

UNIVERSITÀ DEGLI STUDI DI UDINE

LA VALUTAZIONE DELLA PERFORMANCE NEI SETTORI BIBLIOMETRICI

La valutazione dell'attività di ricerca di ricercatori e docenti inquadrati in settori scientifico-disciplinari "bibliometrici", nonché la relativa aggregazione in funzione della composizione dei Collegi di Dottorato, è stata condotta secondo le specifiche di seguito riportate.

1.1 La costruzione del dataset di analisi

Il dataset bibliometrico è stato costruito a partire dalla *Core Collection* di Web of Science. Mediante l'applicazione di un complesso algoritmo di riconoscimento e omologazione degli "address" e di disambiguazione della reale identità degli autori, è stato possibile attribuire ciascuna pubblicazione ai ricercatori accademici che l'hanno realizzata. Tale algoritmo si basa sull'accoppiamento dei dati estratti da WoS con i record del database CINECA sul personale di ruolo degli atenei italiani al 31/12 di ciascun anno. La media armonica¹ di *precision* e *recall* (F-measure) delle authorships disambiguate dall'algoritmo è pari a circa il 97%. Un successivo intervento manuale di filtraggio per individuare e rimuovere le false attribuzioni, generate principalmente da problemi di omonimia, riduce ulteriormente tale tasso di errore.

In particolare, una pubblicazione² viene attribuita a un determinato ricercatore/docente se questi:

1. ha un nominativo compatibile³ con uno degli autori della pubblicazione,
2. era in ruolo al 31/12 dell'anno precedente a quello della pubblicazione,
3. afferisce a una delle università riconosciute nella lista degli address indicati dagli autori della pubblicazione,
4. è inquadrato in un settore scientifico-disciplinare (SSD) compatibile con la subject category della pubblicazione.

¹ Misurata su base campionaria, con un errore di campionamento del 2% e un livello di confidenza, 98%.

² Articles, reviews, letters e proceeding papers. Per motivi di significatività vengono esclusi editorials, meeting abstract, corrections, reprint e ogni document type non associabile a lavori di ricerca propriamente detti.

³ Il formato dei nominativi in CINECA e in WoS è molto diverso. In WoS un autore è indicato nel formato "SURNAME INITIALS": il ricercatore che in CINECA compare come ROSSI Giovanni Maria, come autore di una pubblicazione WoS comparirà molto probabilmente come ROSSI GM. La presenza delle sole iniziali del first name dell'autore amplifica considerevolmente il problema delle omonimie e la complessità del processo di disambiguazione risultante. Inoltre, nel caso di nomi o cognomi multipli (come nell'esempio precedente), alcuni di questi possono essere non riportati in una delle due fonti, o riportati in un ordine diverso. Per non escludere nessun possibile mapping, l'algoritmo proposto assegna una relazione esplorando tutti i possibili criteri di matching (ROSSI GM; ROSSI G; ROSSI M; ROSSI MG).

Per tutti i ricercatori e docenti dell'Università di Udine è stata altresì implementata una verifica puntuale della produzione scientifica individuale, consultando fonti terze rispetto a Web of Science, al fine di correggere ulteriormente eventuali false o mancate attribuzioni.

In particolare per i ricercatore di recente nomina, non in ruolo nel quinquennio in osservazione e per tutti i componenti dei Collegi, esterni ai ruoli accademici, è stato realizzato un censimento della produzione scientifica individuale tramite interrogazioni nominative di WoS.

1.2 La valutazione comparata della performance bibliometrica

La valutazione comparata della performance bibliometrica di singoli ricercatori/docenti è basata sui seguenti indicatori:

- **Output (O)**, numero medio annuo di pubblicazioni realizzate dal ricercatore/docente nel periodo di osservazione.
- **Fractional Output (FO)**, come PO, ma considerando per ciascuna pubblicazione, il reciproco del numero di co-autori e, per le scienze della vita, la posizione nella lista e il carattere della co-authorship (intramuros vs extramuros).
- **Average Impact (IA, IJ)**, per tutte le pubblicazioni di un ricercatore/docente, IA è pari al valore medio della combinazione pesata di citazioni and journal impact factor delle proprie pubblicazioni, entrambi standardizzati alle rispettive distribuzioni di riferimento (anno di pubblicazione e subject category)⁴; IJ è analogo a IA, ma calcolato solo tramite l'impact factor della rivista⁵.
- **Scientific Strength (SS)**, somma dell'impatto standardizzato di tutte le pubblicazioni realizzate dal ricercatore/docente nel periodo di osservazione, calcolato con la combinazione pesata di citazioni and journal impact factor sopra descritta.
- **Fractional Scientific Strength (FSS)**, come SS, ma considerando per ciascuna pubblicazione, il reciproco del numero di co-autori e, per le scienze della vita, la posizione nella lista e il carattere della co-authorship (intramuros vs extramuros).

Per ciascun ricercatore/docente vengono computati in valore assoluto gli indicatori sopra descritti e:

- ✓ Per comparazione con gli stessi dati riferiti a tutti i ricercatori/docenti accademici nazionali dello stesso SSD, si ottiene il rank (assoluto e percentile) della performance del singolo scientist.
- ✓ Per comparazione con gli stessi dati riferiti a tutti i scientist accademici nazionali dello stesso SSD e fascia di inquadramento, si ottiene il rank (assoluto e percentile) della performance del singolo scientist.

⁴ Abramo, G., D'Angelo, C.A., & Felici, G. (2019). Predicting long-term publication impact through a combination of early citations and journal impact factor. *Journal of Informetrics*, 13(1), 32-49.

⁵ Tale indicatore rivela il prestigio medio delle riviste in cui pubblica l'accademico.