si può pensare ad esso semplicemente come un coefficiente di Pearson applicato sui ranghi delle osservazioni (da qui l'assenza di assunzioni sulla distribuzione dei dati).

Coefficiente di correlazione lineare di Pearson

È una misura statistica della correlazione lineare tra due variabili, sotto l'ipotesi che seguano una distribuzione gaussiana. Può essere definito come:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Dove σ_{XY} è la covarianza tra le due variabili, mentre σ_X e σ_Y le loro rispettive deviazioni standard.

Coefficiente di correlazione τ di Kendall

È una misura statistica non parametrica della correlazione tra due variabili, senza alcuna ipotesi all'infuori della misura ordinale dei dati. Può essere definito come:

$$\tau = \frac{n_c - n_d}{\frac{1}{2}n(n - 1)}$$

Dove con n_c e n_d sono il n° di coppie concordanti e discordanti, e n è il numero complessivo di osservazioni. Per sua natura, è simile al coefficiente di Spearman, infatti sono spesso concordi, ma è più conservativo (i.e. valori leggermente più bassi dei coefficienti, $\tau < \rho_s$).

5.3.2 Analisi

L'analisi completa delle correlazioni secondo il coefficiente di Spearman è riportata nella Tab. 9 in Appendice, nella quale sono stati inclusi tutti i coefficienti con $\rho > 0,4$ e $p - value < 0,05$, ordinati in modo decrescente per ρ . La maggior parte di essi presenta $p - value$ estremamente piccoli (nell'ordine di 10^{-16}) che si è deciso di riportare nei rilievi testuali come $p < 0.0001$, poiché in generale è buona norma non fare troppo affidamento ai valori esatti di p sotto un certo livello (da qui la scelta della soglia dello 0,01%). Di seguito si riportano una selezione (Tab. 5) della tabella completa e le considerazioni più rilevanti in merito ai coefficienti $\rho > 0,5$:

Question 1 (row)	Question 2 (column)	Spearman ρ	p-value
Q5. Why do you move away from a research topic - Because I plan to move to a different research institution	Q5. Why do you move away from a research topic - Because I plan to move to a different research group	0,754753888	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it tackles a practical problem	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it presents research results that are practically applicable and useful	0,647994876	<2,22E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on the impact factor of their publication outlet	Q16. How do you choose the citations to be included in your own publications - I make an effort to include highly cited papers among my citations	0,553334475	<2,22E-16

Q16. How do you choose the citations to be included in your own publications - I choose citations based on the impact factor of their publication outlet	Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	0,5516029	<2,22E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on my target publication outlet	Q16. How do you choose the citations to be included in your own publications - I make an effort to include highly cited papers among my citations	0,527987123	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors has a high h index	Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	0,521663964	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors is a well known established researcher in your field	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when I personally know at least one of the authors	0,5115906	<2,22E-16

Tab. 5 Selezione di risultati rilevanti nelle correlazioni di Spearman

- Utilizzo degli strumenti:* in generale, la maggior parte delle correlazioni più forti si sono rilevate, in modo non sorprendente, tra domande che coinvolgono entrambe gli stessi specifici motori di ricerca o archivi online. A titolo di esempio, risulta una correlazione tra la domanda riguardante l'utilizzo di Academia.edu per trovare letture interessanti inerenti al proprio topic di ricerca e quella che richiede di valutare l'efficacia di Academia.edu ($\rho = 0,722290576$; $p < 0.0001$). Poiché da una forte correlazione positiva si può evincere che le due variabili tendano a rimanere concordi (al crescere di una, tende a farlo anche l'altra) si può dedurre che (i pochi che) utilizzano Academia lo trovano anche mediamente efficace. Da altre correlazioni di questo genere sembra emergere che chi utilizza determinati strumenti, lo fa sfruttandone mediamente tutte le funzionalità e lo reputa mediamente efficace. Sulla causalità di queste relazioni ovviamente non è possibile fare inferenze affidabili guardando solo il coefficiente di correlazione (non si può dedurre che chi usa Academia lo fa *perché* lo trova efficace), ma forniscono senz'altro una conferma del quadro generale che ci si aspetterebbe.
- Cambio di istituzione e di gruppo di ricerca:* la correlazione ($\rho = 0,754753888$; $p < 0.0001$) tra il cambio del proprio topic in seguito allo spostamento a un altro gruppo di ricerca, piuttosto che allo spostamento a un altro istituto suggeriscono, come ci si aspetterebbe, che le due cose vadano spesso di pari passo (o si cambia gruppo di ricerca perché si cambia istituto, o è vero il viceversa).

- *Pratica*: la correlazione ($\rho = 0,647994876$; $p < 0.0001$) tra la valutazione di una lettura se affronta un argomento pratico, e se riporta risultati utili e applicabili, mostra una attenzione non sorprendente al nesso tra le due cose. Chi pensa che una pubblicazione valga la pena di essere letta se affronta un problema pratico, si aspetta probabilmente che oltre ad affrontarlo arrivi anche a dei risultati concreti.
- *Attenzione all'IF*: una correlazione meno forte ma ancora tutt'altro che trascurabile ($\rho = 0,553334475$; $p < 0.0001$) tra la scelta delle citazioni da includere nei propri lavori in base all'impact factor dei journals in cui sono pubblicate e lo sforzo di introdurre papers altamente citati conferma l'attenzione del mondo accademico a misure quantitative di popolarità come l'IF.
- *IF e H-index*: si è rilevata una correlazione ($\rho = 0,553334475$; $p < 0.0001$) tra la scelta delle citazioni da includere nel proprio lavoro in base all'IF dei journals in cui è pubblicata l'opera citata e l'h-index degli autori sempre dell'opera citata. Tuttavia visto che dall'analisi globale era emersa una certa divergenza nelle opinioni che coinvolgessero l'IF e l'H-index, è probabile che tale correlazione possa solamente suggerire che nella maggior parte dei casi se un accademico dà valore all'H-index, dà valore anche all'IF, ma non è altrettanto vero il contrario (i.e. coloro che valutano positivamente l'h-index sono meidamente un sottoinsieme di coloro che valutano positivamente l'IF).
- *Influenza editoriale*: altra correlazione non irrilevante ($\rho = 0,527987123$; $p < 0.0001$) che si è riscontrata è tra la scelta delle citazioni da includere nei propri lavori in base al Journal sul quale verrà pubblicato e lo sforzo di includere papers altamente citati. Sembrerebbe quindi che chi si lascia influenzare maggiormente dal journal sul quale deve pubblicare un suo lavoro, mediamente cerca anche di includere papers altamente citati in esso. Lungi dal voler essere affermazioni statisticamente provate, fa comunque emergere delle possibili ipotesi a riguardo: includono papers altamente citati per aumentare le probabilità di essere pubblicati dal journal? Se questo fosse il caso, lo fanno per aumentare la credibilità del proprio lavoro perché attribuiscono un grande valore al n° di citazioni o per cercare di attirarne a loro volta di più ed incrementare i propri indici? Oppure lo fanno perché gli è stato

consigliato di farlo dal Journal stesso? E, ancora, se questo fosse il caso, il journal lo fa per aumentare la propria affidabilità o solo il proprio IF?

- *Bipolarità dell'h-index*: la correlazione ($\rho = 0,521663964$; $p < 0.0001$) tra la valutazione di una lettura sulla base dell'h-index degli autori e l'attenzione all'h-index degli autori nella scelta delle citazioni da includere nel proprio lavoro mostra che (i pochi che) danno valore all'h-index, lo usano anche come driver di scelta (come è anche vero il contrario, chi non gli attribuisce valore non lo usa neanche come driver di scelta).
- *Conoscenze*: la correlazione ($\rho = 0,5115906$; $p < 0.0001$) tra la valutazione di una lettura sulla base della conoscenza personale di uno degli autori, e sulla base del fatto che uno di essi sia un autore affermato e nel proprio campo, può suggerire che gli accademici mediamente cerchino di entrare in contatto diretto con coloro che ritengono sia un ricercatore molto conosciuto nella propria disciplina. Se questo fosse vero, sulle cause che portano a tale comportamento si possono fare solo delle ipotesi: è dovuto a semplice stima e ricerca di uno scambio di opinioni? Oppure alla speranza che nel lungo termine questo dia prestigio anche a se stessi?

Sono emerse altre correlazioni interessanti, ma nell'intervallo $0,4 < \rho < 0,5$, che ai fini di questa analisi non si ritengono particolarmente rilevanti.

L'analisi completa dei coefficienti di Pearson, invece, si trova nella Tab. 10 in Appendice. Tuttavia i risultati che emergono sono totalmente analoghi e concordi a quelli trovati con Spearman. Idem dicasi per l'analisi secondo il coefficiente di correlazione di Kendall (Kendall's τ), riportata in Tab. 11 in Appendice.

5.4 Test

5.4.1 Metodo

Poiché la prima analisi generale delle correlazioni non ha fatto emergere molte considerazioni, è stata fatta una ulteriore analisi di cross-tabulazione sulle domande che hanno sorpreso maggiormente. Per svolgerla i dati delle singole risposte sono stati importati, per semplice comodità d'analisi, su Qualtrics [57], un'altra piattaforma per gestire questionari online che, sebbene presenti limitazioni di gestione nella versione gratuita, permette di accedere a funzionalità di analisi comuni per i

questionari in modo più agevole rispetto ai comuni software statistici (sebbene per alcune altre analisi e verifiche sono stati importati anche in SPSS [58]).

Nelle cross-tabulazioni (o tavole di contingenza) i risultati di due variabili (domande) vengono incrociati, per mostrare in forma matriciale la distribuzione (multivariata) delle frequenze delle due variabili. Nel caso specifico interessava suddividere il campione in modo dicotomico nelle domande di interesse, e quindi i dati relativi a determinate risposte sono stati aggregati in modo che ne risultassero due partizioni distinte. Ad esempio, nella domanda sull'utilizzo dei motori e archivi online per la scelta del proprio topic di ricerca, rispetto all'affermazione "non li uso per scegliere il mio topic di ricerca" sono state aggregate le frequenze di coloro che hanno risposto con 'strongly agree' e 'weakly agree' a determinare la popolazione in generale accordo, e analogamente le frequenze di coloro che hanno risposto con 'strongly disagree' e 'weakly disagree' a determinare la popolazione in generale disaccordo.

Dal nostro punto di vista è stato sorprendente vedere l'incredibile utilizzo a cui vanno incontro Scholar e persino il motore generico di Google. Per contro, tutti gli altri hanno registrato un utilizzo che si stacca notevolmente dalle percentuali ottenute dai due motori Google: sembra quasi che il mondo si possa suddividere grossolanamente in chi usa strumenti di casa Google e chi no. Siccome la spartizione della popolazione negli utilizzatori o meno degli altri motori di ricerca ha condotto a campioni anche molto sbilanciati (e.g. 6,75% "every day" o "every week" su chi afferma di usare Scopus per trovare letture contro 93,25% di "every month" / "a few times in a year" / "never") si pensa possa portare a confronti potenzialmente imparziali per tali motori, mentre per Scholar e Google sarebbero forse più affidabili. Visto il loro successo rispetto agli altri si ritiene possano essere considerati buoni rappresentanti degli utilizzatori degli strumenti di ricerca attualmente disponibili. Oltretutto visto lo stacco che questi strumenti hanno registrato rispetto agli altri, è più probabile che diano una rappresentazione migliore dell'effetto che l'utilizzo di tali strumenti sta avendo sul mondo accademico. In ogni caso l'analisi ha coinvolto tutti i motori per motivi di completezza.

In particolare, dato che si voleva esplorare il comportamento degli utilizzatori, si sono prese come domande di riferimento quella sui metodi di reperimento delle letture e quella sull'utilizzo di motori e archivi online per la scelta del proprio topic di ricerca. Inoltre per entrambi gli utilizzi (letture e scelta del topic) si è voluto confrontare anche

le risposte riguardanti metodi più tradizionali di ricerca: l'utilizzo dei riferimenti delle pubblicazioni già lette nel caso di reperimento delle letture, e l'affermazione di non utilizzare proprio tali strumenti per la scelta del proprio topic di ricerca.

La codifica usata in questa analisi è inversa a quella naturale (e.g. 1='strongly agree', 4='strongly disagree') semplicemente a causa dell'ordine che si era dato alle risposte in fase di design del questionario su Qualtrics, ma questo non ha influenza sui risultati, che andranno banalmente letti in chiave opposta (e.g. risposta media vicina a 1 = risposta più vicina a 'strongly agree').

Dopodiché, nella cross-tabulazione risultante, sono stati eseguiti dei test specifici che potessero supportare le ipotesi. In particolare:

- Dei T-Test per confrontare le risposte medie date alle varie domande dai due campioni risultanti dalle spartizioni.
- Dei Test U di Mann-Whitney, simili ai T-Test ma non parametrici, per assicurarsi che i risultati fossero attendibili.
- Dei test χ^2 d'indipendenza per verificare l'esistenza di una relazione tra due domande e la polarità risultante nei due campioni generati dalla spartizione, come ulteriore rafforzativo delle affermazioni ottenute con i primi due test.

Prima di tutto si definiscono in generale i 3 test.

T-Test a doppio campione (o a campioni indipendenti)

Di ciascuna delle due popolazioni determinate dalla suddivisione (di ciascuna) delle domande di interesse, si è calcolata la risposta media (ovvero il valore medio risultante a seguito della codifica sopracitata) a tutte le altre domande. Successivamente si è effettuato un T-test a doppio campione (i.e. a campioni indipendenti) per determinare se la differenza nelle risposte medie date a ciascuna domanda dalle due popolazioni fosse statisticamente significativa o dovuta al caso (o più nello specifico, la probabilità con cui arrivino dalla stessa distribuzione o meno).

In generale in un T-test a doppio campione le ipotesi sono che le due popolazioni siano distribuite normalmente, abbiano la stessa varianza e siano indipendenti. Il fatto che siano indipendenti è garantito dalla spartizione, che per ogni domanda divide in due popolazioni completamente distinte. Il fatto che abbiano stessa varianza può influire in un diverso calcolo del test, che avverrebbe in un caso con una stima della varianza congiunta, nell'altro su due diverse stime di varianza. L'assunzione di

distribuzione approssimativamente normale potrebbe non essere verificata nel caso di domande del tipo Likert-scale, ed è il principale motivo per cui sono stati eseguiti altri test non parametrici per confermarne i risultati.

L'effettivo calcolo del valore di t avviene nel modo seguente nel caso di stima di varianza congiunta:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S_1^2 (n_1 - 1) + \frac{S_2^2 (n_2 - 1)}{n_1} - 1 + n_2 - 1}}$$

Mentre nel caso di stime distinte delle varianze:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Dove \bar{X}_i , S_i^2 sono media e varianza campionarie.

I risultati di questi test (laddove significativi, i.e. $p\text{-value} < 0,05$, e rilevanti per l'analisi) sono riportati nelle tabelle Tab. 12 e Tab. 14 in Appendice, rispettivamente per le domande sulle fonti di reperimento delle letture e per le domande sull'utilizzo di motori e archivi online nella scelta del proprio topic di ricerca. In esse sono riportate le risposte medie ottenute da ciascun campione nelle varie domande e i valori statistici relativi ai test.

Test U di Mann-Whitney

È un test non parametrico alternativo al T-test, spesso usato quando le assunzioni di quest'ultimo non sono verificate, la cui ipotesi nulla H_0 è che i due campioni arrivino dalla stessa popolazione, mentre l'alternativa è che una delle due distribuzioni sia stocasticamente più grande dell'altra (i.e. una delle due tenda ad avere valori più alti dell'altra). Assume che i valori dai due gruppi siano indipendenti, che siano ordinali, e le cui distribuzioni abbiano la stessa forma. Nel caso non avessero stessa forma il test ha ancora validità nell'affermare che i due gruppi arrivino da popolazioni distinte e quale dei due tenda ad avere valori più alti, mentre se avessero (anche approssimativamente) la stessa forma si può interpretare anche come test di differenza delle mediane dei due gruppi, fornendo anche una quantificazione

attendibile di tale differenza. In ogni caso una misura della differenza tra i due campioni si può misurare anche guardando la classifica media risultante in ciascuno di essi (ed è ciò che sarà fatto). Per il calcolo viene assegnato un rango a tutti i valori (da entrambi i gruppi), assegnando il punto medio di classifiche non rettificate in caso di pareggi (e.g. i ranghi di (3, 5, 5, 9) sono (1, 2.5, 2.5, 4)). Viene poi definito R_1 come somma dei ranghi dei campioni provenienti dal gruppo 1 (quindi la somma R_2 di quelli provenienti dal gruppi 2 è determinata, dato che la somma totale dei ranghi è uguale a $N(N+1)/2$). Infine i valori di U vengono calcolati come segue:

$$U_1 = R_1 - \frac{n_1(n_1 + 1)}{2} ; U_2 = R_2 - \frac{n_2(n_2 + 1)}{2}$$

con n_1 e n_2 le dimensioni dei due gruppi. Per il calcolo della significatività viene poi considerato il più piccolo dei due valori.

I risultati di questi test (laddove significativi, i.e. $p\text{-value} < 0,05$, e rilevanti per l'analisi) sono riportati nelle tabelle Tab. 16 e Tab. 17 in Appendice, rispettivamente per le domande sulle fonti di reperimento delle letture e per le domande sull'utilizzo di motori e archivi online nella scelta del proprio topic di ricerca. In esse sono riportate le classifiche medie ottenute da ciascun campione in ciascuna domanda, e il $p\text{-value}$ relativo ai test.

Test χ^2 d'indipendenza

È un test non parametrico la cui ipotesi nulla H_0 è che le due variabili categoriche in considerazione siano indipendenti. Si applica alle tabelle di contingenza e misura quanto le frequenze osservate si discostino dalle frequenze attese (i.e. quelle che si otterrebbero in modo puramente casuale, ma sempre 'sulla base' dei valori osservati). Ponendo che $O_{i,j}$ sia valore osservato mentre $E_{i,j}$ il valore atteso nella cella (i,j), si definisce il valore atteso di una cella come:

$$E_{i,j} = Np_i p_j$$

con N numero totale di osservazioni, $p_i = \frac{o_i}{N} = \frac{\sum_{j=1}^c o_{i,j}}{N}$ la frazione di osservazioni di tipo i ignorando l'attributo di colonna (frazione del totale di riga), $p_j = \frac{o_j}{N} = \frac{\sum_{i=1}^r o_{i,j}}{N}$ la frazione di osservazioni di tipo j ignorando l'attributo di riga (frazione del totale di colonna).

L'effettivo calcolo del valore χ^2 avviene come segue:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{i,j} - E_{i,j})^2}{E_{i,j}}$$

Dove r è il numero totale di righe, mentre c di colonne.

I risultati di questi test (laddove significativi, i.e. $p\text{-value} < 0,05$, e rilevanti per l'analisi) sono riportati nelle tabelle Tab. 13 e Tab. 15 in Appendice, rispettivamente per le domande sulle fonti di reperimento delle letture e per le domande sull'utilizzo di motori e archivi online nella scelta del proprio topic di ricerca. In esse sono riportate le differenze tra frequenze reali e frequenze attese (per fornire un'idea del trend che si sta verificando) e i valori statistici relativi ai test.

Poiché, come detto sopra, l'applicazione di un T-test in caso di violazione di una delle sue assunzioni può portare a valori inesatti di $p\text{-value}$, è stato eseguito anche il un alternativo test non parametrico equivalente, i.e. il test U di Mann-Whitney. Per farlo i dati relativi alle singole risposte sono stati importati nel software statistico SPSS, nel quale è stato effettivamente svolto tale test nei vari casi analizzati con le stesse aggregazioni.

Si anticipa già che l'analisi del test U di Mann-Whitney ha portato a considerazioni molto simili a quelle ottenute dal T-test (i.e. per la maggioranza delle domande considerate risultano significativi entrambi), in qualche modo confermandone i risultati anche in assenza di certezza sulla validità delle sue assunzioni. Per completezza sono stati riportati in Appendice i risultati di entrambi, ma nell'analisi sono riportate le considerazioni relative solo a Mann-Whitney, poiché è più adeguato al caso specifico (sebbene porti a considerazioni molto simili).

5.4.2 Analisi

5.4.2.1 Fonti di reperimento delle letture

La cross-tabulazione inerente le fonti di reperimento di letture è avvenuta dividendo la popolazione tre diversi modi, in base a come trovino letture interessanti inerenti al proprio topic di ricerca:

- Coloro che utilizzano Scholar frequentemente ('Every day' o 'Every week', e sono ~64%) rispetto a chi invece no ('Every month', 'A few times in a year' o 'Never').
- Coloro che utilizzano Google frequentemente ('Every day' o 'Every week', e sono ~54%) rispetto a chi invece no ('Every month', 'A few times in a year' o 'Never').
- Coloro che utilizzano Scopus frequentemente ('Every day' o 'Every week', e sono ~7%) rispetto a chi invece no ('Every month', 'A few times in a year' o 'Never').
- Coloro che utilizzano WoS frequentemente ('Every day' o 'Every week', e sono ~12%) rispetto a chi invece no ('Every month', 'A few times in a year' o 'Never').
- Coloro che utilizzano i motori verticali frequentemente ('Every day' o 'Every week', e sono ~24%) rispetto a chi invece no ('Every month', 'A few times in a year' o 'Never').
- Coloro che le selezionano frequentemente tra i riferimenti di papers che hanno già letto ('Every day' o 'Every week', e sono ~61%) rispetto a chi invece no ('Every month', 'A few times in a year' o 'Never'). Questo dovrebbe rappresentare un comportamento di ricerca più 'tradizionale'.

Tale suddivisione delle risposte per la determinazione dei cluster è dovuta al fatto che, essendo parte teoricamente predominante dell'attività di ricercatore, il metodo di reperimento delle letture ha un'influenza consistente sul proprio lavoro quando viene usato quasi quotidianamente. Se venisse usato anche 1 volta al mese o meno, implicherebbe che non copre un ruolo predominante nella propria attività di ricerca, sulla quale esercita quindi una bassa influenza.

Ciò che si vuole provare, è una differenza nei trend delle due popolazioni definite dalla spartizione di ognuna delle 6 domande sopra. Il controllo sui valori medi (sia la risposta media effettiva dei T-Test, che la classifica media di Mann-Whitney) tende ad 'appiattire' l'analisi, quindi anche se le due medie rimangono macroscopicamente nella stessa regione logica di risposta (e.g. 'strongly/weakly agree' o 'disagree'), la differenza tra di esse è rilevante, quando statisticamente significativa (i.e. con test positivo), perché sottintende una maggior polarizzazione di una delle due popolazioni

in un senso piuttosto che nell'altro. Va poi tenuto conto che la selezione degli intervistati è avvenuta dai Journal più importanti nel settore, quindi ci si aspetterebbe in ogni caso un comportamento medio di un certo livello. Infine le popolazioni che hanno risposto positivamente alle 6 diverse domande saranno necessariamente parzialmente sovrapposte (i.e. non usano solo Scholar o solo i riferimenti), e questo tenderà inevitabilmente ad attenuare i valori statistici (che acquisiscono quindi maggiore rilevanza, se nonostante questo si presentano differenze significative).

Di seguito si riporta una tabella (Tab. 6) riassuntiva dei risultati: sono indicati i Δ (classifica media di chi ha risposto "every day/week" - classifica media di chi ha risposto "every month/year/never") laddove il test di Mann-Whitney è risultato significativo (con un * se anche il test χ^2 è risultato significativo) e siccome la codifica numerica usata è opposta a quella naturale, $\Delta < 0$ implicano che le risposte dei "frequently" sono più vicine a "Strongly Agree"(=1)/"Always"(=1)/... (viceversa per i $\Delta > 0$). A seguire sono poi elencati i rilievi emersi dai test di Mann-Whitney, inserendo anche quelli relativi ai test χ^2 d'indipendenza laddove rafforzino il risultato con valori significativi:

	How do you find readings related to your research topics? - I use Google Scholar	How do you find readings related to your research topics? - I use Google's general purpose engine	How do you find readings related to your research topics? - I use Scopus	How do you find readings related to your research topics? - I use Web of Science	How do you find readings related to your research topics? - I use my field's vertical search engines	How do you find readings related to your research topics? - I select interesting readings among the references of papers
How would you define a "hot research topic"?-A research topic is hot if many researchers are working on it	-7,48%	-8,06%*	-	-	-	-
How would you define a "hot research topic"?-A research topic is hot if papers focusing on it are more likely to be cited	-	-	-	-	-	7,48%
How do you choose your research topics?-I do research on what is an "hot topic" in my fields	-	-7,63%	-	-	-	-
How do you choose your research topics?-I do research on topics that are most likely to be cited	-6,68%	-	-	-8,98%*	-	-
How do you choose your research topics?-I choose my research topics as a consequence of suggestions from peer/senior members of my research group	-	-	-	-	-8,25%*	-
How do you choose your research topics?-I choose my research topics based on the directions provided by the most impactful authors in my field	-8,29%*	-	-13,01%	-	-7,56%	-
Why do you move away from a research topic?-Because peer/senior members in my research group	-5,83%	-6,10%*	-	-8,63%*	-12,68%*	-

advised me to do so						
Why do you move away from a research topic?- Because of a personal loss of intellectual interest	-	-	-	-	-	-6,65%
What are the drivers that make you think that a publication is worth reading?-A publication is worth reading when it has a high number of citations	-6,73%	-	-19,90%*	-14,46%*	-7,41%	-
What are the drivers that make you think that a publication is worth reading?-A publication is worth reading when at least one of the authors has a high h-index	-	-	-16,72%*	-12,44%*	-	-
How do you choose the citations to be included in your own publications?-I choose citations based on the impact factor of their publication outlet	-	-	-	-11,41%*	-	-
How do you choose the citations to be included in your own publications?-I choose citations where at least one of the authors has a high h-index	-5,71%	-7,12%*	-14,69%*	-8,85%*	-7,08%*	-
How do you choose the citations to be included in your own publications?-I choose citations based on my target publication outlet	-8,55%*	-	-	-11,13%	-	-
How do you choose the citations to be included in your own publications?-I make an effort to include highly cited papers among my citations	-9,57%*	-	-	-11,40%*	-9,65%*	-
How do you choose the citations to be included in your own publications?-I cite authors who are most likely to cite me back	-8,51%*	-8,40%*	-	-9,88%*	-5,68%*	-

Tab. 6 Sommario dei risultati dei test di Mann-Whitney e χ^2 per il reperimento delle letture

1. Coloro che utilizzano gli attuali strumenti, rispetto a coloro che seguono i riferimenti di papers già letti, presentano un **comportamento più gregario** (e quindi meno indipendente). Si è rilevata una differenza statisticamente significativa nelle seguenti domande:

- Chi usa frequentemente Scholar (Δ nelle classifiche della media delle risposte di -7,48% tra chi lo usa di frequente e chi no) o Google (Δ nelle classifiche della media delle risposte di -8,06%) è più incline a pensare che un topic di ricerca sia “hot” se molti ricercatori ci stanno lavorando. Non è invece stata rilevata una differenza statisticamente significativa per chi usa i gli altri strumenti o riferimenti.
- Chi usa frequentemente Scholar (Δ di -8,29%), Scopus (Δ di -13,01%) o i motori verticali (Δ di -7,56%) è più incline a scegliere un topic di ricerca in base alle direzioni degli autori di maggiore impatto nel proprio campo. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o i riferimenti.

Tale riscontro è confermato per Scholar anche da un test χ^2 di indipendenza (p – value = 0,02) per il quale risulta esserci una relazione tra la domanda sull’utilizzo dello strumento e quella sull’assecondare gli autori d’impatto, con $\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ in risposte che rappresentano un maggiore

'agreement' per coloro che usano Scholar. Questo attesta che è probabile che esista una relazione tra le due variabili, il cui risultato è che chi usa Scholar è più influenzabile da tali autori.

- Chi usa frequentemente Scholar (Δ di -5,83%), Google (Δ di -6,10%), WoS (Δ di -8,63%) o i motori verticali (Δ di -12,68%) è più incline ad abbandonare un topic di ricerca in base alle direzioni dei membri peer o senior nel loro gruppo. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o i riferimenti.

Invece il test χ^2 di indipendenza risulta significativo per Google (p – value = 0,03), Scopus (p – value = 0,01), WoS (p – value = 0,01) e i motori verticali (p – value < 0,01), con $\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ in risposte che rappresentano un maggiore 'agreement' per gli utilizzatori di tali strumenti.

- Chi usa frequentemente i riferimenti è più incline (Δ di -13,4%) ad abbandonare un topic di ricerca per perdita di interesse intellettuale. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli strumenti online.
- Chi usa frequentemente i motori verticali (Δ di -8,25%) è più incline a scegliere un topic di ricerca in base alle direzioni dei membri peer o senior nel loro gruppo. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o i riferimenti.

2. Coloro che utilizzano gli attuali strumenti, rispetto a coloro che seguono i riferimenti di papers già letti, presentano un **comportamento più opportunistico**. Si è rilevata una differenza statisticamente significativa nelle seguenti domande:

- Chi usa frequentemente i riferimenti è meno incline (Δ di 7,48%) a pensare che un topic di ricerca sia "hot" se è più probabile che venga citato. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli strumenti online.
- Chi usa frequentemente Scholar (Δ di -6,68%) o WoS (Δ di -8,98%) è più incline a fare ricerca su topic che è più probabile che vengano citati. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o i riferimenti.

Il test χ^2 di indipendenza risulta significativo e conferma il trend per WoS (p – value = 0,02), con $\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ in risposte che rappresentano un maggiore frequenza.

- Chi usa frequentemente Scholar (Δ di -8,55%) o WoS (Δ di -11,13%) è più incline a scegliere le proprie citazioni in base al proprio outlet di pubblicazione (il Journal sul quale si intende pubblicare il lavoro). Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o i riferimenti.

Anche in questo caso si ha una ulteriore conferma per Scholar dal test χ^2 d'indipendenza (p – value = 0,04), per il quale quindi una relazione tra le due domande sembra sussistere. In particolare i $\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ sulle risposte 'frequently' o più per coloro che usano Scholar confermano il trend sottolineato dal T-test.

- Chi usa frequentemente Scholar (Δ di -8,51%), Google (Δ di -8,40%), WoS (Δ di -9,88%) o i motori verticali (Δ di -5,68%) è più incline citare autori che è più probabile che li ri-citino a loro volta. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o i riferimenti.

Conferma della relazione tra le due domande anche secondo un test χ^2 d'indipendenza, con $\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ sulle risposte che rappresentano una maggiore frequenza per coloro che usano Scholar e Google (sebbene alcune celle abbiano frequenza attesa minore di 5, che potrebbe portare a una minore attendibilità del test, sono comunque in numero minore o uguale alla soglia del 20%). Anche per WoS e i verticali è significativo ma il numero di celle con freq. attesa < 5 è maggiore della soglia del 20%, quindi non si considerano affidabili.

3. Coloro che utilizzano gli attuali strumenti, rispetto a coloro che seguono i riferimenti di papers già letti, risultano **più sensibili agli indicatori bibliometrici**.

Si è rilevata una differenza statisticamente significativa nelle seguenti domande:

- Chi usa frequentemente Scholar (Δ di -6,73%), Scopus (Δ di -19,90%), WoS (Δ di -14,46%) o i motori verticali (Δ di -7,41%) è più incline a pensare che una pubblicazione valga la pena di essere letta se ha un alto numero di citazioni. Non è invece stata rilevata una differenza statisticamente significativa per chi

usa gli altri strumenti o i riferimenti.

Il test χ^2 di indipendenza risulta significativo per Scopus (p – value < 0,01) e WoS (p – value = 0,01), con $\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ in risposte che rappresentano un maggiore frequenza.

- Chi usa frequentemente Scholar (Δ di -5,71%), Google (Δ di -7,12%), Scopus (Δ di -14,69%), WoS (Δ di -8,85%) o i motori verticali (Δ di -7,08%) è più incline a scegliere citazioni a pubblicazioni nelle quali uno degli autori abbia un alto h-index. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o i riferimenti. Per il test χ^2 laddove significativo si sono presentate celle con freq. attesa < 5 in numero maggiore della soglia del 20%.

- Chi usa frequentemente Scholar (Δ di -9,57%), WoS (Δ di -11,40%) o i motori verticali (Δ di -9,65%) è più incline a fare uno sforzo per inserire papers molto citati tra le proprie citazioni. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o i riferimenti.

Il test χ^2 di indipendenza risulta significativo per Scholar (p – value = 0,02), Scopus (p – value < 0,01), WoS (p – value < 0,01) e i verticali (p – value = 0,05), con $\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ in risposte che rappresentano un maggiore frequenza.

- Chi usa frequentemente Scopus (Δ di -16,72%) o WoS (Δ di -12,44%) è più incline a pensare che una pubblicazione valga la pena di essere letta se uno degli autori ha un alto h-index. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o i riferimenti.

Il test χ^2 di indipendenza risulta significativo per Scopus (p – value = 0,02) e WoS (p – value < 0,01), con $\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ in risposte che rappresentano un maggiore frequenza.

- Chi usa frequentemente WoS (Δ di -14,98%) è più incline a scegliere le citazioni in base all'IF del journal in cui è pubblicato l'articolo da citare. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o i riferimenti.

Il test χ^2 di indipendenza risulta significativo per WoS (p – value < 0,01), con

$\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ in risposte che rappresentano un maggiore frequenza.

5.4.2.2 Utilizzo di motori e archivi online per la scelta del topic di ricerca

L'analisi di cross-tabulazione è stata fatta anche rispetto a coloro che utilizzano motori di ricerca e archivi online per scegliere il proprio topic di ricerca (in particolare Google e Scholar perché sono quelli con un maggiore utilizzo) e chi invece no (o quasi). In particolare sono stati anche qui suddivisi in 3 modi:

- Coloro che *non* utilizzano proprio (o poco) tali strumenti per scegliere il proprio topic di ricerca ('strongly/weakly agree', e sono ~61%) rispetto a chi li utilizza ('strongly/weakly disagree'). Questo dovrebbe rappresentare una metodologia più "tradizionale" di ricerca.
- Coloro che utilizzano (anche poco) Google Scholar a tale scopo ('strongly/weakly agree', e sono ~57%) rispetto a chi invece no ('strongly/weakly disagree').
- Coloro che utilizzano (anche poco) Google a tale scopo ('strongly/weakly agree', e sono ~45%) rispetto a chi invece no ('strongly/weakly disagree').
- Coloro che utilizzano (anche poco) Scopus a tale scopo ('strongly/weakly agree', e sono ~14%) rispetto a chi invece no ('strongly/weakly disagree').
- Coloro che utilizzano (anche poco) WoS a tale scopo ('strongly/weakly agree', e sono ~24%) rispetto a chi invece no ('strongly/weakly disagree').
- Coloro che utilizzano (anche poco) i motori verticali a tale scopo ('strongly/weakly agree', e sono ~34%) rispetto a chi invece no ('strongly/weakly disagree').

Di seguito si riporta una tabella (Tab. 7) riassuntiva dei risultati: sono indicati i $\Delta(\text{classifica media di chi ha risposto "strongly/weakly agree"} - \text{classifica media di chi ha risposto "strongly/weakly disagree"})$ laddove il test di Mann-Whitney è risultato significativo (con un * se anche il test χ^2 è risultato significativo) e siccome la codifica numerica usata è opposta a quella naturale, $\Delta < 0$ implicano che le risposte degli "agree" sono più vicine a "Strongly Agree" (=1) / "Always" (=1) / ... (viceversa per i

$\Delta > 0$). A seguire sono poi elencati i rilievi emersi dai test di Mann-Whitney, inserendo anche quelli relativi ai test χ^2 d'indipendenza laddove rafforzino il risultato con valori significativi:

	Do you use on-line archives and search engines to choose a research topic? - I do not use them to choose a research topic	Do you use on-line archives and search engines to choose a research topic? - I use Google Scholar	Do you use on-line archives and search engines to choose a research topic? - I use Google's general purpose engine	Do you use on-line archives and search engines to choose a research topic? - I use Scopus	Do you use on-line archives and search engines to choose a research topic? - I use Web of Science	Do you use on-line archives and search engines to choose a research topic? - I use my field's vertical search engines
How would you define a "hot research topic"?-A research topic is hot if many researchers are working on it	-	-7,09%	-7,91%*	-	-	-
How would you define a "hot research topic"?-A research topic is hot if papers focusing on it are more likely to be cited	6,38%*	-	-	-	-	-
How do you choose your research topics?-I do research on what is an "hot topic" in my fields	-	-	-8,50%*	12,36%*	-7,39%*	-
How do you choose your research topics?-I do research on topics that are most likely to be cited	7,11%*	-7,34%*	-5,75%	-8,56%	-8,05%*	-
How do you choose your research topics?-I choose my research topics as a consequence of suggestions from peer/senior members of my research group	-	-6,69%*	-8,34%*	-	-	-8,81%*
How do you choose your research topics?-I choose my research topics based on the directions provided by the most impactful authors in my field	7,52%*	13,01%*	10,23%*	14,49%*	-7,69%*	10,45%*
Why do you move away from a research topic?-Because it is no longer a "hot topic" in my field	6,89%	-	-	-9,75%	-	-6,98%
Why do you move away from a research topic?-Because peer/senior members in my research group advised me to do so	6,93%*	10,51%*	11,66%*	-	-6,82%*	12,83%*
What are the drivers that make you think that a publication is worth reading?-A publication is worth reading when it has a high number of citations	-	-6,32%*	-5,89%	12,24%*	12,40%*	-9,45%*
What are the drivers that make you think that a publication is worth reading?-A publication is worth reading when at least one of the authors has a high h-index	-	-	-8,16%*	-	-	-
What are the drivers that make you think that a publication is worth reading?-A publication is worth reading when it is published in a journal that has a high impact factor	6,54%*	-	-	-	-9,13%*	-
How do you choose the citations to be included in your own publications?-I choose citations based on the impact factor of their publication outlet	8,73%*	-	-	-	-7,39%*	-
How do you choose the citations to be included in your own publications?-I choose citations where at least one of the authors has a high h-index	6,46%*	-5,71%	-7,90%*	11,54%*	-	-8,57%*
How do you choose the citations to be included in your own publications?-I choose citations based on my target publication outlet	-	-6,56%	-	-	-	-

How do you choose the citations to be included in your own publications?-I make an effort to include highly cited papers among my citations	-	-	-	-	-	-7,31%
---------------------------------------------------------------------------------------------------------------------------------------------	---	---	---	---	---	--------

Tab. 7 Sommario dei risultati dei test di Mann-Whitney e χ^2 per l'utilizzo degli attuali strumenti per la scelta del proprio topic di ricerca

1. Coloro che utilizzano gli attuali strumenti per scegliere il proprio topic di ricerca, rispetto a chi non li usa proprio (o quasi), presentano un **comportamento più gregario** (e quindi meno indipendente). Si è rilevata una differenza statisticamente significativa nelle seguenti domande:

- Chi usa Scholar (Δ di -7,09%) o Google (Δ di -7,91%) è più incline a pensare che un topic di ricerca sia “hot” se molti ricercatori ci stanno lavorando. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o chi sostiene di non usarli (o quasi) a tale scopo.
- Chi usa Scholar (Δ di -6,69%), Google (Δ di -8,34%) o i motori verticali (Δ di -8,81%) è più incline a scegliere un topic di ricerca in seguito a suggerimenti dei membri peer o senior nel loro gruppo. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o chi sostiene di non usarli (o quasi) a tale scopo.

Il test χ^2 di indipendenza risulta significativo per Scholar ($p - \text{value} = 0,02$), Google ($p - \text{value} < 0,01$) e i motori verticali ($p - \text{value} < 0,01$), con $\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ in risposte che rappresentano un maggiore frequenza (ci sono celle con freq. attesa minore di 5, ma sono in numero minore o uguale della soglia 20%).

- Chi usa Scholar (Δ di -13,01%), Google (Δ di -10,23%), Scopus (Δ di -14,49%), WoS (Δ di -7,69%) o i motori verticali (Δ di -10,45%) è più incline a scegliere un topic di ricerca in base alle direzioni degli autori di maggiore impatto nel proprio campo. E' stata anche rilevata una differenza statisticamente significativa opposta (i.e. meno incline a farlo) per chi sostiene di non usare (o quasi) questi strumenti a tale scopo (Δ di 7,52%).

Inoltre il test χ^2 d'indipendenza mostra che esiste una relazione tra questa domanda e gli strumenti. In particolare: chi non usa questi strumenti presenta $\Delta(\text{freq. reali} - \text{freq. attese}) < 0$ in risposte che rappresentano frequenze maggiori ($p - \text{value} = 0,02$), viceversa (i.e. trend opposto) per chi usa Scholar ($p - \text{value} < 0,01$) e Google ($p - \text{value} < 0,01$), Scopus ($p - \text{value} < 0,01$), WoS ($p - \text{value} = 0,01$) e i verticali ($p - \text{value} < 0,01$).

- Chi usa Scholar (Δ di -10,51%), Google (Δ di -11,66%), WoS (Δ di -6,82%) o i motori verticali (Δ di -12,83%) è più incline ad abbandonare un topic di ricerca in base alle direzioni dei membri peer o senior nel loro gruppo. E' stata anche rilevata una differenza statisticamente significativa opposta (i.e. meno incline a farlo) per chi sostiene di non usare (o quasi) questi strumenti a tale scopo (Δ di 6,93%), mentre non è risultata significativa per gli altri strumenti.
Inoltre il test χ^2 d'indipendenza mostra che esiste una relazione tra questa domanda e gli strumenti. In particolare: chi non usa questi strumenti presenta $\Delta(\text{freq. reali} - \text{freq. attese}) < 0$ in risposte che rappresentano maggiore 'agreement' (p - value = 0,03), viceversa, invece, per chi usa Scholar (p - value < 0,01), Google (p - value < 0,01), WoS (p - value < 0,04) o i verticali (p - value < 0,01).
2. Coloro che utilizzano gli attuali strumenti per scegliere il proprio topic di ricerca, rispetto a chi non lo fa proprio (o quasi), presentano un **comportamento più opportunistico**. Si è rilevata una differenza statisticamente significativa nelle seguenti domande:
- Chi non li usa proprio (o quasi) a tale scopo è meno incline (Δ di 6,38%) a pensare che un topic di ricerca sia "hot" se è più probabile che venga citato. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti.
Una relazione tra le due domande è confermata anche da un test χ^2 d'indipendenza (p - value = 0,02), con $\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ sulle risposte che rappresentano 'disagreement' per coloro che non usano questi strumenti a tale scopo.
 - Chi usa Scholar (Δ di -7,34%), Google (Δ di -5,75%), Scopus (Δ di -8,56%) o WoS (Δ di -8,05%) a tale scopo è più incline a fare ricerca su topic che è più probabile che vengano citati. Qui è stata anche rilevata una differenza statisticamente significativa opposta (i.e. meno incline a farlo) per chi sostiene di non usare (o quasi) questi strumenti a tale scopo (Δ di 7,11%), mentre non è risultata significativa per gli altri strumenti.
Una relazione con questa domanda si è rilevata anche con il test χ^2 d'indipendenza sia per coloro che non usano proprio (o quasi) questi strumenti (per i quali i $\Delta(\text{freq. reali} - \text{freq. attese}) < 0$ nelle risposte relative a frequenze

più alte) ($p - \text{value} = 0,05$) sia per coloro che usano Scholar ($p - \text{value} = 0,01$) che WoS ($p - \text{value} = 0,03$) (nel quale il trend dei Δ è opposto).

- Chi usa Scopus (Δ di -9,75%) o i motori verticali (Δ di -6,98%) è più incline ad abbandonare un topic di ricerca perché non è più “hot”. È stata anche rilevata una differenza statisticamente significativa opposta (i.e. meno incline a farlo) per chi sostiene di non usare (o quasi) questi strumenti a tale scopo (Δ di 6,89%), mentre non è risultata significativa per gli altri strumenti.
 - Chi usa Scholar (Δ di -6,56%) è più incline a scegliere le proprie citazioni in base al proprio outlet di pubblicazione (il Journal sul quale si intende pubblicare il lavoro). Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o chi sostiene di non usarli (o quasi) a tale scopo.
3. Coloro che utilizzano gli attuali strumenti per scegliere il proprio topic di ricerca, rispetto a chi non lo fa proprio (o quasi), risultano **più sensibili agli indicatori bibliometrici**. Si è rilevata una differenza statisticamente significativa nelle seguenti domande:
- Chi usa Scholar (Δ di -6,32%), Google (Δ di -5,89%), Scopus (Δ di -12,24%) o WoS (Δ di -12,40%) o i motori verticali (Δ di -9,45%) è più incline a pensare che una pubblicazione valga la pena di essere letta se ha un alto numero di citazioni. Non è invece stata rilevata una differenza statisticamente significativa per chi sostiene di non usarli (o quasi) a tale scopo.
Il test χ^2 d'indipendenza conferma una relazione tra la domanda e chi usa Scholar ($p - \text{value} = 0,02$), Scopus ($p - \text{value} = 0,02$), WoS ($p - \text{value} < 0,01$) e motori verticali ($p - \text{value} = 0,01$), con $\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ sulle risposte che rappresentano maggiore ‘agreement’.
 - Chi usa WoS (Δ di -7,39%) è più incline a scegliere le citazioni in base all’impact factor del Journal in cui è pubblicato l’articolo da citare. È stata anche rilevata una differenza statisticamente significativa opposta (i.e. meno incline a farlo) per chi sostiene di non usare (o quasi) questi strumenti a tale scopo (Δ di 8,73%), mentre non è risultata significativa per gli altri strumenti. Relazioni confermate dal test χ^2 d'indipendenza per chi usa WoS ($p - \text{value} = 0,02$), con $\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ in risposte che rappresentano una

maggior frequenza, viceversa per chi non usa questi strumenti a tale scopo ($p - \text{value} = 0,04$) hanno trend opposto.

- Chi usa Scholar (Δ di -5,71%), Google (Δ di -7,90%), Scopus (Δ di -11,54%) o i motori verticali (Δ di -8,57%) è più incline a scegliere citazioni a pubblicazioni nelle quali uno degli autori abbia un alto h-index. È invece stata rilevata una differenza statisticamente significativa opposta per chi sostiene di non usare (o quasi) questi strumenti a tale scopo (Δ di 6,46%), mentre non è risultata significativa per gli altri strumenti.

Una relazione con la domanda è confermata dal test χ^2 d'indipendenza sia per coloro che non usano questi strumenti ($p - \text{value} = 0,02$, $\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ in risposte che rappresentano frequenze minori), sia, Δ con trend opposto, per chi usa Google ($p - \text{value} = 0,03$), WoS ($p - \text{value} = 0,03$) e i verticali ($p - \text{value} = 0,01$) (per Scopus è significativo ma le freq. attese < 5 superano la soglia).

- Chi usa Google (Δ di -8,16%) è più incline a pensare che una pubblicazione valga la pena di essere letta se almeno uno degli autori ha un alto h-index. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o chi sostiene di non usarli (o quasi) a tale scopo.
- Chi usa WoS (Δ di -9,13%) è più incline a pensare che una pubblicazione valga la pena di essere letta se è in un journal con un alto IF. È invece stata rilevata una differenza statisticamente significativa opposta per chi sostiene di non usare (o quasi) questi strumenti a tale scopo (Δ di 6,54%), mentre non è risultata significativa per gli altri strumenti.

Il test χ^2 d'indipendenza conferma una relazione tra la domanda e chi usa WoS ($p - \text{value} = 0,03$), con $\Delta(\text{freq. reali} - \text{freq. attese}) > 0$ sulle risposte che rappresentano maggiore 'agreement', e un trend opposto per coloro che sostengono di non usare questi strumenti a tale scopo ($p - \text{value} = 0,05$).

- Chi usa i motori verticali (Δ di -7,31%) è più incline a sforzarsi di inserire papers altamente citati tra le proprie citazioni. Non è invece stata rilevata una differenza statisticamente significativa per chi usa gli altri strumenti o chi sostiene di non usarli (o quasi) a tale scopo.

5.5 Conclusione

Complessivamente, l'analisi ha dato buoni risultati.

In generale ha dato risposta alle domande di ricerca che ci si era posti, e ha portato a concludere che gli sforzi per future implementazioni possano concentrarsi sull'applicazione dell'analisi testuale dei paper accademici delle conferenze allo scopo di identificare gli attuali "hot topics" del proprio campo scientifico.

Inoltre ha verificato le ipotesi formulate, vedendo una maggiore inclinazione degli utilizzatori degli attuali strumenti verso comportamenti di carattere vizioso.

6 Conclusioni

Questo lavoro ha portato a rilievi per certi versi contrastanti. Da un lato il problema del reperimento di informazioni più rilevanti nella documentazione scientifica sembra un problema sentito dagli accademici. Lo dimostra la straripante letteratura riguardante gli indicatori bibliometrici e i motori di ricerca, che vede tanti criticismi, ma anche molti articoli di carattere propositivo su possibili alternative e miglioramenti. Dall'altro emerge uno stato di generale "inerzia" ovvero, nonostante la consapevolezza delle problematiche esistenti, sembra esserci globalmente una scarsa propensione al cambiamento. Questo emerge in primo luogo dalla bassa attenzione generale che hanno riscosso le alternative proposte in letteratura, in secondo luogo dai risultati stessi del questionario di questo lavoro. È infatti emerso che gli accademici:

- Non usano quasi per nulla funzionalità avanzate come l'analisi dei risultati di Scopus, o i servizi di alerting forniti da Scopus stesso e dai social network accademici.
- Non sfruttano, a differenza di quanto suggerito dai dati di Alexa.com, i social network accademici, quindi una direzione in questo senso non sembrerebbe promettente.
- Sembrano usare in schiacciante maggioranza Scholar (ma persino il generico Google) rispetto agli altri strumenti a disposizione, sia nella scelta del proprio argomento di ricerca sia nel reperimento delle letture, sebbene sia probabilmente il più limitato.
- Oltretutto valutano incredibilmente efficace la ricerca avanzata di Scholar stesso, indicando una soddisfazione diffusa della semplicità che espone questo strumento.

In definitiva sembra quindi che il mondo accademico non presenti il bisogno che si era immaginato di uno strumento più avanzato che permetta di recuperare contenuti più rilevanti in meno tempo, e che piuttosto si ritenga globalmente soddisfatto della semplicità di utilizzo di uno strumento come Scholar. Sembrerebbe quindi vano lo sforzo implementativo di uno strumento che cerchi di sopperire a una lacuna che, evidentemente, non è percepita come tale. In compenso, tuttavia, è sembrato

riscuotere un notevole interesse un potenziale strumento che raccolga automaticamente i topic più frequenti della letteratura recente. Sembrerebbe dunque più promettente il futuro sviluppo di uno strumento di identificazione degli “hot topic”, la cui progettazione può essere guidata dai risultati ottenuti dal presente questionario nel capire quali definizioni gli accademici ritenessero più vicine al loro concetto di “hot”.

È poi emerso che l'utilizzo degli attuali strumenti disponibili eserciti effettivamente un'influenza negativa sui comportamenti potenzialmente viziosi che possono esibire gli accademici, in quanto i suoi utilizzatori hanno presentato una maggiore inclinazione verso di essi. Quindi, nonostante un nuovo strumento non sembri andare incontro alle esigenze degli accademici, va comunque considerato che potrebbe presentare un'influenza minore su questo genere di abitudini, se solo andasse incontro a un effettivo utilizzo che permettesse di validare quest'idea.

7 Bibliografia

- [1] E. Garfield, "Journal impact factor: A brief review.", *Canadian Medical Association Journal*, vol. 161, no. 8, pp. 979–980, 1999.
- [2] L. Page, S. Brin, R. Motowani, and T. Winograd, "The PageRank Citation Ranking: Bringing Order to the Web (Tech. Rep.)", *Stanford Digital Library Technologies Project*, 1999.
- [3] G. Pinski and F. Narin, "Citation influence for journal aggregates of scientific publications: theory, with application to the literature of physics.", *Information Processing and Management*, vol. 12, no. 5, pp. 297–312, 1976.
- [4] EigenFactor. (2007) Eigenfactor Score and Article Influence Score: Detailed methods. [Online]. <http://www.eigenfactor.org/methods.pdf>
- [5] J. E. Hirsch, "An index to quantify an individual's scientific research output.", *Proceedings of the National Academy of Sciences*, vol. 102, no. 46, pp. 16569–16572, 2005.
- [6] M. Amin and M. Mabe, "Impact Factors: Use & Abuse", *Medicina (Buenos Aires)*, vol. 63, no. 4, pp. 347-345, 2003.
- [7] P. Gray, "Journal Self-Citation I: Overview of the Journal Self-Citation Papers-The Wisdom of the IS Crowd.", *Communications of the Association for Information Systems*, vol. 25, no. 1, p. 1, 2009.
- [8] J. Hartley, "To cite or not to cite: author self-citations and the impact factor.", *Scientometrics*, vol. 92, no. 2, pp. 313-317, 2011.
- [9] G. Buela-Casal, I. Zych, and A. Medina, "Analysis of the influence of the two types of the journal articles; theoretical and empirical on the impact factor of a journal.", *Scientometrics*, vol. 80, no. 1, pp. 265-282, 2009.
- [10] S. Bauin and H. Rothman, "'Impact' of journals as proxies for citation counts.", *Leiden: DSWO Press*, vol. 39, p. 225, 1992.
- [11] F. Narin and K.S. Hamilton, "Bibliometric performance measures.", *Scientometrics*, vol. 36, no. 3, pp. 293-310, 1996.
- [12] L. Lange, "Effects of disciplines and countries on citation habits. An analysis of empirical papers in behavioural sciences.", *Scientometrics*, vol. 8, no. 3-4, pp. 205-215, 1985.
- [13] A.R. Weale, M. Bailey, and P.A. Lear, "The level of non-citation of articles within a journal as a measure of quality: a comparison to the impact factor.", *BMC medical research methodology*, vol. 4, no. 1, p. 14, 2004.
- [14] P.O. Seglen, "Why the impact factor of journals should not be used for evaluating research.", *BMJ*, vol. 314, no. 7079, pp. 498-502, 1997.
- [15] T. Misteli, "Eliminating the impact of the Impact Factor", *The Journal of Cell Biology*, vol. 201, no. 5, pp. 651-652, 2013.
- [16] J.D. Dimitrov, S.V. Kaveri, and J. Bayry, "Metrics: journal's impact factor skewed by a single paper.", *Nature* 466, 179, 2010., vol. 466, no. 179, 2010.
- [17] K. Moustafa, "The disaster of the impact factor", *Science and engineering ethics*, vol. 21, no. 1, pp. 139-142, 2014.
- [18] I. Rafols, L. Leydesdorff, A. O'Hare, P. Nightingale, and A. Stirling, "How journal

- rankings can suppress interdisciplinary research: A comparison between innovation studies and Business & Management", *Research Policy*, vol. 41, no. 7, pp. 1262-1282, 2012.
- [19] J. Last. Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans. [Online]. <http://www.ncehr-cnerh.org/english/communiqué2/Negative.html>
- [20] M. E. Falagas and V. G. Alexiou, "The top-ten in journal impact factor manipulation", *Arch Immunol Ther Exp (Warsz)*, vol. 56, no. 4, pp. 223-226, 2008.
- [21] H.F. Moed, "Measuring contextual citation impact of scientific journals", *Journal of Informetrics*, vol. 4, no. 3, pp. 265-277, 2010.
- [22] Johan Bollen, Marko A. Rodriguez, and Herbert Van De Sompel, "Journal status", *Scientometrics*, vol. 69, no. 3, pp. 669-687, 2006.
- [23] J. Rizkallah and D.D. Sin, "Integrative approach to quality assessment of medical journals using impact factor, eigenfactor, and article influence scores.", *PLoS One*, vol. 5, no. 4, 2010.
- [24] SCImago Journal Rank. Description of SCImago Journal Rank Indicator. [Online]. <http://www.scimagojr.com/SCImagoJournalRank.pdf>
- [25] Tibor Braun, Wolfgang Glanzel, and Andràs Schubert, "A Hirsch-type index for journals", *Scientometrics*, vol. 69, no. 1, pp. 169-173, 2006.
- [26] L.A. Zhivotovsky and K.V. Krutovsky, "Self-citation can inflate h-index.", *Scientometrics*, vol. 77, no. 2, pp. 373-375.
- [27] C. Bartneck and S. Kokkelmans, "Detecting h-index manipulation through self-citation analysis.", *Scientometrics*, vol. 87, no. 1, pp. 85-98.
- [28] M.C. Wendl, "H-index: however ranked, citations need context.", *Nature*, vol. 449, no. 7161, p. 403, 2007.
- [29] L. Bornmann and H.D. Daniel, "What do we know about the h index?", *Journal of the American Society for Information Science and Technology*, vol. 58, no. 9, pp. 1381-1385, 2007.
- [30] L. Bornmann, "H-index Research in Scientometrics: A Summary", *Journal of Informetrics*, vol. 8, no. 3, pp. 749-750, 2014.
- [31] R. Costas and M. Bordons, "The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level.", *Journal of Informetrics*, vol. 1, no. 3, pp. 193-203, 2007.
- [32] L. Egghe, "An improvement of the h-index: The g-index", *ISSI newsletter*, vol. 2, no. 1, pp. 8-9, 2006.
- [33] L. Egghe, "Theory and practise of the g-index", *Scientometrics*, vol. 69, no. 1, pp. 131-152, 2006.
- [34] B. Jin, "h-index: an evaluation indicator proposed by scientist.", *Science Focus*, vol. 1, no. 1, pp. 8-9, 2006.
- [35] L. Bornmann, R. Mutz, and H.D. Daniel., "Are there better indices for evaluation purposes than the h index? A comparison of nine different variants of the h index using data from biomedicine.", *Journal of the American Society for Information Science and Technology*, vol. 59, no. 5, pp. 830-837, 2008.
- [36] B. Jin, L. Liang, R. Rousseau, and L. Egghe, "The R- and AR- indices: Complementing the h-index.", *Chinese Science Bulletin*, vol. 52, no. 6, pp. 855-863, 2007.
- [37] M. Schreiber, "A modification of the h-index: The hm-index accounts for multi-authored manuscripts", *Journal of Informetrics*, vol. 2, no. 3, pp. 211-216, 2008.

- [38] A. Noruzi, "Google Scholar: The new generation of Citation Indexes", *Libri*, vol. 55, no. 4, pp. 170-180, 2005.
- [39] P. Jacsó, "Google Scholar: the pros and the cons", *Online Information Review*, vol. 29, no. 2, pp. 208-214, 2005.
- [40] M. Shultz, "Comparing test searches in PubMed and Google Scholar", *Journal of the Medical Library Association*, vol. 95, no. 4, p. 442, 2007.
- [41] P. Jacsó, "Google Scholar revisited", *Online Information Review*, vol. 32, no. 1, pp. 102-114, 2008.
- [42] J. Bar-Ilan, "Which h-index? – A comparison of WoS, Scopus and Google Scholar", *Scientometrics*, vol. 74, no. 2, pp. 257-271, 2008.
- [43] P. Mayr and A. Walter, "An exploratory study of Google Scholar", *Online information review*, vol. 31, no. 6, pp. 814-830, 2007.
- [44] K. Yang and L. Meho, "Citation analysis: A comparison of Google Scholar, Scopus and Web of Science", *Proceedings of the American Society for Information Science and Technology*, vol. 43, no. 1, pp. 1-15, 2006.
- [45] P. Jacsó, "The pros and cons of computing the h-index using Google Scholar", *Online Information Review*, vol. 32, no. 3, pp. 437-452, 2008.
- [46] P. Jacsó, "As we may search—comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases", *Current Science*, vol. 89, no. 9, pp. 1537-1547, 2005.
- [47] Y. Gavel and L. Iselid, "Web of Science and Scopus: a journal title overlap study", *Online information review*, vol. 32, no. 1, pp. 8-21, 2008.
- [48] E. Vieira and J. Gomes, "A comparison of Scopus and Web of Science for a typical university", *Scientometrics*, vol. 81, no. 2, pp. 587-600, 2009.
- [49] J.F. Burnham, "Scopus database: a review", *Biomedical Digital Libraries*, vol. 3, no. 1, p. 1, 2006.
- [50] N. Bakkalbasi, K. Bauer, J. Glover, and L. Wang, "Three options for citation tracking: Google Scholar, Scopus and Web of Science.", *Biomedical Digital Libraries*, vol. 3, no. 1, p. 7, 2006.
- [51] M.A. Penagos, *Master Thesis - First approach for a Content-based impact of academic research.*: Politecnico di Milano.
- [52] Scopus. (2014) Elsevier. [Online]. http://www.elsevier.com/_data/assets/pdf_file/0007/69451/sc_content-coverage-guide_july-2014.pdf
- [53] Thomson Reuters. (2010) Web of Science Coverage Expansion. [Online]. <http://community.thomsonreuters.com/t5/Citation-Impact-Center/Web-of-Science-Coverage-Expansion/ba-p/10663>
- [54] M. Baez, A. Birukou, F. Casati, and M. Marchese, "Addressing Information Overload in the Scientific Community", *IEEE Computer Society*, vol. 14, no. 6, pp. 31-38, 2010.
- [55] SurveyFace. [Online]. www.surveyface.com
- [56] R Foundation. R. [Online]. www.r-project.org
- [57] Qualtrics. [Online]. www.qualtrics.com
- [58] IBM. IBM SPSS Statistics. [Online]. <http://www-01.ibm.com/software/it/analytics/spss/>

8 Appendice

Tab. 8 Risultati globali del questionario

	2. How would you define a "hot research topic"?			
	Strongly agree	Weakly agree	Weakly disagree	Strongly disagree
1. A research topic is hot if many researchers are working on it	38.65% (126)	48.47% (158)	9.51% (31)	3.37% (11)
2. A research topic is hot if many companies are investing in research on it	31.9% (104)	47.55% (155)	14.72% (48)	5.83% (19)
3. A research topic is hot if it is included in many calls for funded research	38.04% (124)	46.93% (153)	11.96% (39)	3.07% (10)
4. A research topic is hot if it is the focus of many research publications	32.21% (105)	46.63% (152)	18.1% (59)	3.07% (10)
5. A research topic is hot if papers focusing on it are more likely to be cited	26.99% (88)	43.25% (141)	24.54% (80)	5.21% (17)
6. A research topic is hot if it provides more opportunities to cooperate with other researchers	17.79% (58)	40.18% (131)	34.05% (111)	7.98% (26)
	3. How do you choose your research topics?			
	Always	Frequently	Seldom	Never
1. I do research on what I think is a "hot topic" in my field	5.52% (18)	40.18% (131)	48.47% (158)	5.83% (19)
2. I do research on topics that are most likely to be cited	2.15% (7)	27.61% (90)	55.21% (180)	15.03% (49)
3. My research topics are a consequence of research contracts with companies	2.76% (9)	21.17% (69)	41.41% (135)	34.66% (113)
4. My research topics are a consequence of peer reviewed funded projects	6.44% (21)	28.83% (94)	41.72% (136)	23.01% (75)
5. I choose my research topics as a consequence of suggestions from peer/senior members of my research group	2.76% (9)	26.69% (87)	48.77% (159)	21.78% (71)
6. I choose my research topics according to my personal research interests	70.25% (229)	28.22% (92)	1.53% (5)	0% (0)
7. I choose my research topics according to cooperation opportunities with researchers outside of my research field	2.45% (8)	33.44% (109)	51.23% (167)	12.88% (42)
8. I choose my research topics according to cooperation	7.98% (26)	57.98% (189)	27.61% (90)	6.44% (21)

opportunities with other researchers in my research field				
9. I choose my research topics based on the directions provided by the most impactful authors in my field	2.45% (8)	23.01% (75)	48.16% (157)	26.38% (86)
	4. How frequently do you change your set of research topics?			
	Every year		4.29% (14)	
	Every 2 years		10.43% (34)	
	Every 3 years		22.39% (73)	
	Every 4 years		18.1% (59)	
	Every 5 years		21.47% (70)	
	Between 5 and 10 years		21.47% (70)	
	Never		1.84% (6)	
	5. Why do you move away from a research topic?			
	Strongly agree	Weakly agree	Weakly disagree	Strongly disagree
1. Because it is no longer a "hot topic" in my field	5.83% (19)	27.61% (90)	39.57% (129)	26.99% (88)
2. Because the related project has ended	27.61% (90)	46.32% (151)	17.18% (56)	8.9% (29)
3. Because of lack of funds	7.06% (23)	28.53% (93)	31.9% (104)	32.52% (106)
4. Because peer/senior members in my research group advised me to do so	3.37% (11)	10.43% (34)	27.61% (90)	58.59% (191)
5. Because I plan to move to a different research institution	3.68% (12)	15.64% (51)	21.17% (69)	59.51% (194)
6. Because I plan to move to a different research group	4.6% (15)	16.87% (55)	23.31% (76)	55.21% (180)
7. Because of a personal loss of intellectual interest	55.52% (181)	33.13% (108)	6.13% (20)	5.21% (17)
	6. Do you use online archives and search engines to choose a research topic?			
	Strongly agree	Weakly agree	Weakly disagree	Strongly disagree
1. I do not use them to choose a research topic	41.41% (135)	19.33% (63)	19.63% (64)	19.63% (64)
2. I use Google Scholar	30.67% (100)	26.69% (87)	14.11% (46)	28.53% (93)
3. I use Google's general purpose search engine	18.4% (60)	27.3% (89)	19.33% (63)	34.97% (114)
4. I use Scopus	3.68% (12)	10.12% (33)	17.48% (57)	68.71% (224)
5. I use the analytic functionalities of Scopus ("Analyze search results" button)	1.53% (5)	4.29% (14)	17.79% (58)	76.38% (249)
6. I use Web of Science	7.98% (26)	16.26% (53)	17.79% (58)	57.98% (189)
7. I use ResearchGate.net	2.15% (7)	15.34% (50)	26.38% (86)	56.13% (183)
8. I use Academia.edu	1.23% (4)	6.75% (22)	20.55% (67)	71.47% (233)

9. I use my field's vertical search engines (e.g. IEEE for engineering, Pubmed for medicine, etc.)	14.72% (48)	19.02% (62)	14.11% (46)	52.15% (170)	
	7. Do you think that it would be useful to have a tool that allows you to automatically gather the most frequent topics in scientific papers?				
	Yes	May be yes	May be no	No	
	26.07% (85)	0,5 (163)	11.66% (38)	12.27% (40)	
	8. How do you find readings related to your research topics? - Conferences				
	More than 4 times a year	2-4 times a year	Once a year	Less that once a year	Never
	10.43% (34)	49.39% (161)	29.14% (95)	11.04% (36)	0% (0)
2. I serve as a reviewer for conferences in my field	29.14% (95)	45.4% (148)	17.79% (58)	4.91% (16)	2.76% (9)
	9. How do you find readings related to your research topics? - I serve as an editor for journals in my field				
	More than 3	Three	Two	One	None
	9.2% (30)	8.59% (28)	16.26% (53)	24.85% (81)	41.1% (134)
	10. How do you find readings related to your research topics? - Other sources				
	Every day	Every week	Every month	A few times in a year	Never
1. I use Google Scholar	21.17% (69)	43.25% (141)	19.33% (63)	10.74% (35)	5.52% (18)
2. I use Google's general purpose search engine	22.7% (74)	30.98% (101)	21.47% (70)	14.11% (46)	10.74% (35)
3. I use Scopus	1.53% (5)	5.21% (17)	11.66% (38)	20.86% (68)	60.74% (198)
4. I use Web of Science	2.76% (9)	9.2% (30)	15.03% (49)	27.91% (91)	45.09% (147)
5. I use ResearchGate.net	0.31% (1)	10.12% (33)	17.79% (58)	26.38% (86)	45.4% (148)
6. I use Academia.edu	0% (0)	1.23% (4)	7.98% (26)	18.1% (59)	72.7% (237)
7. I use my field's vertical search engines (e.g. IEEE for engineering, Pubmed for medicine, etc.)	4.6% (15)	19.02% (62)	18.1% (59)	19.33% (63)	38.96% (127)
8. I select interesting readings among the references of papers that I have read	17.48% (57)	43.87% (143)	25.46% (83)	10.43% (34)	2.76% (9)
9. I receive paper versions of the relevant journals in my field	3.99% (13)	10.12% (33)	17.79% (58)	23.31% (76)	44.79% (146)
	11. How do you rate the effectiveness of the following tools to search for readings related to your research topics?				
	Excellent	Good	Fair	Bad	I never use it
1. Google Scholar's advanced search	31.9% (104)	47.85% (156)	13.19% (43)	1.84% (6)	5.21% (17)
2. Scopus advanced search	3.68% (12)	15.95% (52)	12.88% (42)	1.84% (6)	65.64% (214)
3. Web of Science's advanced search	6.44% (21)	19.33% (63)	21.47% (70)	4.29% (14)	48.47% (158)
4. The advanced search of your field's vertical search engines (e.g. IEEE for engineering, Pubmed for medicine, etc.)	10.43% (34)	25.77% (84)	20.86% (68)	2.45% (8)	40.49% (132)

5. ResearchGate.net	0.92% (3)	10.43% (34)	26.69% (87)	5.21% (17)	56.75% (185)
6. Academia.edu	0.61% (2)	5.52% (18)	12.88% (42)	2.15% (7)	78.83% (257)
	12. What type of readings do you look for?				
	Always	Frequently	Seldom	Never	
1. Academic papers	85.58% (279)	14.11% (46)	0.31% (1)	0% (0)	
2. Company white papers	2.76% (9)	21.17% (69)	52.45% (171)	23.62% (77)	
3. Patents	0% (0)	1.84% (6)	22.7% (74)	75.46% (246)	
4. Slides	1.53% (5)	16.26% (53)	50.92% (166)	31.29% (102)	
5. Videos	1.53% (5)	11.66% (38)	48.47% (158)	38.34% (125)	
	13. What are the characteristics of your readings?				
	Always	Frequently	Seldom	Never	
1. They discuss the results of theoretical research	33.44% (109)	54.29% (177)	11.96% (39)	0.31% (1)	
2. They discuss the results of empirical research	33.13% (108)	63.19% (206)	3.68% (12)	0% (0)	
3. They describe case studies	11.35% (37)	50.92% (166)	32.82% (107)	4.91% (16)	
	14. Do you use push services that alert you when new research is available that might be of interest for you?				
	Always	Frequently	Seldom	Never	
1. I use Scopus alerting services that let me know when there is a new publication that matches one of my searches	2.45% (8)	3.68% (12)	9.51% (31)	84.36% (275)	
2. I use ResearchGate alerting services that let me know when a researcher in my network has uploaded a new publication	6.75% (22)	11.96% (39)	16.87% (55)	64.42% (210)	
3. I use Academia.edu alerting services that let me know when a researcher in my network has uploaded a new publication	1.23% (4)	1.84% (6)	11.66% (38)	85.28% (278)	
	15. What are the drivers that make you think that a publication is worth reading?				
	Strongly agree	Weakly agree	Weakly disagree	Strongly disagree	
1. A publication is worth reading when it has a high number of citations	22.7% (74)	49.08% (160)	19.94% (65)	8.28% (27)	
2. A publication is worth reading when it tackles a practical problem	34.36% (112)	49.39% (161)	14.11% (46)	2.15% (7)	
3. A publication is worth reading when it tackles a theoretical problem	38.96% (127)	50% (163)	8.59% (28)	2.45% (8)	
4. A publication is worth reading when it is closely related to your research topics	80.98% (264)	16.56% (54)	2.15% (7)	0.31% (1)	
5. A publication is worth reading when it presents research results that are practically applicable	46.32% (151)	39.88% (130)	11.66% (38)	2.15% (7)	

and useful				
6. A publication is worth reading when at least one of the authors has a high h-index	4.91% (16)	21.47% (70)	32.21% (105)	41.41% (135)
7. A publication is worth reading when at least one of the authors is a well known established researcher in your field	15.34% (50)	48.47% (158)	26.99% (88)	9.2% (30)
8. A publication is worth reading when I personally know at least one of the authors	13.19% (43)	40.8% (133)	28.22% (92)	17.79% (58)
9. A publication is worth reading when at least one of the authors belongs to my research circle in either ResearchGate or Academia.edu	2.45% (8)	18.71% (61)	23.62% (77)	55.21% (180)
10. A publication is worth reading when it is published in a journal that has a high impact factor	22.7% (74)	46.63% (152)	20.25% (66)	10.43% (34)
	16. How do you choose the citations to be included in your own publications?			
	Always	Frequently	Seldom	Never
1. I choose citations based on their closeness to my research topics	62.27% (203)	33.13% (108)	3.37% (11)	1.23% (4)
2. I choose citations that provide important evidence to support specific statements in my publication	73.31% (239)	24.54% (80)	1.23% (4)	0.92% (3)
3. I choose citations based on my personal knowledge of at least one of the authors	3.99% (13)	19.33% (63)	45.4% (148)	31.29% (102)
4. I choose citations based on the impact factor of their publication outlet	7.36% (24)	32.52% (106)	31.29% (102)	28.83% (94)
5. I choose citations where at least one of the authors has a high h-index	0.61% (2)	6.44% (21)	25.46% (83)	67.48% (220)
6. I choose citations based on my target publication outlet	13.19% (43)	43.87% (143)	30.98% (101)	11.96% (39)
7. I make an effort to include highly cited papers among my citations	7.98% (26)	28.53% (93)	35.28% (115)	28.22% (92)
8. I cite authors who are most likely to cite me back	1.23% (4)	5.52% (18)	18.71% (61)	74.54% (243)
	17. How do you work to disseminate your own research work?			
	Always	Frequently	Seldom	Never

1. I personally go to the conferences where my papers are published	30.37% (99)	51.84% (169)	14.72% (48)	3.07% (10)
2. I reach out to other researchers in my field on ResearchGate	3.68% (12)	11.35% (37)	30.37% (99)	54.6% (178)
3. I reach out to other researchers in my field on Academia.edu	1.23% (4)	2.45% (8)	15.03% (49)	81.29% (265)
4. I create a network of contacts through cooperation in common research projects	11.66% (38)	44.79% (146)	25.15% (82)	18.4% (60)
5. I circulate my own papers before publication to get early feedback	8.59% (28)	32.21% (105)	39.88% (130)	19.33% (63)
6. I circulate my own papers before publication to engage other researchers	4.91% (16)	20.55% (67)	44.79% (146)	29.75% (97)

Tab. 9 Coefficienti di correlazione di Spearman e relativi p-value

Question 1 (row)	Question 2 (column)	Spearman rho	p-value
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	0,788300753	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use my field s vertical search engines. e.g. IEEE for engineering. Pubmed for medicine. etc.	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - The advanced search of your field s vertical search engines. e.g. IEEE for engineering. Pubmed for medicine. etc.	0,772161067	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Scopus	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	0,769858837	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Web of Science	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Web of Science s advanced search	0,766332388	<2.22E-16
Q5. Why do you move away from a research topic - Because I plan to move to a different research institution	Q5. Why do you move away from a research topic - Because I plan to move to a different research group	0,754753888	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	0,744656324	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Google Scholar	Q6. Do you use online archives and search engines to choose a research topic - I use Google s general purpose search engine	0,728464842	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Academia.edu	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Academia.edu	0,722290576	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q10. How do you find readings related to your research topics - Other sources. I use Scopus	0,720783293	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use my field s vertical search engines. e.g. IEEE for engineering.	Q10. How do you find readings related to your research topics - Other sources. I use my field s vertical search engines. e.g. IEEE for engineering. Pubmed for	0,705643535	<2.22E-16

Pubmed for medicine. etc.	medicine. etc.		
Q17. How do you work to disseminate your own research work - I circulate my own papers before publication to get early feedback	Q17. How do you work to disseminate your own research work - I circulate my own papers before publication to engage other researchers	0,696071982	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	Q10. How do you find readings related to your research topics - Other sources. I use Web of Science	0,663332403	<2.22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Academia edu	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on Academia edu	0,65621376	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	0,650324345	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	0,649752259	<2.22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it tackles a practical problem	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it presents research results that are practically applicable and useful	0,647994876	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	0,631458461	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use my field s vertical search engines. e.g. IEEE for engineering. Pubmed for medicine. etc.	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - The advanced search of your field s vertical search engines. e.g. IEEE for engineering. Pubmed for medicine. etc.	0,618838608	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use ResearchGate alerting services that let me know when a researcher in my network has uploaded a new publication	0,616800189	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Web of Science s advanced search	0,616484642	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	0,608623207	<2.22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Academia edu	0,606454253	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q10. How do you find readings related to your research topics - Other sources. I use Scopus	0,597739697	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on	0,590722799	<2.22E-16

	Academia edu		
Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	0,585577011	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	0,585129082	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	0,581063986	<2.22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Academia edu	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	0,579818845	<2.22E-16
Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on Academia edu	0,569737494	<2.22E-16
Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use ResearchGate alerting services that let me know when a researcher in my network has uploaded a new publication	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on ResearchGate	0,554362416	<2.22E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on the impact factor of their publication outlet	Q16. How do you choose the citations to be included in your own publications - I make an effort to include highly cited papers among my citations	0,553334475	<2.22E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on the impact factor of their publication outlet	Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	0,5516029	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on ResearchGate	0,545843422	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	0,538386941	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	0,534615874	<2.22E-16
Q3. How do you choose your research topics - I choose my research topics as a consequence of suggestions from peer senior members of my research group	Q5. Why do you move away from a research topic - Because peer senior members in my research group advised me to do so	0,534191489	<2.22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use ResearchGate alerting services that let me know when a researcher in my network has uploaded	0,530852139	<2.22E-16

	a new publication		
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	0,529586613	<2.22E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on my target publication outlet	Q16. How do you choose the citations to be included in your own publications - I make an effort to include highly cited papers among my citations	0,527987123	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	0,526813209	<2.22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on ResearchGate	0,52438736	<2.22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Web of Science s advanced search	0,522076786	<2.22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors has a high h index	Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	0,521663964	<2.22E-16
Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on ResearchGate	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on Academia edu	0,521080136	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Academia edu	0,517787576	<2.22E-16
Q12. What type of readings do you look for - Slides	Q12. What type of readings do you look for - Videos	0,514656246	<2.22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors is a well known established researcher in your field	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when I personally know at least one of the authors	0,5115906	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Google Scholar	Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	0,511334419	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	0,509932697	<2.22E-16
Q2. How would you define a. hot research topic - A research topic is hot if many researchers are working on it	Q2. How would you define a. hot research topic - A research topic is hot if it is the focus of many research publications	0,501561821	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	0,498463094	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	0,491602153	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Scopus	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Scopus alerting services that let me know when there is a new publication that matches one of my	0,488034487	<2.22E-16

	searches		
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	0,486398727	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Scopus	Q10. How do you find readings related to your research topics - Other sources. I use Web of Science	0,486184657	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	0,485245466	<2.22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it is published in a journal that has a high impact factor	Q16. How do you choose the citations to be included in your own publications - I choose citations based on the impact factor of their publication outlet	0,484921485	<2.22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Scopus alerting services that let me know when there is a new publication that matches one of my searches	0,479754806	<2.22E-16
Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use ResearchGate alerting services that let me know when a researcher in my network has uploaded a new publication	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	0,472790778	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	0,467273712	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on Academia edu	0,466317743	<2.22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors has a high h index	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors is a well known established researcher in your field	0,465047926	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Google Scholar	Q10. How do you find readings related to your research topics - Other sources. I use Google s general purpose search engine	0,462565303	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	0,460514575	<2.22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Google Scholar	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Google Scholar s advanced search	0,457804888	<2.22E-16
Q2. How would you define a. hot research topic - A research topic is hot if many companies are investing in research on it	Q2. How would you define a. hot research topic - A research topic is hot if it is included in many calls for funded research	0,457507581	<2.22E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	Q16. How do you choose the citations to be included in your own publications - I make an effort to include highly cited papers among my citations	0,453808963	<2.22E-16

Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when I personally know at least one of the authors	Q16. How do you choose the citations to be included in your own publications - I choose citations based on my personal knowledge of at least one of the authors	0,449844271	<2.22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it is published in a journal that has a high impact factor	Q16. How do you choose the citations to be included in your own publications - I make an effort to include highly cited papers among my citations	0,443823695	<2.22E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on my personal knowledge of at least one of the authors	Q16. How do you choose the citations to be included in your own publications - I cite authors who are most likely to cite me back	0,438462585	<2.22E-16
Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Scopus alerting services that let me know when there is a new publication that matches one of my searches	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	0,437230229	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Google s general purpose search engine	Q10. How do you find readings related to your research topics - Other sources. I use Google s general purpose search engine	0,437218875	<2.22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors has a high h index	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors belongs to my research circle in either ResearchGate or Academia edu	0,436006337	<2.22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Google Scholar	Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	0,433523118	2,22045E-16
Q5. Why do you move away from a research topic - Because peer senior members in my research group advised me to do so	Q5. Why do you move away from a research topic - Because I plan to move to a different research group	0,432515264	2,22045E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on the impact factor of their publication outlet	Q16. How do you choose the citations to be included in your own publications - I choose citations based on my target publication outlet	0,431774765	4,44089E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on my personal knowledge of at least one of the authors	Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	0,430216938	4,44089E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Google s general purpose search engine	Q6. Do you use online archives and search engines to choose a research topic - I use my field s vertical search engines. e.g. IEEE for engineering. Pubmed for medicine. etc.	0,42729336	6,66134E-16
Q13. What are the characteristics of your readings - They discuss the results of theoretical research	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it tackles a theoretical problem	0,424162358	1,11022E-15
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Scopus alerting services that let me know when there is a new publication that matches one of my searches	0,424154729	1,11022E-15
Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	Q16. How do you choose the citations to be included in your own publications - I cite authors who are most likely to cite me back	0,422713935	1,33227E-15

Q3. How do you choose your research topics - I choose my research topics according to cooperation opportunities with researchers outside of my research field	Q3. How do you choose your research topics - I choose my research topics according to cooperation opportunities with other researchers in my research field	0,422528118	1,55431E-15
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	0,422496289	1,55431E-15
Q3. How do you choose your research topics - My research topics are a consequence of peer reviewed funded projects	Q5. Why do you move away from a research topic - Because of lack of funds	0,42054674	2,22045E-15
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors belongs to my research circle in either ResearchGate or Academia edu	Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	0,415829659	4,66294E-15
Q3. How do you choose your research topics - I do research on topics that are most likely to be cited	Q5. Why do you move away from a research topic - Because it is no longer a hot topic in my field	0,413397044	6,88338E-15
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q6. Do you use online archives and search engines to choose a research topic - I use my field s vertical search engines. e g. IEEE for engineering. Pubmed for medicine. etc.	0,412039876	8,43769E-15
Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	Q6. Do you use online archives and search engines to choose a research topic - I use my field s vertical search engines. e g. IEEE for engineering. Pubmed for medicine. etc.	0,409817338	1,24345E-14
Q6. Do you use online archives and search engines to choose a research topic - I use Google s general purpose search engine	Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	0,408369333	1,55431E-14
Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	Q6. Do you use online archives and search engines to choose a research topic - I use my field s vertical search engines. e g. IEEE for engineering. Pubmed for medicine. etc.	0,406590849	2,08722E-14
Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Academia edu	0,404710978	2,79776E-14
Q6. Do you use online archives and search engines to choose a research topic - I use Google Scholar	Q6. Do you use online archives and search engines to choose a research topic - I use my field s vertical search engines. e g. IEEE for engineering. Pubmed for medicine. etc.	0,403080761	3,64153E-14

Tab. 10 Coefficienti di correlazione di Pearson e relativi p-value

Question 1 (row)	Question 2 (column)	Pearson rho	p-value
Q10. How do you find readings related to your research topics - Other sources. I use my field s vertical search engines. e g. IEEE for engineering. Pubmed for medicine. etc.	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - The advanced search of your field s vertical search engines. e g. IEEE for engineering. Pubmed for medicine. etc.	0,763210952	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	0,757063687	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources.	Q11. How do you rate the effectiveness of the following tools to search for	0,747922838	<2,22E-16

I use Academia edu	readings related to your research topics - Academia edu		
Q10. How do you find readings related to your research topics - Other sources. I use Scopus	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	0,737024367	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Web of Science	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Web of Science s advanced search	0,734518945	<2,22E-16
Q5. Why do you move away from a research topic - Because I plan to move to a different research institution	Q5. Why do you move away from a research topic - Because I plan to move to a different research group	0,732993603	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Google Scholar	Q6. Do you use online archives and search engines to choose a research topic - I use Google s general purpose search engine	0,729584754	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	0,727479339	<2,22E-16
Q17. How do you work to disseminate your own research work - I circulate my own papers before publication to get early feedback	Q17. How do you work to disseminate your own research work - I circulate my own papers before publication to engage other researchers	0,712971687	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use my field s vertical search engines. e.g. IEEE for engineering. Pubmed for medicine. etc.	Q10. How do you find readings related to your research topics - Other sources. I use my field s vertical search engines. e.g. IEEE for engineering. Pubmed for medicine. etc.	0,71184957	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it tackles a practical problem	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it presents research results that are practically applicable and useful	0,675996661	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q10. How do you find readings related to your research topics - Other sources. I use Scopus	0,647215545	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	0,64419198	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	0,641770959	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	0,636540711	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	Q10. How do you find readings related to your research topics - Other sources. I use Web of Science	0,636132002	<2,22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Academia edu	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on Academia edu	0,635257661	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use my field s vertical search engines. e.g. IEEE for engineering. Pubmed for medicine. etc.	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - The advanced search of your field s vertical search engines. e.g. IEEE for engineering. Pubmed for medicine. etc.	0,618445635	<2,22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics	0,617992342	<2,22E-16

- ResearchGate net	- Academia edu		
Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on Academia edu	0,610599995	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Web of Science s advanced search	0,609110713	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	0,603158236	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use ResearchGate alerting services that let me know when a researcher in my network has uploaded a new publication	0,576746941	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	0,574749589	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	0,573210597	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	0,566448092	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on Academia edu	0,561497748	<2,22E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on the impact factor of their publication outlet	Q16. How do you choose the citations to be included in your own publications - I make an effort to include highly cited papers among my citations	0,559841633	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Academia edu	0,551254332	<2,22E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on the impact factor of their publication outlet	Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	0,550388157	<2,22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Academia edu	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	0,549165487	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q10. How do you find readings related to your research topics - Other sources. I use Scopus	0,548709154	<2,22E-16
Q3. How do you choose your research topics - I choose my research topics as a consequence of suggestions from peer senior members of my research group	Q5. Why do you move away from a research topic - Because peer senior members in my research group advised me to do so	0,546418905	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors has a high h index	Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	0,545342922	<2,22E-16

Q16. How do you choose the citations to be included in your own publications - I choose citations based on my target publication outlet	Q16. How do you choose the citations to be included in your own publications - I make an effort to include highly cited papers among my citations	0,544140458	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	0,539598286	<2,22E-16
Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on ResearchGate	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on Academia edu	0,538879752	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Google Scholar	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Google Scholar s advanced search	0,531597555	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors is a well known established researcher in your field	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when I personally know at least one of the authors	0,52801609	<2,22E-16
Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use ResearchGate alerting services that let me know when a researcher in my network has uploaded a new publication	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on ResearchGate	0,525305271	<2,22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Web of Science s advanced search	0,522851944	<2,22E-16
Q12. What type of readings do you look for - Slides	Q12. What type of readings do you look for - Videos	0,512013674	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on ResearchGate	0,511825264	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	0,508973598	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Google Scholar	Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	0,505849123	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	0,498063654	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Scopus	Q10. How do you find readings related to your research topics - Other sources. I use Web of Science	0,495984912	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors has a high h index	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors is a well known established researcher in your field	0,495450169	<2,22E-16
Q2. How would you define a. hot research topic - A research topic is hot if many researchers are working on it	Q2. How would you define a. hot research topic - A research topic is hot if it is the focus of many research publications	0,492902517	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it is published in a journal that has a high impact factor	Q16. How do you choose the citations to be included in your own publications - I choose citations based on the impact factor of their publication outlet	0,491467088	<2,22E-16

Q10. How do you find readings related to your research topics - Other sources. I use Scopus	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Scopus alerting services that let me know when there is a new publication that matches one of my searches	0,488222271	<2,22E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	Q16. How do you choose the citations to be included in your own publications - I cite authors who are most likely to cite me back	0,484680891	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	0,483400255	<2,22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on ResearchGate	0,47956261	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	0,477227062	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	0,476717055	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Google Scholar	Q10. How do you find readings related to your research topics - Other sources. I use Google s general purpose search engine	0,468595356	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on Academia edu	0,468063176	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	0,466196924	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	0,464294076	<2,22E-16
Q3. How do you choose your research topics - I choose my research topics according to cooperation opportunities with researchers outside of my research field	Q3. How do you choose your research topics - I choose my research topics according to cooperation opportunities with other researchers in my research field	0,464215428	<2,22E-16
Q2. How would you define a. hot research topic - A research topic is hot is many companies are investing in research on it	Q2. How would you define a. hot research topic - A research topic is hot if it is included in many calls for funded research	0,464179128	<2,22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use ResearchGate alerting services that let me know when a researcher in my network has uploaded a new publication	0,463936538	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	0,461153179	<2,22E-16
Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use ResearchGate alerting services that let me know when a researcher in my network has uploaded a new publication	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	0,458293736	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when I personally know at least one of the	Q16. How do you choose the citations to be included in your own publications - I choose citations based on my personal knowledge of at least one of the authors	0,457839191	<2,22E-16

authors			
Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	Q16. How do you choose the citations to be included in your own publications - I make an effort to include highly cited papers among my citations	0,455259472	<2,22E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on my personal knowledge of at least one of the authors	Q16. How do you choose the citations to be included in your own publications - I cite authors who are most likely to cite me back	0,45069313	<2,22E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on my personal knowledge of at least one of the authors	Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	0,450302392	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it is published in a journal that has a high impact factor	Q16. How do you choose the citations to be included in your own publications - I make an effort to include highly cited papers among my citations	0,446705669	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors has a high h index	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors belongs to my research circle in either ResearchGate or Academia edu	0,445405096	<2,22E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on the impact factor of their publication outlet	Q16. How do you choose the citations to be included in your own publications - I choose citations based on my target publication outlet	0,444499493	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	0,444110125	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Google s general purpose search engine	Q10. How do you find readings related to your research topics - Other sources. I use Google s general purpose search engine	0,44225055	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	0,441873848	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors belongs to my research circle in either ResearchGate or Academia edu	Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	0,433797568	2,22045E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Scopus alerting services that let me know when there is a new publication that matches one of my searches	0,433609217	2,22045E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Google Scholar	Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	0,430186212	4,44089E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Google s general purpose search engine	Q6. Do you use online archives and search engines to choose a research topic - I use my field s vertical search engines. e.g. IEEE for engineering. Pubmed for medicine. etc.	0,41604653	4,44089E-15
Q13. What are the characteristics of your readings - They discuss the results of theoretical research	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it tackles a theoretical problem	0,414617002	5,77316E-15
Q3. How do you choose your research topics - I do research on what I think is a. hot topic. in my field	Q3. How do you choose your research topics - I do research on topics that are most likely to be cited	0,409773886	1,24345E-14

Q3. How do you choose your research topics - My research topics are a consequence of peer reviewed funded projects	Q5. Why do you move away from a research topic - Because of lack of funds	0,409020364	1,42109E-14
Q6. Do you use online archives and search engines to choose a research topic - I use Google Scholar	Q6. Do you use online archives and search engines to choose a research topic - I use my field s vertical search engines. e g. IEEE for engineering. Pubmed for medicine. etc.	0,408360004	1,55431E-14
Q3. How do you choose your research topics - I do research on topics that are most likely to be cited	Q5. Why do you move away from a research topic - Because it is no longer a. hot topic. in my field	0,404882371	2,73115E-14
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	0,402555883	3,95239E-14
Q6. Do you use online archives and search engines to choose a research topic - I use Google s general purpose search engine	Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	0,401464939	4,68514E-14

Tab. 11 Coefficienti di correlazione di Kendall e relativi p-value

Question 1 (row)	Question 2 (column)	Kendall tau	p-value
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	0,761432161	<2,22E-16
Q5. Why do you move away from a research topic - Because I plan to move to a different research institution	Q5. Why do you move away from a research topic - Because I plan to move to a different research group	0,715322664	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Scopus	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	0,704685759	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Web of Science	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Web of Science s advanced search	0,690141703	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Academia edu	0,686027112	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use my field s vertical search engines. e g. IEEE for engineering. Pubmed for medicine. etc.	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - The advanced search of your field s vertical search engines. e g. IEEE for engineering. Pubmed for medicine. etc.	0,682726193	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q10. How do you find readings related to your research topics - Other sources. I use Scopus	0,676535009	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	0,675032922	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Google Scholar	Q6. Do you use online archives and search engines to choose a research topic - I use Google s general purpose search engine	0,666452941	<2,22E-16
Q17. How do you work to disseminate your own research work - I circulate my own papers before publication to get early feedback	Q17. How do you work to disseminate your own research work - I circulate my own papers before publication to engage other researchers	0,649604826	<2,22E-16
Q6. Do you use online archives and search engines to choose a research	Q10. How do you find readings related to your research topics - Other sources.	0,633235862	<2,22E-16

topic - I use my field s vertical search engines. e g. IEEE for engineering. Pubmed for medicine. etc.	I use my field s vertical search engines. e g. IEEE for engineering. Pubmed for medicine. etc.		
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Academia edu	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on Academia edu	0,629978499	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	0,626077538	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it tackles a practical problem	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it presents research results that are practically applicable and useful	0,616155297	7,54952E-15
Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	0,608218093	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	Q10. How do you find readings related to your research topics - Other sources. I use Web of Science	0,60331346	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	0,602343346	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	0,591990812	<2,22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Academia edu	0,581286906	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on Academia edu	0,566257362	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q10. How do you find readings related to your research topics - Other sources. I use Scopus	0,56043182	<2,22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Academia edu	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	0,557775925	<2,22E-16
Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on Academia edu	0,557772081	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use ResearchGate alerting services that let me know when a researcher in my network has uploaded a new publication	0,556631122	2,22045E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	0,555453271	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics	0,552348376	<2,22E-16

	- Web of Science s advanced search		
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	0,55111416	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use my field s vertical search engines. e g. IEEE for engineering. Pubmed for medicine. etc.	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - The advanced search of your field s vertical search engines. e g. IEEE for engineering. Pubmed for medicine. etc.	0,55013486	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	0,529205694	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	0,511572607	<2,22E-16
Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use ResearchGate alerting services that let me know when a researcher in my network has uploaded a new publication	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on ResearchGate	0,511126099	1,06581E-14
Q16. How do you choose the citations to be included in your own publications - I choose citations based on the impact factor of their publication outlet	Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	0,506615878	1,75193E-13
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	0,50343262	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	0,501032888	<2,22E-16
Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on ResearchGate	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on Academia edu	0,499983674	2,22045E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Academia edu	0,496123808	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use ResearchGate net	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on ResearchGate	0,494785357	<2,22E-16
Q3. How do you choose your research topics - I choose my research topics as a consequence of suggestions from peer senior members of my research group	Q5. Why do you move away from a research topic - Because peer senior members in my research group advised me to do so	0,49383481	<2,22E-16
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Web of Science s advanced search	0,490893177	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	0,490684648	<2,22E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on the impact factor of their publication outlet	Q16. How do you choose the citations to be included in your own publications - I make an effort to include highly cited papers among my citations	0,490597148	<2,22E-16
Q12. What type of readings do you look for - Slides	Q12. What type of readings do you look for - Videos	0,48505002	1,48881E-12

Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors has a high h index	Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	0,484288664	5,66214E-14
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use ResearchGate alerting services that let me know when a researcher in my network has uploaded a new publication	0,483812524	2,05413E-12
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on ResearchGate	0,474467701	4,44089E-16
Q16. How do you choose the citations to be included in your own publications - I choose citations based on my target publication outlet	Q16. How do you choose the citations to be included in your own publications - I make an effort to include highly cited papers among my citations	0,471796331	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	0,46863684	<2,22E-16
Q2. How would you define a. hot research topic - A research topic is hot if many researchers are working on it	Q2. How would you define a. hot research topic - A research topic is hot if it is the focus of many research publications	0,468456745	4,50751E-14
Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	0,465560213	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	0,462434214	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors is a well known established researcher in your field	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when I personally know at least one of the authors	0,461350193	6,63502E-09
Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	0,457449736	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Scopus	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Scopus alerting services that let me know when there is a new publication that matches one of my searches	0,45565361	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Google Scholar	Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	0,455434364	1,11022E-15
Q6. Do you use online archives and search engines to choose a research topic - I use Web of Science	Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	0,45378158	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Scopus	Q10. How do you find readings related to your research topics - Other sources. I use Web of Science	0,45124271	<2,22E-16
Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use ResearchGate alerting services that let me know when a researcher in my network has uploaded a new publication	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	0,450157548	1,48881E-12
Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Scopus advanced search	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Scopus alerting services that let me	0,448887568	<2,22E-16

	know when there is a new publication that matches one of my searches		
Q6. Do you use online archives and search engines to choose a research topic - I use Academia edu	Q17. How do you work to disseminate your own research work - I reach out to other researchers in my field on Academia edu	0,447929237	<2,22E-16
Q6. Do you use online archives and search engines to choose a research topic - I use Scopus	Q6. Do you use online archives and search engines to choose a research topic - I use ResearchGate net	0,435121724	<2,22E-16
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when it is published in a journal that has a high impact factor	Q16. How do you choose the citations to be included in your own publications - I choose citations based on the impact factor of their publication outlet	0,429620297	<2,22E-16
Q10. How do you find readings related to your research topics - Other sources. I use Academia edu	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - ResearchGate net	0,426529225	<2,22E-16
Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Scopus alerting services that let me know when there is a new publication that matches one of my searches	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Academia edu alerting services that let me know when a researcher in my network has uploaded a new publication	0,425826512	2,13163E-14
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors has a high h index	Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when at least one of the authors is a well known established researcher in your field	0,423353731	4,03011E-13
Q2. How would you define a. hot research topic - A research topic is hot if many companies are investing in research on it	Q2. How would you define a. hot research topic - A research topic is hot if it is included in many calls for funded research	0,41810056	1,89989E-08
Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	Q16. How do you choose the citations to be included in your own publications - I make an effort to include highly cited papers among my citations	0,414782481	1,94711E-12
Q16. How do you choose the citations to be included in your own publications - I choose citations where at least one of the authors has a high h index	Q16. How do you choose the citations to be included in your own publications - I cite authors who are most likely to cite me back	0,407590645	2,84217E-14
Q10. How do you find readings related to your research topics - Other sources. I use Google Scholar	Q11. How do you rate the effectiveness of the following tools to search for readings related to your research topics - Google Scholar s advanced search	0,406545986	2,20988E-07
Q6. Do you use online archives and search engines to choose a research topic - I use the analytic functionalities of Scopus. Analyze search results. button	Q14. Do you use push services that alert you when new research is available that might be of interest for you - I use Scopus alerting services that let me know when there is a new publication that matches one of my searches	0,405941082	2,58904E-13
Q16. How do you choose the citations to be included in your own publications - I choose citations based on my personal knowledge of at least one of the authors	Q16. How do you choose the citations to be included in your own publications - I cite authors who are most likely to cite me back	0,405628342	1,30554E-10
Q15. What are the drivers that make you think that a publication is worth reading - A publication is worth reading when I personally know at least one of the authors	Q16. How do you choose the citations to be included in your own publications - I choose citations based on my personal knowledge of at least one of the authors	0,402458817	6,39575E-10
Q10. How do you find readings related to your research topics - Other sources. I use Google Scholar	Q10. How do you find readings related to your research topics - Other sources. I use Google s general purpose search engine	0,401992985	6,93665E-09
Q16. How do you choose the citations to be included in your own publications - I	Q16. How do you choose the citations to be included in your own publications - I	0,400785649	4,67916E-07

choose citations based on my personal knowledge of at least one of the authors	choose citations where at least one of the authors has a high h index		
--------------------------------------------------------------------------------	-----------------------------------------------------------------------	--	--

Tab. 12 T-Test rilevanti per l'analisi sulle fonti di lettura

		How do you find readings related to your research topics? - Online Sources - I use Google Scholar	How do you find readings related to your research topics? - Online Sources - I use Google's general purpose search engine	How do you find readings related to your research topics? - Online Sources - I use Scopus	How do you find readings related to your research topics? - Online Sources - I use Web of Science	How do you find readings related to your research topics? - Online Sources - I use my field's vertical search engines (e.g. IEEE for engineering, Pubmed for medicine, etc.)	How do you find readings related to your research topics? - Online Sources - I select interesting readings among the references of papers that I have read
		Frequently Rarely	Frequently Rarely	Frequently Rarely	Frequently Rarely	Frequently Rarely	Frequently Rarely
	Cluster %	64,42% 35,58%	53,68% 46,32%	6,75% 93,25%	11,96% 88,04%	23,62% 76,38%	61,35% 38,65%
How would you define a "hot research topic"? - A research topic is hot if many researchers are working on it	Vertical Means	1,7 1,92	1,67 1,89	1,59 1,79	1,77 1,78	1,74 1,79	1,73 1,85
	T-Test	A-B: T Value: 2,52 DF: 210,59 P-Value: 0,01	A-B: T Value: 2,64 DF: 311,65 P-Value: 0,01	-	-	-	-
How would you define a "hot research topic"? - A research topic is hot if papers focusing on it are more likely to be cited	Vertical Means	2,07 2,1	2,11 2,05	2 2,09	1,9 2,1	2,14 2,06	2,17 1,94
	T-Test	-	-	-	-	-	A-B: T Value: 2,29 DF: 262,54 P-Value: 0,02
How do you choose your research topics? - I do research on what is an "hot topic" in my fields	Vertical Means	2,53 2,58	2,46 2,65	2,32 2,56	2,31 2,58	2,43 2,58	2,54 2,56
	T-Test	-	A-B: T Value: 2,53 DF: 321,02 P-Value: 0,01	-	A-B: T Value: 2,10 DF: 46,35 P-Value: 0,04	-	-
How do you choose your research topics? - I do research on topics that are most likely to be cited	Vertical Means	2,77 2,94	2,77 2,9	2,64 2,85	2,59 2,86	2,73 2,86	2,88 2,76
	T-Test	A-B: T Value: 2,07 DF: 228,13 P-Value: 0,04	-	-	A-B: T Value: 2,26 DF: 48,08 P-Value: 0,03	-	-
How do you choose your research topics? - I choose my research topics as a consequence of suggestions from peer/senior members of my research group	Vertical Means	2,88 2,92	2,89 2,91	2,73 2,91	2,79 2,91	2,71 2,95	2,86 2,95
	T-Test	-	-	-	-	A-B: T Value: 2,52 DF: 137,83 P-Value: 0,01	-
How do you choose your research topics? - I choose my research topics based on the directions provided by the most impactful authors in	Vertical Means	2,9 3,13	2,93 3,05	2,64 3,01	2,85 3	2,82 3,04	2,96 3,02
	T-Test	A-B: T Value: 2,49 DF: 224,14	-	A-B: T Value: 2,15 DF: 23,94 P-	-	A-B: T Value: 2,24 DF: 131,78	-

my field		P-Value: 0,01		Value: 0,04		P-Value: 0,03	
Why do you move away from a research topic? - Because peer/senior members in my research group advised me to do so	Vertical Means	3,35 3,53	3,35 3,49	3 3,44	3,13 3,45	3,12 3,51	3,42 3,41
	T-Test	A-B: T Value: 2,09 DF: 268,56 P-Value: 0,04	-	-	-	A-B: T Value: 3,43 DF: 111,06 P- Value<0,01	-
Why do you move away from a research topic? - Because of a personal loss of intellectual interest	Vertical Means	1,57 1,69	1,58 1,65	1,64 1,61	1,56 1,62	1,62 1,61	1,52 1,75
	T-Test	-	-	-	-	-	A-B: T Value: 2,41 DF: 225,38 P-Value: 0,02
What are the drivers that make you think that a publication is worth reading? - A publication is worth reading when it has a high number of citations	Vertical Means	2,06 2,28	2,13 2,15	1,59 2,18	1,74 2,19	1,97 2,19	2,13 2,14
	T-Test	A-B: T Value: 2,23 DF: 217,39 P-Value: 0,03	-	A-B: T Value: 3,32 DF: 24,62 P- Value<0,01	A-B: T Value: 3,31 DF: 51,14 P- Value<0,01	-	-
What are the drivers that make you think that a publication is worth reading? - A publication is worth reading when at least one of the authors has a high h-index	Vertical Means	3,08 3,15	3,02 3,2	2,55 3,14	2,69 3,16	3,01 3,13	3,12 3,07
	T-Test	-	-	A-B: T Value: 2,69 DF: 23,39 P- Value: 0,01	A-B: T Value: 2,63 DF: 45,28 P- Value: 0,01	-	-
What are the drivers that make you think that a publication is worth reading? - A publication is worth reading when it is published in a journal that has a high impact factor	Vertical Means	2,12 2,29	2,18 2,19	1,95 2,2	1,92 2,22	2,03 2,23	2,17 2,21
	T-Test	-	-	-	A-B: T Value: 2,12 DF: 52,07 P- Value: 0,04	-	-
How do you choose the citations to be included in your own publications? - I choose citations based on the impact factor of their publication outlet	Vertical Means	2,74 2,95	2,75 2,89	2,59 2,83	2,44 2,87	2,69 2,86	2,85 2,75
	T-Test	-	-	-	A-B: T Value: 2,45 DF: 46,16 P- Value: 0,02	-	-
How do you choose the citations to be included in your own publications? - I choose citations where at least one of the authors has a high h-index	Vertical Means	3,55 3,69	3,51 3,7	3,18 3,63	3,38 3,63	3,44 3,65	3,58 3,62
	T-Test	A-B: T Value: 2,01 DF: 264,76 P-Value: 0,05	A-B: T Value: 2,80 DF: 321,89 P-Value: 0,01	A-B: T Value: 2,27 DF: 22,37 P- Value: 0,03	-	A-B: T Value: 2,19 DF: 106,71 P-Value: 0,03	-
How do you choose the citations to be included in your own publications? - I choose citations based on my target publication outlet	Vertical Means	2,32 2,59	2,34 2,51	2,14 2,44	2,1 2,46	2,3 2,45	2,44 2,38
	T-Test	A-B: T Value: 2,62 DF: 230,76 P-Value: 0,01	-	-	A-B: T Value: 2,54 DF: 50,15 P- Value: 0,01	-	-
How do you choose the citations to be included in your own publications? - I make an effort to include highly cited papers among my citations	Vertical Means	2,72 3,04	2,81 2,87	2,45 2,87	2,44 2,89	2,58 2,92	2,89 2,75
	T-Test	A-B: T Value: 3,02 DF: 240,60 P- Value<0,01	-	-	A-B: T Value: 2,49 DF: 45,11 P- Value: 0,02	A-B: T Value: 2,70 DF: 122,12 P-Value: 0,01	-
How do you choose the citations to be included in your own publications? - I cite authors who are most likely to cite me back	Vertical Means	3,58 3,82	3,56 3,79	3,41 3,68	3,44 3,7	3,57 3,69	3,63 3,72
	T-Test	A-B: T Value: 3,66 DF: 312,76 P- Value<0,01	A-B: T Value: 3,33 DF: 316,98 P- Value<0,01	-	-	-	-

Tab. 13 Frequenze reali-attese e relativo test χ^2 d'indipendenza per l'analisi delle fonti di lettura

	Cluster %	How do you find readings related to your research topics? - Online Sources - I use Google Scholar		How do you find readings related to your research topics? - Online Sources - I use Google's general purpose search engine		How do you find readings related to your research topics? - Online Sources - I use Scopus		How do you find readings related to your research topics? - Online Sources - I use Web of Science		How do you find readings related to your research topics? - Online Sources - I use my field's vertical search engines (e.g. IEEE for engineering, Pubmed for medicine, etc.)		How do you find readings related to your research topics? - Online Sources - I select interesting readings among the references of papers that I have read	
		Every day, Every week A	Every month, A few times in a year, Never B	Every day, Every week A	Every month, A few times in a year, Never B	Every day, Every week A	Every month, A few times in a year, Never B	Every day, Every week A	Every month, A few times in a year, Never B	Every day, Every week A	Every month, A few times in a year, Never B	Every day, Every week A	Every month, A few times in a year, Never B
	64,42 %	35,58 %	53,68 %	46,32 %	6,75 %	93,25 %	11,96 %	88,04 %	23,62 %	76,38 %	61,35 %	38,65 %	
How would you define a "hot research topic"? - A research topic is hot if many researchers are working on it	Strongly Agree	7,83	-7,83	12,36	-12,36	1,5	-1,5	-0,07	0,07	-0,76	0,76	2,7	-2,7
	Weakly Agree	-0,78	0,78	-8,82	8,82	0,34	-0,34	0,1	-0,1	3,68	-3,68	0,07	-0,07
	Weakly Disagree	-4,97	4,97	-1,64	1,64	-1,09	1,09	0,29	-0,29	-2,32	2,32	0,98	-0,98
	Strongly Disagree	-2,09	2,09	-1,9	1,9	-0,74	0,74	-0,32	0,32	-0,6	0,6	-3,75	3,75
	Chi Square	7,34*		8,53		1,70*		0,11*		1,65*		5,76*	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,06		0,04		0,64		0,99		0,65		0,12	
How do you choose your research topics? - I do research on what is an "hot topic" in my fields	Always	-0,60	0,60	2,34	-2,34	1,79	-1,79	4,85	-4,85	3,75	-3,75	4,96	-4,96
	Frequently	2,61	-2,61	9,68	-9,68	0,16	-0,16	-2,67	2,67	0,06	-0,06	-7,37	7,37
	Seldom	0,22	-0,22	-10,82	10,82	-0,66	0,66	0,10	-0,10	-2,32	2,32	1,07	-1,07
	Never	-2,24	2,24	-1,20	1,20	-1,28	1,28	-2,27	2,27	-1,49	1,49	1,34	-1,34
	Chi Square	1,47		7,38		4,24*		15,49*		5,16*		7,94	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,69		0,06		0,24		0		0,16		0,05	
How do you choose your research topics? - I do research on topics that are most likely to be cited	Always	-0,51	0,51	-0,76	0,76	0,53	-0,53	2,16	-2,16	0,35	-0,35	-1,29	1,29
	Frequently	8,02	-8,02	6,69	-6,69	1,93	-1,93	1,23	-1,23	5,74	-5,74	-4,21	4,21
	Seldom	-1,95	1,95	-0,63	0,63	-1,15	1,15	0,47	-0,47	-4,52	4,52	3,57	-3,57
	Never	-5,56	5,56	-5,3	5,3	-1,31	1,31	-3,86	3,86	-1,57	1,57	1,94	-1,94
	Chi Square	6,13*		4,65*		1,96*		9,41*		3,03*		2,46*	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,11		0,2		0,58		0,02		0,39		0,48	
How do you choose your research topics? - I choose my research topics as a consequence of suggestions from peer/senior members of my research group	Always	1,20	-1,20	0,17	-0,17	-0,61	0,61	-0,08	0,08	0,87	-0,87	0,48	-0,48
	Frequently	-0,04	0,04	0,30	-0,30	2,13	-2,13	1,59	-1,59	3,45	-3,45	2,63	-2,63
	Seldom	-0,42	0,42	0,65	-0,65	1,27	-1,27	0,98	-0,98	4,44	-4,44	0,45	-0,45
	Never	-0,74	0,74	-1,11	1,11	-2,79	2,79	-2,49	2,49	-8,77	8,77	-3,56	3,56
	Chi Square	0,74*		0,10*		3,38*		1,17*		7,92*		1,20*	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,86		0,99		0,34		0,76		0,05		0,75	

How do you choose your research topics? - I choose my research topics based on the directions provided by the most impactful authors in my field	Always	-0,15	0,15	-1,29	1,29	0,46	-0,46	0,04	-0,04	1,11	-1,11	-0,91	0,91
	Frequently	5,69	-5,69	3,74	-3,74	3,94	-3,94	2,03	-2,03	2,29	-2,29	1,99	-1,99
	Seldom	5,87	-5,87	5,72	-5,72	-1,6	1,6	1,22	-1,22	4,92	-4,92	3,68	-3,68
	Never	-11,4	11,4	-8,17	8,17	-2,8	2,8	-3,29	3,29	-8,31	8,31	-4,76	4,76
	Chi Square	9,44*		5,55*		5,42*		1,81*		6,55*		2,13*	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,02		0,14		0,14		0,61		0,09		0,55	
Why do you move away from a research topic? - Because peer/senior members in my research group advised me to do so	Strongly Agree	1,91	-1,91	0,1	-0,1	2,26	-2,26	0,68	-0,68	2,4	-2,4	-0,75	0,75
	Weakly Agree	2,1	-2,1	-0,25	0,25	2,71	-2,71	5,93	-5,93	3,97	-3,97	0,14	-0,14
	Weakly Disagree	4,02	-4,02	11,69	-11,69	-3,07	3,07	-2,77	2,77	7,74	-7,74	1,79	-1,79
	Strongly Disagree	-8,04	8,04	-11,53	11,53	-1,89	1,89	-3,85	3,85	-14,11	14,11	-1,18	1,18
	Chi Square	4,28*		8,91		12,75*		11,78*		14,95*		0,40*	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,23		0,03		0,01		0,01		0		0,94	
What are the drivers that make you think that a publication is worth reading? - A publication is worth reading when it has a high number of citations	Strongly Agree	4,33	-4,33	3,28	-3,28	7,01	-7,01	8,15	-8,15	5,52	-5,52	-0,4	0,4
	Weakly Agree	3,93	-3,93	-3,89	3,89	-2,8	2,8	-3,14	3,14	1,21	-1,21	2,84	-2,84
	Weakly Disagree	-3,87	3,87	-0,89	0,89	-3,39	3,39	-2,78	2,78	-6,35	6,35	-3,88	3,88
	Strongly Disagree	-4,39	4,39	1,51	-1,51	-0,82	0,82	-2,23	2,23	-0,38	0,38	1,44	-1,44
	Chi Square	5,65		1,35		14,52*		11,98*		5,81		1,52	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,13		0,72		0		0,01		0,12		0,68	
What are the drivers that make you think that a publication is worth reading? - A publication is worth reading when at least one of the authors has a high h-index	Strongly Agree	1,69	-1,69	2,41	-2,41	1,92	-1,92	3,09	-3,09	2,22	-2,22	-0,82	0,82
	Weakly Agree	-2,09	2,09	4,42	-4,42	4,28	-4,28	5,63	-5,63	0,47	-0,47	3,06	-3,06
	Weakly Disagree	4,36	-4,36	-1,37	1,37	-2,09	2,09	-4,56	4,56	-0,8	0,8	-7,42	7,42
	Strongly Disagree	-3,96	3,96	-5,47	5,47	-4,11	4,11	-4,15	4,15	-1,89	1,89	5,18	-5,18
	Chi Square	2,35		3,55		10,46*		13,04*		1,91*		3,79	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,5		0,31		0,02		0		0,59		0,29	
How do you choose the citations to be included in your own publications? - I choose citations based on the impact factor of their publication outlet	Always	2,54	-2,54	1,12	-1,12	1,38	-1,38	6,13	-6,13	2,33	-2,33	1,28	-1,28
	Frequently	5,72	-5,72	3,1	-3,1	0,85	-0,85	-1,68	1,68	2,96	-2,96	-6,03	6,03
	Seldom	-3,71	3,71	2,25	-2,25	-0,88	0,88	-0,2	0,2	-3,09	3,09	0,42	-0,42
	Never	-4,55	4,55	-6,46	6,46	-1,34	1,34	-4,25	4,25	-2,2	2,2	4,33	-4,33
	Chi Square	4,07		2,56		1,80*		16,94*		2,52		2,58	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,25		0,46		0,62		0		0,47		0,46	
How do you choose the citations to be included in your own publications? - I choose citations where at least one of	Always	0,71	-0,71	0,93	-0,93	1,87	-1,87	1,76	-1,76	1,53	-1,53	0,77	-0,77
	Frequently	0,47	-0,47	2,73	-2,73	-0,42	0,42	-1,51	1,51	1,04	-1,04	2,12	-2,12
	Seldom	7,53	-7,53	7,44	-7,44	4,4	-4,4	6,07	-6,07	5,4	-5,4	-3,92	3,92
	Never	-8,72	8,72	-11,1	11,1	-5,85	5,85	-6,32	6,32	-7,96	7,96	1,03	-1,03
	Chi Square	5,64*		8,09*		33,94*		21,69*		10,29*		2,96*	
	Degrees	3		3		3		3		3		3	

the authors has a high h-index	of Freedom												
	p-value	0,13		0,04		0		0		0,02		0,4	
How do you choose the citations to be included in your own publications ? - I choose citations based on my target publication outlet	Always	3,3	-3,3	2,92	-2,92	1,1	-1,1	3,86	-3,86	1,84	-1,84	-2,38	2,38
	Frequently	8,88	-8,88	6,24	-6,24	1,35	-1,35	1,89	-1,89	2,22	-2,22	2,27	-2,27
	Seldom	-8,06	8,06	-7,22	7,22	0,18	-0,18	-3,08	3,08	-0,86	0,86	-1,96	1,96
	Never	-4,12	4,12	-1,94	1,94	-2,63	2,63	-2,67	2,67	-3,21	3,21	2,07	-2,07
	Chi Square	8,22		4,35		3,48*		6,14*		2,14		1,33	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,04		0,23		0,32		0,1		0,54		0,72	
How do you choose the citations to be included in your own publications ? - I make an effort to include highly cited papers among my citations	Always	3,25	-3,25	-0,96	0,96	5,25	-5,25	7,89	-7,89	3,86	-3,86	1,05	-1,05
	Frequently	6,09	-6,09	1,08	-1,08	-3,28	3,28	-4,13	4,13	5,03	-5,03	-9,06	9,06
	Seldom	1,92	-1,92	6,27	-6,27	-0,76	0,76	0,24	-0,24	-2,16	2,16	4,45	-4,45
	Never	-11,26	11,26	-6,39	6,39	-1,21	1,21	-4,01	4,01	-6,73	6,73	3,56	-3,56
	Chi Square	9,67		3,35		18,98*		26,13*		7,64		5,2	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,02		0,34		0		0		0,05		0,16	
How do you choose the citations to be included in your own publications ? - I cite authors who are most likely to cite me back	Always	1,42	-1,42	0,85	-0,85	1,73	-1,73	1,52	-1,52	1,06	-1,06	1,55	-1,55
	Frequently	2,4	-2,4	3,34	-3,34	-1,21	1,21	-1,15	1,15	-2,25	2,25	-1,04	1,04
	Seldom	8,71	-8,71	9,25	-9,25	2,88	-2,88	6,7	-6,7	8,59	-8,59	4,58	-4,58
	Never	-12,53	12,53	-13,44	13,44	-3,4	3,4	-7,07	7,07	-7,4	7,4	-5,08	5,08
	Chi Square	11,85*		11,86*		16,11*		15,14*		11,06*		4,67*	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,01		0,01		0		0		0,01		0,2	

Tab. 14 T-Test rilevanti per l'analisi sull'utilizzo dei motori di ricerca e archivi online nella scelta del topic di ricerca

		Do you use on-line archives and search engines to choose a research topic? - I do not use them to choose a research topic	Do you use on-line archives and search engines to choose a research topic? - I use Google Scholar	Do you use on-line archives and search engines to choose a research topic? - I use Google's general purpose search engine	Do you use on-line archives and search engines to choose a research topic? - I use Scopus	Do you use on-line archives and search engines to choose a research topic? - I use Web of Science	Do you use on-line archives and search engines to choose a research topic? - I use my field's vertical search engines (e.g. IEEE for engineering, Pubmed for medicine, etc.)
		Agree Disagree	Agree Disagree	Agree Disagree	Agree Disagree	Agree Disagree	Agree Disagree
	Cluster %	60,74% 39,26%	57,36% 42,64%	45,71% 54,29%	13,80% 86,20%	24,23% 75,77%	33,74% 66,26%
How would you define a "hot research topic"? - A research topic is hot if many researchers are working on it	Vertical Means	1,76 1,8	1,69 1,89	1,66 1,88	1,67 1,79	1,67 1,81	1,7 1,81
	T-Test	-	A-B: T Value: 2,38 DF: 281,82 P-Value:	A-B: T Value: 2,64 DF: 321,18 P-Value:	-	-	-

			0,02	0,01			
How would you define a "hot research topic"? - A research topic is hot if papers focusing on it are more likely to be cited	Vertical Means	2,17 1,95	2,06 2,11	2,12 2,05	2,09 2,08	2,09 2,08	2,14 2,05
	T-Test	A-B: T Value: 2,42 DF: 307,57 P-Value: 0,02	-	-	-	-	-
How do you choose your research topics? - I do research on what is an "hot topic" in my fields	Vertical Means	2,59 2,48	2,49 2,63	2,42 2,66	2,27 2,59	2,38 2,6	2,45 2,59
	T-Test	-	-	A-B: T Value: 3,20 DF: 323,70 P- Value<0,01	A-B: T Value: 3,21 DF: 63,03 P- Value<0,01	A-B: T Value: 2,48 DF: 131,54 P-Value: 0,01	-
How do you choose your research topics? - I do research on topics that are most likely to be cited	Vertical Means	2,91 2,71	2,74 2,95	2,74 2,9	2,62 2,86	2,66 2,89	2,77 2,86
	T-Test	A-B: T Value: 2,56 DF: 282,66 P-Value: 0,01	A-B: T Value: 2,62 DF: 277,06 P-Value: 0,01	A-B: T Value: 2,07 DF: 320,10 P-Value: 0,04	A-B: T Value: 2,30 DF: 61,49 P-Value: 0,02	A-B: T Value: 2,70 DF: 144,12 P-Value: 0,01	-
How do you choose your research topics? - I choose my research topics as a consequence of suggestions from peer/senior members of my research group	Vertical Means	2,92 2,86	2,81 3,01	2,76 3,01	2,76 2,92	2,82 2,92	2,72 2,99
	T-Test	-	A-B: T Value: 2,43 DF: 296,92 P-Value: 0,02	A-B: T Value: 3,02 DF: 318,06 P- Value<0,01	-	-	A-B: T Value: 3,05 DF: 224,35 P- Value<0,01
How do you choose your research topics? - I choose my research topics based on the directions provided by the most impactful authors in my field	Vertical Means	3,08 2,84	2,83 3,19	2,83 3,11	2,64 3,04	2,81 3,04	2,78 3,09
	T-Test	A-B: T Value: 2,60 DF: 243,90 P-Value: 0,01	A-B: T Value: 4,31 DF: 285,87 P- Value<0,01	A-B: T Value: 3,36 DF: 322,89 P- Value<0,01	A-B: T Value: 3,70 DF: 66,19 P- Value<0,01	A-B: T Value: 2,39 DF: 137,72 P-Value: 0,02	A-B: T Value: 3,45 DF: 219,07 P- Value<0,01
Why do you move away from a research topic? - Because it is no longer a "hot topic" in my field	Vertical Means	2,96 2,74	2,84 2,92	2,81 2,94	2,6 2,92	2,71 2,93	2,74 2,95
	T-Test	A-B: T Value: 2,24 DF: 267,26 P-Value: 0,03	-	-	A-B: T Value: 2,16 DF: 56,40 P-Value: 0,04	-	A-B: T Value: 2,08 DF: 218,27 P-Value: 0,04
Why do you move away from a research topic? - Because peer/senior members in my research group advised me to do so	Vertical Means	3,51 3,27	3,28 3,6	3,23 3,57	3,29 3,43	3,24 3,47	3,15 3,55
	T-Test	A-B: T Value: 2,44 DF: 234,47 P-Value: 0,02	A-B: T Value: 3,69 DF: 321,37 P- Value<0,01	A-B: T Value: 3,83 DF: 290,77 P- Value<0,01	-	A-B: T Value: 2,04 DF: 117,72 P-Value: 0,04	A-B: T Value: 3,98 DF: 186,04 P- Value<0,01
What are the drivers that make you think that a publication is worth reading? - A publication is worth reading when it has a high number of citations	Vertical Means	2,16 2,1	2,04 2,27	2,03 2,23	1,8 2,19	1,84 2,23	1,94 2,24
	T-Test	-	A-B: T Value: 2,27 DF: 263,57 P-Value: 0,02	A-B: T Value: 2,03 DF: 320,37 P-Value: 0,04	A-B: T Value: 3,52 DF: 71,49 P- Value<0,01	A-B: T Value: 4,10 DF: 162,96 P- Value<0,01	A-B: T Value: 3,22 DF: 253,51 P- Value<0,01
What are the drivers that make you think that a publication is worth reading? - A publication is worth reading when at least one of the authors has a high h-index	Vertical Means	3,14 3,05	3,08 3,13	2,97 3,21	2,93 3,13	2,92 3,16	3,01 3,15
	T-Test	-	-	A-B: T Value: 2,49 DF: 314,78 P-Value: 0,01	-	-	-
What are the drivers that make you think that a publication is worth reading? - A publication is worth reading when it is published in a journal that has a high impact factor	Vertical Means	2,26 2,07	2,17 2,2	2,19 2,18	2,02 2,21	1,94 2,26	2,1 2,23
	T-Test	-	-	-	-	A-B: T Value: 3,15 DF: 160,14 P- Value<0,01	-
How do you choose the citations to be included in your own	Vertical Means	2,93 2,63	2,75 2,9	2,75 2,87	2,6 2,85	2,61 2,88	2,69 2,88

publications? - I choose citations based on the impact factor of their publication outlet	T-Test	A-B: T Value: 2,84 DF: 261,19 P-Value<0,01	-	-	-	A-B: T Value: 2,17 DF: 121,12 P-Value: 0,03	-
How do you choose the citations to be included in your own publications? - I choose citations where at least one of the authors has a high h-index	Vertical Means	3,66 3,5	3,54 3,68	3,49 3,69	3,38 3,63	3,51 3,63	3,45 3,68
	T-Test	A-B: T Value: 2,18 DF: 243,87 P-Value: 0,03	-	A-B: T Value: 2,80 DF: 292,54 P-Value: 0,01	A-B: T Value: 2,46 DF: 58,05 P-Value: 0,02	-	A-B: T Value: 2,90 DF: 180,99 P-Value<0,01
How do you choose the citations to be included in your own publications? - I choose citations based on my target publication outlet	Vertical Means	2,39 2,46	2,33 2,54	2,32 2,5	2,38 2,42	2,33 2,45	2,32 2,47
	T-Test	-	A-B: T Value: 2,18 DF: 279,17 P-Value: 0,03	-	-	-	-
How do you choose the citations to be included in your own publications? - I make an effort to include highly cited papers among my citations	Vertical Means	2,9 2,73	2,78 2,91	2,73 2,93	2,67 2,86	2,73 2,87	2,67 2,92
	T-Test	-	-	-	-	-	A-B: T Value: 2,28 DF: 214,19 P-Value: 0,02

Tab. 15 Frequenze reali-attese e relativo test χ^2 d'indipendenza per l'analisi sull'utilizzo dei motori di ricerca nella scelta del proprio topic

	Do you use on-line archives and search engines to choose a research topic? - I do not use them to choose a research topic		Do you use on-line archives and search engines to choose a research topic? - I use Google Scholar		Do you use on-line archives and search engines to choose a research topic? - I use Google's general purpose search engine		Do you use on-line archives and search engines to choose a research topic? - I use Scopus		Do you use on-line archives and search engines to choose a research topic? - I use Web of Science		Do you use on-line archives and search engines to choose a research topic? - I use my field's vertical search engines (e.g. IEEE for engineering, Pubmed for medicine, etc.)		
	Strongly Agree	Weakly Disagree, Strongly Disagree	Strongly Agree	Weakly Disagree, Strongly Disagree	Strongly Agree	Weakly Disagree, Strongly Disagree	Strongly Agree	Weakly Disagree, Strongly Disagree	Strongly Agree	Weakly Disagree, Strongly Disagree	Strongly Agree	Weakly Disagree, Strongly Disagree	
Cluster %	60,74 %	39,26%	57,36 %	42,64%	45,71 %	54,29%	13,80 %	86,20%	24,23 %	75,77%	33,74 %	66,26%	
How would you define a "hot research topic"? - A research topic is hot if many researchers are working on it	Strongly Agree	-0,53	0,53	8,72	-8,72	10,41	-10,41	1,61	-1,61	5,47	-5,47	0,48	-0,48
	Weakly Agree	1,04	-1,04	-2,63	2,63	-4,21	4,21	1,19	-1,19	-2,29	2,29	5,69	-5,69
	Weakly Disagree	2,17	-2,17	-4,78	4,78	-5,17	5,17	-2,28	2,28	-3,51	3,51	-4,46	4,46
	Strongly Disagree	-2,68	2,68	-1,31	1,31	-1,03	1,03	-0,52	0,52	0,33	-0,33	-1,71	1,71
	Chi Square	3,42*		6,30*		7,78		1,86*		3,69*		4,99*	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,33		0,1		0,05		0,6		0,3		0,17	
How would you define a "hot research topic"? - A research	Strongly Agree	-1,45	1,45	-1,48	1,48	-7,22	7,22	1,85	-1,85	0,67	-0,67	-5,69	5,69
	Weakly Agree	-9,64	9,64	7,12	-7,12	10,56	-10,56	-3,46	3,46	-1,17	1,17	5,42	-5,42
	Weakly Disagree	6,41	-6,41	-5,89	5,89	-5,56	5,56	0,96	-0,96	-0,39	0,39	0,01	-0,01

topic is hot if papers focusing on it are more likely to be cited	Strongly Disagree	4,67	-4,67	0,25	-0,25	2,23	-2,23	0,65	-0,65	0,88	-0,88	0,26	-0,26
	Chi Square	10,41		3,36		8,31		1,35*		0,34*		2,6	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,02		0,34		0,04		0,72		0,95		0,46	
How do you choose your research topics? - I do research on what is an "hot topic" in my fields	Always	1.07	-1.07	0.67	-0.67	1.77	-1.77	1.52	-1.52	3.64	-3.64	1.93	-1.93
	Frequently	-7.56	7.56	2.86	-2.86	8.13	-8.13	6.92	-6.92	2.25	-2.25	2.80	-2.80
	Seldom	4.04	-4.04	3.37	-3.37	-2.21	2.21	-5.81	5.81	-2.29	2.29	-1.31	1.31
	Never	2.46	-2.46	-6.90	6.90	-7.68	7.68	-2.62	2.62	-3.60	3.60	-3.41	3.41
	Chi Square	3,87		10,89		15,38		8,98*		8,12*		3,98	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,28		0,01		0		0,03		0,04		0,26	
How do you choose your research topics? - I do research on topics that are most likely to be cited	Always	-1,25	1,25	0,98	-0,98	0,8	-0,8	1,03	-1,03	1,3	-1,3	-1,36	1,36
	Frequently	-4,66	4,66	3,37	-3,37	3,87	-3,87	2,58	-2,58	3,19	-3,19	6,63	-6,63
	Seldom	-2,33	2,33	6,75	-6,75	2,73	-2,73	1,15	-1,15	3,38	-3,38	-2,74	2,74
	Never	8,24	-8,24	-11,11	11,11	-7,4	7,4	-4,76	4,76	-7,87	7,87	-2,53	2,53
	Chi Square	7,89*		12,41*		5,70*		5,86*		9,18*		4,14*	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,05		0,01		0,13		0,12		0,03		0,25	
How do you choose your research topics? - I choose my research topics as a consequence of suggestions from peer/senior members of my research group	Always	-1.47	1.47	2.84	-2.84	2.89	-2.89	0.76	-0.76	0.82	-0.82	2.96	-2.96
	Frequently	2.16	-2.16	1.10	-1.10	2.24	-2.24	2.99	-2.99	0.92	-0.92	2.64	-2.64
	Seldom	-4.57	4.57	5.79	-5.79	7.33	-7.33	-1.95	1.95	1.47	-1.47	5.35	-5.35
	Never	3.88	-3.88	-9.73	9.73	-12.45	12.45	-1.80	1.80	-3.21	3.21	-10.96	10.96
	Chi Square	2,67*		10,03*		14,12*		1,98*		1,32*		13,09*	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,45		0,02		0		0,58		0,72		0	
How do you choose your research topics? - I choose my research topics based on the directions provided by the most impactful authors in my field	Always	-2,86	2,86	1,41	-1,41	0,34	-0,34	-0,1	0,1	2,06	-2,06	2,3	-2,3
	Frequently	-7,55	7,55	6,98	-6,98	5,72	-5,72	6,65	-6,65	-0,17	0,17	5,69	-5,69
	Seldom	5,64	-5,64	10,94	-10,94	10,24	-10,24	2,33	-2,33	7,95	-7,95	4,02	-4,02
	Never	4,77	-4,77	-19,33	19,33	-16,31	16,31	-8,87	8,87	-9,84	9,84	-12,02	12,02
	Chi Square	9,43*		24,56*		16,97*		12,94*		11,22*		12,87*	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,02		0		0		0		0,01		0	
Why do you move away from a research	Strongly Agree	-3,68	3,68	1,69	-1,69	1,97	-1,97	0,48	-0,48	0,33	-0,33	2,29	-2,29
	Weakly Agree	-0,65	0,65	6,5	-6,5	5,46	-5,46	2,31	-2,31	6,76	-6,76	6,53	-6,53

topic? - Because peer/senior members in my research group advised me to do so	Weakly Disagree	-5,66	5,66	7,37	-7,37	10,87	-10,87	-0,42	0,42	-0,81	0,81	8,63	-8,63
	Strongly Disagree	9,99	-9,99	-15,56	15,56	-18,3	18,3	-2,37	2,37	-6,29	6,29	-17,45	17,45
	Chi Square	8,90*		13,79*		17,31		1,76*		8,54*		18,57*	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,03		0		0		0,62		0,04		0	
What are the drivers that make you think that a publication is worth reading? - A publication is worth reading when it has a high number of citations	Strongly Agree	3,06	-3,06	0,55	-0,55	5,18	-5,18	3,79	-3,79	7,07	-7,07	5,03	-5,03
	Weakly Agree	-10,18	10,18	11,22	-11,22	1,87	-1,87	4,91	-4,91	5,23	-5,23	8,01	-8,01
	Weakly Disagree	6,52	-6,52	-6,29	6,29	-3,71	3,71	-5,97	5,97	-7,75	7,75	-8,93	8,93
	Strongly Disagree	0,6	-0,6	-5,49	5,49	-3,34	3,34	-2,73	2,73	-4,54	4,54	-4,11	4,11
	Chi Square	6,04		10,28		4,07		9,82*		13,8		11,61	
Degrees of Freedom	3		3		3		3		3		3		
p-value	0,11		0,02		0,25		0,02		0		0,01		
What are the drivers that make you think that a publication is worth reading? - A publication is worth reading when at least one of the authors has a high h-index	Strongly Agree	-0,72	0,72	-1,18	1,18	-0,31	0,31	-0,21	0,21	2,12	-2,12	0,6	-0,6
	Weakly Agree	1,48	-1,48	2,85	-2,85	9,01	-9,01	3,34	-3,34	5,04	-5,04	3,38	-3,38
	Weakly Disagree	-7,77	7,77	1,77	-1,77	3,01	-3,01	1,51	-1,51	-2,44	2,44	1,57	-1,57
	Strongly Disagree	7,01	-7,01	-3,44	3,44	-11,7	11,7	-4,63	4,63	-4,71	4,71	-5,55	5,55
	Chi Square	4,2		1,31		9,13		2,88*		4,71*		1,96	
Degrees of Freedom	3		3		3		3		3		3		
p-value	0,24		0,73		0,03		0,41		0,19		0,58		
What are the drivers that make you think that a publication is worth reading? - A publication is worth reading when it is published in a journal that has a high impact factor	Strongly Agree	-7,94	7,94	-1,45	1,45	-1,82	1,82	3,79	-3,79	5,07	-5,07	5,03	-5,03
	Weakly Agree	0,68	-0,68	3,81	-3,81	3,53	-3,53	-2,98	2,98	3,17	-3,17	-3,29	3,29
	Weakly Disagree	7,91	-7,91	-0,86	0,86	-3,17	3,17	1,89	-1,89	-1,99	1,99	0,73	-0,73
	Strongly Disagree	-0,65	0,65	-1,5	1,5	1,46	-1,46	-2,69	2,69	-6,24	6,24	-2,47	2,47
	Chi Square	7,62		0,82		1,38		4,37*		8,81		2,69	
Degrees of Freedom	3		3		3		3		3		3		
p-value	0,05		0,84		0,71		0,22		0,03		0,44		
How do you choose the citations to be included in your own publications? - I choose citations based on the impact factor of their	Always	-4,58	4,58	2,23	-2,23	1,03	-1,03	1,69	-1,69	6,18	-6,18	1,9	-1,9
	Frequently	-6,38	6,38	2,2	-2,2	1,55	-1,55	2,37	-2,37	-0,69	0,69	3,23	-3,23
	Seldom	3,05	-3,05	0,49	-0,49	3,38	-3,38	-0,08	0,08	-0,72	0,72	1,58	-1,58
	Never	7,91	-7,91	-4,92	4,92	-5,96	5,96	-3,98	3,98	-4,78	4,78	-6,72	6,72
	Chi Square	8,44		2,1		2,25		2,86*		10,05		3,37	
Degrees of Freedom	3		3		3		3		3		3		
p-value	0,04		0,55		0,52		0,41		0,02		0,34		

publication outlet													
How do you choose the citations to be included in your own publications? - I choose citations where at least one of the authors has a high-index	Always	-1,21	1,21	-0,15	0,15	0,09	-0,09	-0,28	0,28	1,52	-1,52	1,33	-1,33
	Frequently	0,25	-0,25	1,95	-1,95	3,4	-3,4	1,1	-1,1	-1,09	1,09	1,91	-1,91
	Seldom	-9,41	9,41	7,39	-7,39	9,06	-9,06	8,54	-8,54	4,89	-4,89	8,99	-8,99
	Never	10,38	-10,38	-9,2	9,2	-12,55	12,55	-9,37	9,37	-5,31	5,31	-12,23	12,23
	Chi Square	9,63*		5,05*		9,11*		11,55*		8,83*		12,11*	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,02		0,17		0,03		0,01		0,03		0,01	
How do you choose the citations to be included in your own publications? - I make an effort to include highly cited papers among my citations	Always	-3,79	3,79	0,09	-0,09	2,12	-2,12	1,41	-1,41	5,7	-5,7	3,23	-3,23
	Frequently	1,52	-1,52	4,65	-4,65	4,49	-4,49	1,16	-1,16	-6,54	6,54	4,62	-4,62
	Seldom	-4,85	4,85	1,03	-1,03	0,44	-0,44	1,13	-1,13	4,13	-4,13	-0,8	0,8
	Never	7,12	-7,12	-5,77	5,77	-7,05	7,05	-3,7	3,7	-3,29	3,29	-7,04	7,04
	Chi Square	5,59		2,47		3,75		2,11*		10,76		5,25	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,13		0,48		0,29		0,55		0,01		0,15	
How do you choose the citations to be included in your own publications? - I cite authors who are most likely to cite me back	Always	-0,43	0,43	-1,29	1,29	-0,83	0,83	-0,55	0,55	1,03	-1,03	0,65	-0,65
	Frequently	2,07	-2,07	-0,33	0,33	0,77	-0,77	0,52	-0,52	-3,36	3,36	0,93	-0,93
	Seldom	-4,05	4,05	9,01	-9,01	8,12	-8,12	2,58	-2,58	4,22	-4,22	3,42	-3,42
	Never	2,41	-2,41	-7,39	7,39	-8,06	8,06	-2,54	2,54	-1,89	1,89	-4,99	4,99
	Chi Square	2,42*		8,10*		6,26*		1,91*		6,53*		2,00*	
	Degrees of Freedom	3		3		3		3		3		3	
	p-value	0,49		0,04		0,1		0,59		0,09		0,57	

Tab. 16 Test U di Mann-Whitney significativi più rilevanti per le fonti di reperimento delle letture

Clusters' Mean Ranks and Mann-Whitney significance	How do you find readings related to your research topics? - I use Google Scholar	How do you find readings related to your research topics? - I use Google's general purpose engine	How do you find readings related to your research topics? - I use Scopus	How do you find readings related to your research topics? - I use Web of Science	How do you find readings related to your research topics? - I use my field's vertical search engines	How do you find readings related to your research topics? - I select interesting readings among the references of papers	
Cluster Frequencies (Frequently Rarely)	210 116	175 151	22 304	39 287	77 249	200 126	
How would you define a "hot research topic"?-A research topic is hot if many researchers are working on it	Frequently	154,82	151,33	145,25	163,54	161,16	159,88
	Rarely	179,22	177,61	164,82	163,49	164,22	169,24
	P-value	0,014	0,006	0,301	0,998	0,784	0,337
How would you define a "hot	Frequently	161,09	166,12	157,73	142,85	171,54	172,92

research topic"?-A research topic is hot if papers focusing on it are more likely to be cited	Rarely	167,86	160,47	163,92	166,31	161,01	148,54
	P-value	0,509	0,566	0,752	0,121	0,363	0,016
How do you choose your research topics?-I do research on what is an "hot topic" in my fields	Frequently	161,38	151,98	139,52	141,03	151,02	163,99
	Rarely	167,34	176,85	165,24	166,55	167,36	162,72
	P-value	0,546	0,009	0,173	0,08	0,142	0,896
How do you choose your research topics?-I do research on topics that are most likely to be cited	Frequently	155,75	155,67	140,48	137,72	150,27	168,64
	Rarely	177,53	172,58	165,17	167	167,59	155,34
	P-value	0,026	0,072	0,187	0,043	0,117	0,167
How do you choose your research topics?-I choose my research topics as a consequence of suggestions from peer/senior members of my research group	Frequently	162,14	162,39	141,73	151,46	142,95	159,43
	Rarely	165,96	164,78	165,08	165,14	169,86	169,96
	P-value	0,705	0,805	0,225	0,358	0,018	0,288
How do you choose your research topics?-I choose my research topics based on the directions provided by the most impactful authors in my field	Frequently	153,88	156,52	123,95	147,05	144,67	160,17
	Rarely	180,92	171,59	166,36	165,74	169,32	168,79
	P-value	0,007	0,12	0,028	0,21	0,03	0,386
Why do you move away from a research topic?-Because peer/senior members in my research group advised me to do so	Frequently	156,73	154,29	135,14	138,72	131,92	162,95
	Rarely	175,75	174,18	165,55	166,87	173,27	164,38
	P-value	0,048	0,031	0,097	0,047	0	0,879
Why do you move away from a research topic?-Because of a personal loss of intellectual interest	Frequently	160,8	160,82	164,18	154,56	164,67	155,12
	Rarely	168,39	166,61	163,45	164,71	163,14	176,81
	P-value	0,434	0,534	0,969	0,478	0,889	0,023
What are the drivers that make you think that a publication is worth reading?-A publication is worth reading when it has a high number of citations	Frequently	155,7	162,1	103	121,99	145,05	162,69
	Rarely	177,63	165,12	167,88	169,14	169,2	164,79
	P-value	0,03	0,756	0,001	0,002	0,034	0,833
What are the drivers that make you think that a publication is worth reading?-A publication is worth reading when at least one of the authors has a high h-index	Frequently	161,05	155,74	112,68	127,78	156,27	165,8
	Rarely	167,93	172,49	167,18	168,35	165,74	159,85
	P-value	0,503	0,089	0,005	0,007	0,413	0,555
How do you choose the citations to be included in your own publications?-I choose citations based on the impact factor of their publication outlet	Frequently	156,5	156,96	142,91	130,76	151,58	167,68
	Rarely	176,17	171,08	164,99	167,95	167,19	156,87
	P-value	0,059	0,158	0,266	0,015	0,183	0,29
How do you choose the citations to be included in your own publications?-I choose citations where at least one of the authors has a high h-index	Frequently	156,88	152,75	118,84	138,1	145,87	163,49
	Rarely	175,49	175,96	166,73	166,95	168,95	163,52
	P-value	0,038	0,007	0,005	0,029	0,022	0,996
How do you choose the citations to be included in your own publications?-I choose citations based on my target publication outlet	Frequently	153,59	154,79	136,91	131,54	151,91	165,4
	Rarely	181,45	173,59	165,42	167,84	167,08	160,48
	P-value	0,007	0,056	0,145	0,016	0,189	0,625
How do you choose the citations to be included in your own publications?-I make an effort to include highly cited papers among my citations	Frequently	152,4	159,98	134,32	130,79	139,46	169,19
	Rarely	183,59	167,58	165,61	167,94	170,93	154,46
	P-value	0,003	0,446	0,115	0,015	0,007	0,15
How do you choose the	Frequently	153,63	150,82	138,23	135,14	149,36	159,46

citations to be included in your own publications?-I cite authors who are most likely to cite me back	Rarely	181,36	178,19	165,33	167,35	167,87	169,92
	P-value	0,001	0,001	0,087	0,008	0,048	0,2

Tab. 17 Test U di Mann-Whitney significativi più rilevanti sull'utilizzo di tools nella scelta del proprio topic di ricerca

Clusters' Mean Ranks and Mann-Whitney significativity		Do you use on-line archives and search engines to choose a research topic?- I do not use them to choose a research topic	Do you use on-line archives and search engines to choose a research topic?- I use Google Scholar	Do you use on-line archives and search engines to choose a research topic?- I use Google's general purpose engine	Do you use on-line archives and search engines to choose a research topic? - I use Scopus	Do you use on-line archives and search engines to choose a research topic? - I use Web of Science	Do you use on-line archives and search engines to choose a research topic? - I use my field's vertical search engines
Cluster Frequencies (Agree Disagree)		198 128	187 139	149 177	45 281	79 247	110 216
How would you define a "hot research topic"?-A research topic is hot if many researchers are working on it	Agree	163,35	153,65	149,5	152,31	149,96	157,25
	Disagree	163,73	176,75	175,28	165,29	167,83	166,69
	P-value	0,969	0,016	0,007	0,346	0,107	0,347
How would you define a "hot research topic"?-A research topic is hot if papers focusing on it are more likely to be cited	Agree	171,67	161,14	167,3	163,44	163,75	169,81
	Disagree	150,86	166,68	160,3	163,51	163,42	160,28
	P-value	0,038	0,577	0,477	0,996	0,977	0,359
How do you choose your research topics?-I do research on what is an "hot topic" in my fields	Agree	168,94	157,24	148,45	128,76	145,25	153,25
	Disagree	155,09	171,92	176,17	169,06	169,34	168,72
	P-value	0,153	0,125	0,004	0,003	0,029	0,122
How do you choose your research topics?-I do research on topics that are most likely to be cited	Agree	172,6	153,3	153,33	139,43	143,61	155
	Disagree	149,42	177,23	172,06	167,35	169,86	167,83
	P-value	0,016	0,012	0,047	0,04	0,016	0,196
How do you choose your research topics?-I choose my research topics as a consequence of suggestions from peer/senior members of my research group	Agree	165,68	154,2	148,73	147,84	155,63	144,48
	Disagree	160,13	176,01	175,93	166,01	166,02	173,19
	P-value	0,575	0,025	0,005	0,194	0,357	0,005
How do you choose your research topics?-I choose my research topics based on the directions provided by the most impactful authors in my field	Agree	173,12	145,42	145,39	122,78	144,51	140,93
	Disagree	148,61	187,82	178,75	170,02	169,57	175
	P-value	0,013	0	0,001	0,001	0,026	0,001
Why do you move away from a research topic?-Because it is no longer a "hot topic" in my field	Agree	172,32	159,44	155,8	136,11	147,49	148,41
	Disagree	149,86	168,96	169,98	167,89	168,62	171,18
	P-value	0,027	0,341	0,153	0,027	0,067	0,029
Why do you move away from a research topic?-Because peer/senior members in my research group advised me to do so	Agree	172,37	148,89	142,86	152,03	146,66	135,78
	Disagree	149,79	183,15	180,88	165,34	168,89	177,62
	P-value	0,017	0	0	0,319	0,038	0

What are the drivers that make you think that a publication is worth reading?- A publication is worth reading when it has a high number of citations	Agree	165,88	154,72	153,08	129,12	132,88	143,09
	Disagree	159,82	175,31	172,27	169,01	173,29	173,89
	P-value	0,541	0,036	0,049	0,005	0	0,003
What are the drivers that make you think that a publication is worth reading?- A publication is worth reading when at least one of the authors has a high h-index	Agree	167,56	160,78	149,06	145,26	147,25	154,04
	Disagree	157,21	167,15	175,66	166,42	168,7	168,32
	P-value	0,304	0,521	0,007	0,137	0,061	0,169
What are the drivers that make you think that a publication is worth reading?- A publication is worth reading when it is published in a journal that has a high impact factor	Agree	171,87	162,6	164,12	149,06	140,94	155,48
	Disagree	150,55	164,72	162,97	165,81	170,71	167,58
	P-value	0,033	0,83	0,907	0,237	0,009	0,242
How do you choose the citations to be included in your own publications?-I choose citations based on the impact factor of their publication outlet	Agree	174,67	157,68	157,33	143,03	145,25	151,54
	Disagree	146,22	171,33	168,7	166,78	169,34	169,59
	P-value	0,005	0,176	0,255	0,1	0,038	0,087
How do you choose the citations to be included in your own publications?-I choose citations where at least one of the authors has a high h-index	Agree	171,77	155,56	149,51	131,08	152,81	144,98
	Disagree	150,71	174,19	175,27	168,69	166,92	172,93
	P-value	0,017	0,032	0,003	0,003	0,159	0,002
How do you choose the citations to be included in your own publications?-I choose citations based on my target publication outlet	Agree	159,83	154,38	153,46	160,33	155,24	154,2
	Disagree	169,17	175,77	171,95	164,01	166,14	168,24
	P-value	0,352	0,031	0,06	0,796	0,341	0,176
How do you choose the citations to be included in your own publications?-I make an effort to include highly cited papers among my citations	Agree	169,56	157,64	153,14	147,18	155,99	147,71
	Disagree	154,13	171,38	172,22	166,11	165,9	171,54
	P-value	0,13	0,172	0,056	0,19	0,394	0,024