

POLITECNICO DI TORINO

DIPARTIMENTO DI INGEGNERIA GESTIONALE E DELLA PRODUZIONE



**Politecnico
di Torino**

Corso di Laurea Magistrale in Ingegneria Gestionale

Tesi di Laurea Magistrale

Relazione tra ambito di specializzazione e collaborazione interuniversitaria nella ricerca scientifica italiana

Relatore

Prof. Federico Caviggioli

Candidato

Davide Ponzetti

Anno Accademico 2022/2023

Indice

Elenco delle figure.....	4
Elenco delle tabelle.....	7
Introduzione.....	9
Capitolo I.....	11
1.1. Ricerca scientifica nel mondo.....	11
1.2. Ricerca scientifica in Italia.....	13
1.3. Il ruolo delle università.....	15
1.4. Collaborazione nella ricerca scientifica.....	16
1.5. Materiali e dataset.....	20
Capitolo II.....	27
2.1. Ambiti di ricerca in Italia.....	27
2.2. Ricerca scientifica negli istituti italiani.....	30
2.3. Indice di specializzazione.....	37
2.4. Indice di collaborazione.....	45
2.5. Confronto specializzazione e collaborazione.....	52
Capitolo III.....	63
3.1. Analisi di regressione.....	63
3.2. Modello impiegato.....	64
3.3. I risultati del modello.....	67
Conclusioni.....	72
Appendice.....	75
Bibliografia.....	79

Elenco delle figure

Figura 1 – Evoluzione della spesa in ricerca e sviluppo come % del PIL per le principali economie mondiali dal 1996 al 2020 (fonte The World Bank Data).....	12
Figura 2 – Spesa in ricerca e sviluppo come % del PIL per i Paesi appartenenti all’Unione Europea nel 2020 (fonte Eurostat)	13
Figura 3 – Spesa in ricerca e sviluppo come % del PIL per l’Italia e l’Europa tra il 2000 e il 2020 (fonte Eurostat)	14
Figura 4 – Interazione strategica Triple Helix in un Paese sviluppato.....	18
Figura 5 – Numero di occorrenze per tipologia di documento all’interno del database	23
Figura 6 – Numero di istituti accademici con almeno una pubblicazione nell’anno considerato.....	24
Figura 7 – Distribuzione del numero di articoli per le 30 aree di ricerca con più pubblicazioni tra il 2015 e 2020	27
Figura 8 – Distribuzione del numero di articoli cooperativi per le 30 aree di ricerca più collaborative (2015-2020).....	28
Figura 9 – Distribuzione del numero di articoli scientifici pubblicati per ciascun istituto analizzato (2015-2020)	30
Figura 10 – Distribuzione del numero di aree di ricerca trattate rispetto al numero di articoli scientifici pubblicati per ciascuna università (2015-2020).....	32
Figura 11 – Rappresentazione della distribuzione dell’indice di concentrazione HHI ..	33
Figura 12 – Distribuzione indice di concentrazione HHI rispetto al numero di pubblicazioni totali per ciascuna università (2015-2020).....	34
Figura 13 – Distribuzione indice di concentrazione HHI rispetto al numero di aree di ricerca per ciascuna università (2015-2020).....	35
Figura 14 – Distribuzione numero di articoli collaborativi rispetto al numero di pubblicazioni totali per ciascuna università (2015-2020).....	36
Figura 15 – Distribuzione % articoli collaborativi rispetto al numero di pubblicazioni totali per ciascuna università (2015-2020).....	36

Figura 16 – Distribuzione indice di specializzazione RSA rispetto al numero di articoli pubblicati per la categoria “Engineering” per ogni anno compreso tra il 2015 e il 2020	40
Figura 17 – Distribuzione indice di specializzazione RSA rispetto al numero di articoli pubblicati nell’anno 2019 per le categorie: “Clinical Medicine”, “Economics”, “Law, Political and Social Sciences” e “Physics”	41
Figura 18 – Andamento del numero di università sovra specializzate in ciascuna macro categoria (2015-2020).....	43
Figura 19 – Andamento del numero di università per numero di categorie in cui sono sovra specializzate (2015-2020)	43
Figura 20 – Distribuzione del numero medio di categorie in cui si è specializzata rispetto al numero totale di articoli pubblicati per ciascuna università (2015-2020)	45
Figura 21 – Distribuzione indice di collaborazione RCA rispetto al numero di articoli pubblicati per la categoria “Engineering” per ogni anno compreso tra il 2015 e il 2020	47
Figura 22 – Distribuzione indice di collaborazione RCA rispetto al numero di articoli pubblicati nell’anno 2019 per le categorie: “Clinical Medicine”, “Economics”, “Law, Political and Social Sciences” e “Physics”	48
Figura 23 – Distribuzione indice di specializzazione RSA rispetto all’indice di collaborazione RCA per la categoria “Engineering” per ogni anno compreso tra il 2015 e il 2020	49
Figura 24 – Distribuzione indice di specializzazione RSA rispetto all’indice di collaborazione RCA nell’anno di pubblicazione 2019 per le categorie: “Clinical Medicine”, “Economics”, “Law, Political and Social Sciences” e “Physics”	50
Figura 25 – Andamento del numero di università maggiormente collaborative in ciascuna macro categoria (2015-2020).....	51
Figura 26 – Andamento del numero di università per numero di categorie in cui sono maggiormente collaborative (2015-2020)	51
Figura 27 – Classificazione delle università sulla base dei valori associati di RSA e RCA	55

Figura 28 – Andamento del numero di università maggiormente collaborative per l'indice RCA2 in ciascuna macro categoria (2015-2020).....	58
Figura 29 – Andamento del numero di università per numero di categorie in cui sono maggiormente collaborative secondo l'indice RCA2 (2015-2020).....	59
Figura 30 – Suddivisione delle università del database per macro regione.....	66

Elenco delle tabelle

Tabella 1 – Informazioni riguardanti la produzione scientifica italiana tramite la suddivisione in macro categorie (2015-2020)	25
Tabella 2 – Prime dieci aree di ricerca per rapporto articoli collaborativi su articoli totali	29
Tabella 3 – Classificazione per le macro categorie “Biology”, “Chemistry”, “Earth&SpaceSciences” e “Psychology” (2015-2020).....	55
Tabella 4 – Classificazione per le macro categorie “Biology”, “Chemistry”, “Earth&SpaceSciences” e “Psychology” rispetto all’indice RCA2 (2015-2020)	60
Tabella 5 – Matrice di correlazione tra i regressori impiegati nel modello di regressione	67
Tabella 6 – Risultati analisi di regressione per i gruppi specificati.....	68

Introduzione

Le università rappresentano, grazie alla loro attività di ricerca, il luogo essenziale dove poter sostenere lo sviluppo scientifico, tecnologico e sociale. Per favorire tale compito, un aspetto che negli ultimi anni ha acquisito un ruolo sempre di maggior rilievo è stata la collaborazione. Data la crescente complessità e interdisciplinarietà della scienza, l'operato della singola università non è stato più sufficiente per apportare cambiamenti significativi. Per tale motivo l'interesse degli studiosi si è concentrato in maniera sempre più consistente verso gli effetti e le conseguenze derivati dal suo rapporto con altri enti, come aziende, istituzioni governative o semplicemente altri istituti accademici. Seguendo tale filone, l'obiettivo del presente elaborato è quello di esaminare l'aspetto della collaborazione interuniversitaria e come esso possa essere influenzato dal livello di specializzazione delle università nei vari ambiti scientifici. Per condurre tale studio si è deciso di restringere l'analisi al contesto scientifico italiano, concentrandosi sul comportamento e sulle tenenze delle università italiane. A partire da un set di dati a disposizione, sono state costruite delle metriche che potessero rappresentare in una maniera quantificabile i fattori della specializzazione e della collaborazione e che potessero essere utilizzate per impostare un confronto. I valori così ottenuti sono stati oggetto di un'analisi dettagliata dalla quale sono state ricavate delle supposizioni riguardo la tipologia di relazione esistente tra questi due fattori. Tali ipotesi, infine, sono state verificate da un punto di vista statistico attraverso l'applicazione di un modello di regressione.

In particolare, nel primo capitolo, dopo una breve presentazione della funzione e del ruolo della ricerca scientifica nel panorama internazionale e italiano e di come le università svolgano un compito essenziale in tale ambito, viene introdotto e analizzato il set di dati di partenza impiegato per lo svolgimento del presente studio. Successivamente, nel secondo capitolo vengono costruite delle metriche per la definizione dell'aspetto della specializzazione e della collaborazione e formulate delle ipotesi sulla loro interazione. Infine, nel terzo e ultimo capitolo, le conclusioni ottenute in precedenza sono state sottoposte ad un controllo statistico, attraverso gli strumenti conseguiti dall'analisi di regressione, per valutarne la consistenza e l'affidabilità.

Capitolo I

1.1. Ricerca scientifica nel mondo

Secondo quanto riportato sul sito online dell'enciclopedia Treccani: «Con la locuzione "ricerca scientifica" comunemente s'intende l'insieme delle attività destinate alla scoperta e utilizzazione delle conoscenze scientifiche. Essa comprende sia la "r. fondamentale", che è lo studio sistematico della natura e delle sue leggi a fini puramente conoscitivi a prescindere da scopi immediatamente pratici, sia la "r. applicata", volta invece a individuare e sperimentare le possibili applicazioni pratiche delle conoscenze acquisite, sia infine l'attività di "sviluppo", legata alla produzione su scala industriale delle innovazioni tecnologiche.» [1]. Essa rappresenta uno dei fattori chiave per lo sviluppo e la crescita della società nel medio-lungo periodo. Mediante la capacità di portare innovazione, per esempio attraverso l'applicazione tecnologica delle sue scoperte scientifiche, si manifesta come il motore dello sviluppo tecnico, scientifico, culturale e sociale.

Da quando è stata posta maggiore attenzione alle implicazioni sociali e di benessere collettivo che la ricerca scientifica poteva apportare alla comunità, soprattutto negli ultimi due secoli, essa ha attirato un notevole interesse e riguardo nei suoi confronti. Sono nate molteplici iniziative volte alla sua diffusione e sviluppo. Inoltre, la presenza di un sistema industriale che necessitava di ingenti quote di capitale per svilupparsi e crescere, ha richiesto un uso più efficiente delle sue risorse, sia umane sia materiali, che ha portato all'introduzione, in dosi sempre più rilevanti, della scienza nel processo produttivo. Tale dinamica ha spinto le imprese private ad occuparsi attivamente della ricerca scientifica e tecnologica, quella che oggi viene definita come *Research and Development* o R&D.

Parallelamente al settore privato, anche quello pubblico e governativo ha attraversato un periodo di cambiamenti. Si è passati da una situazione di astensione ad una in cui i governi intervenivano direttamente e diffusamente indirizzando gli ambiti di investimenti e non solo. La diffusione dei progressi scientifici e tecnologici in ogni aspetto della vita quotidiana e la richiesta di investimenti che impiegassero quantità di risorse economiche e umane sempre più ingenti e per un lungo periodo di tempo, ha reso l'intervento del governano inevitabile per poter conciliare interessi diversi e spesso contrastanti [1].

L'attenzione posta sull'aspetto della ricerca scientifica può essere dimostrata dalla quantità di denaro in essa investita. Un indicatore che spesso viene utilizzato per misurare gli sforzi di innovazione dei singoli Paesi è la spesa pubblica e privata in ricerca e sviluppo, calcolata in percentuale sul PIL (Prodotto Interno Lordo o, in inglese, *Gross Domestic Product*). In Figura 1 viene riportata l'evoluzione di questo valore per i principali Stati del mondo secondo i dati pubblicati sul sito della Banca Mondiale [2].

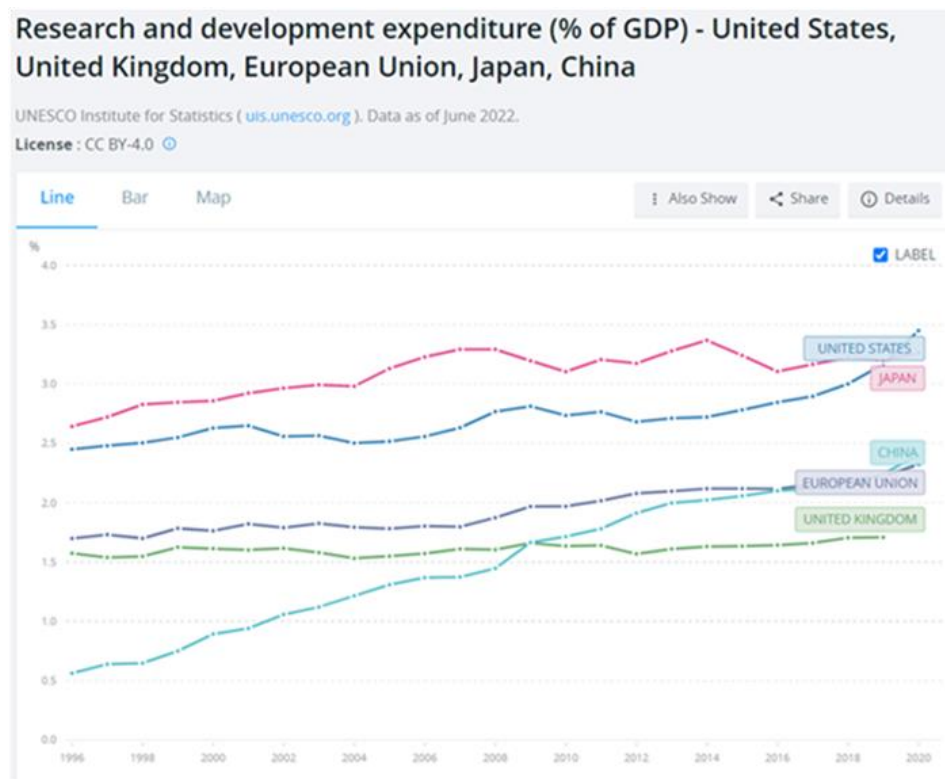


Figura 1 – Evoluzione della spesa in ricerca e sviluppo come % del PIL per le principali economie mondiali dal 1996 al 2020 (fonte The World Bank Data)

Come si evince dalla figura, negli ultimi 25 anni vi è stato un graduale aumento della spesa in questo ambito, con un incremento più significativo negli ultimi 3, tra il 2018 e 2020. A livello mondiale si è partiti da una percentuale che si attestava intorno all'1,96% nel 1996, fino ad arrivare al 2,63% nel 2020. Questo permette di capire come l'attenzione mondiale si stia focalizzando su queste tematiche e come il primo passo per ottenere il progresso tecnologico e sociale sia quello di investire nella ricerca scientifica.

Tuttavia, rimane ancora molta diffomità a livello di singolo Paese. In termini assoluti, gli Stati Uniti sono i *leader* indiscussi di questa classifica, con oltre 720 miliardi di dollari investiti in ricerca e sviluppo nel solo anno 2020 [3], anche se la Cina sta notevolmente riducendo questo divario, in quanto è riuscita a decuplicare la sua spesa nell'arco temporale di due decenni, raggiungendo quota 580 miliardi di dollari [3]. Anche il

Giappone, seppur a livello relativo, riesce a tenere testa, anzi, a performare leggermente meglio degli Stati Uniti. Diversa, invece, appare la situazione in Europa. Nonostante vi sia stato un graduale incremento nel corso degli ultimi anni, il dislivello rimane ancora elevato. Con la strategia Europa 2020 [4], si era concepito un piano per l'uscita dalla recessione iniziata nel 2008 che aveva messo in evidenza alcune delle criticità e contraddizioni che caratterizzavano l'economia continentale. In questo contesto, l'obiettivo di potenziare gli investimenti pubblici e privati in ricerca e sviluppo doveva essere il propulsore della ripresa economica del continente. L'asticella era stata fissata al 3% entro il 2020 ma, come si può evincere sia dalla Figura 1 che dalla Figura 2, la media europea si è fermata al 2,3%. Questo valore è stato sicuramente impattato da Stati importanti come l'Italia e la Spagna che, pur possedendo un'economia a bassa produttività e con un tasso di crescita annua molto al di sotto della media europea, non sono riuscite ad andare oltre, rispettivamente, all'1,5% e 1,4% [5].

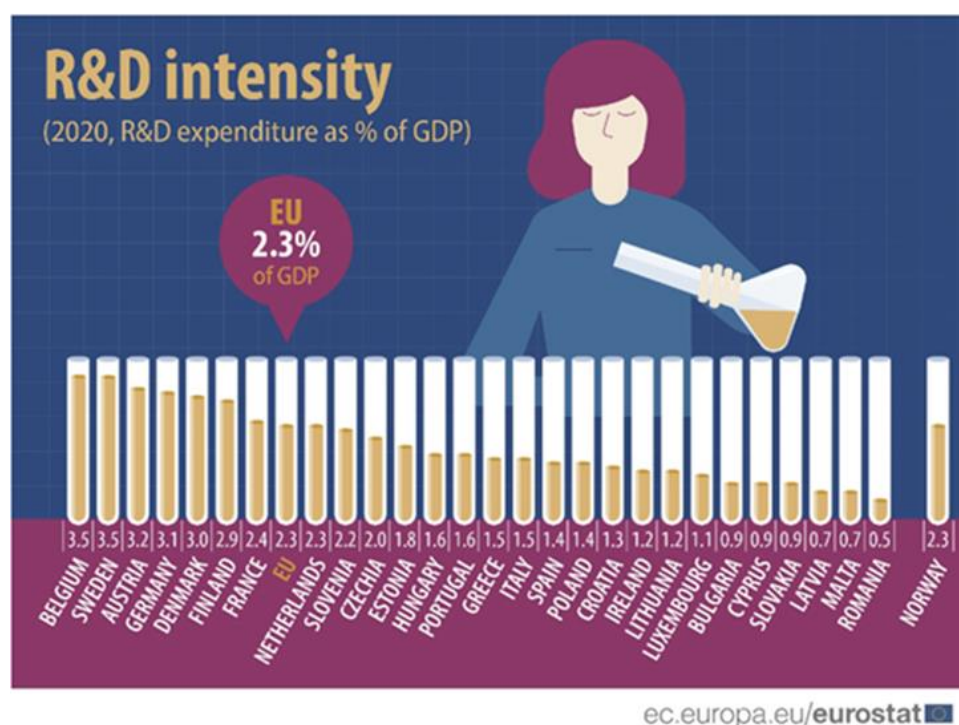


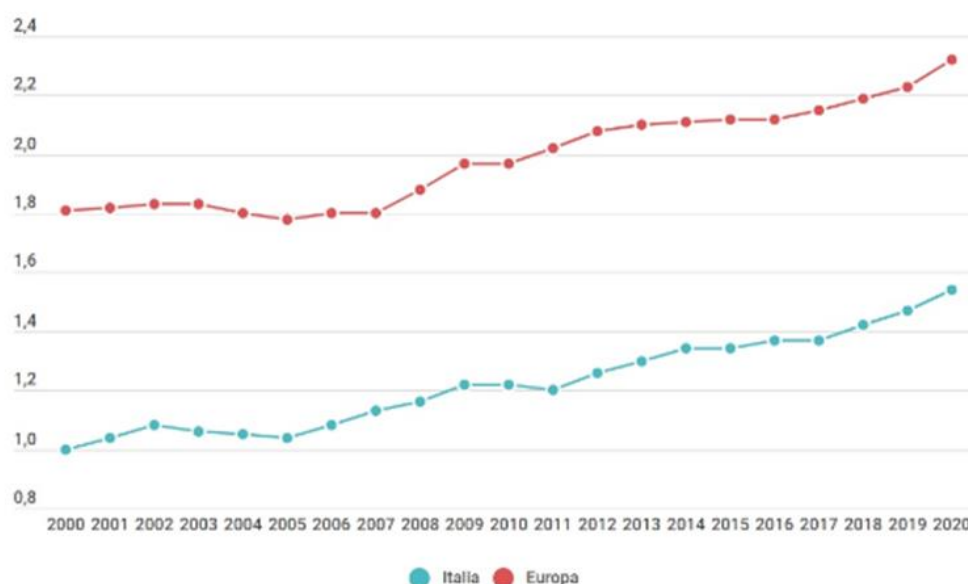
Figura 2 – Spesa in ricerca e sviluppo come % del PIL per i Paesi appartenenti all'Unione Europea nel 2020 (fonte Eurostat)

1.2. Ricerca scientifica in Italia

In Italia, secondo quanto previsto dal decreto legislativo 5 giugno 1998, n. 204, le politiche sulla ricerca vengono definite dal Programma Nazionale della Ricerca (PNR), il

quale stabilisce e coordina gli interventi e i finanziamenti per la ricerca. Tale documento si propone l'obiettivo di guidare la competitività industriale e lo sviluppo del Paese per rafforzarne la presenza sulla scena europea e mondiale [6].

Con il 2020 si è concluso il programma quinquennale, definito PNR 2015-2020 [7], che ha posto una serie di obiettivi innovativi, tra i quali l'investimento in capitale umano per far crescere il numero di ricercatori, una maggiore collaborazione tra settore pubblico e privato, il sostegno alle infrastrutture di ricerca ed una serie di iniziative volte a risollevare il Mezzogiorno. Tramite il programma sono stati stanziati fondi per circa 2,5 miliardi di euro nei primi tre anni che si sono aggiunti ai finanziamenti che normalmente il Ministero dell'Università e della Ricerca indirizza ad Università ed Enti Pubblici di Ricerca. Questo ha permesso di stimolare ulteriormente gli investimenti in Ricerca e Sviluppo. Secondo i dati provenienti dall'Eurostat, infatti, negli ultimi due decenni la percentuale del PIL investita in R&D è cresciuta considerevolmente. Tuttavia, non è stato sufficiente per colmare il gap con le nazioni più avanzate.



Fonte: Eurostat - I dati indicano la percentuale del PIL investita

Figura 3 – Spesa in ricerca e sviluppo come % del PIL per l'Italia e l'Europa tra il 2000 e il 2020 (fonte Eurostat)

Per tale motivo, con il nuovo programma PNR 2021-2027 [8], firmato il 15 dicembre 2020, si è optato di riprendere il filo conduttore del programma precedente, cercando di porre obiettivi ancora più ambiziosi e ponendo una sempre maggiore attenzione alla loro realizzazione. Inoltre, essendo stato tale programma discusso e realizzato durante la

pandemia di Covid-19, è divenuto uno dei principali strumenti per far fronte a questa situazione di emergenza ed ha focalizzato l'attenzione sul fatto che l'investimento nella ricerca sia la migliore soluzione per superare la crisi e rilanciare l'economia del Paese.

1.3. Il ruolo delle università

Dagli studi sulla crescita economica (tra i quali Scherer 1986 [9], Aghion&Howitt 1995 [10] e Aliu Mulaj&Dedaj 2023 [11]) si evince un sempre maggiore progresso tecnico quale motore dello sviluppo. La capacità innovativa di un paese si traduce nel saper offrire beni e servizi caratterizzati da una qualità progressivamente più alta, i quali possono sostenere la crescita di lungo periodo e l'aumento della produttività complessiva, non solo a livello industriale.

L'innovazione è un fenomeno complesso che vede coinvolti molti soggetti e istituzioni. È il risultato di diversi fattori che possono spaziare dall'intuizione umana alla sperimentazione estesa, dall'applicazione degli studi più avanzati fino ad arrivare alla pura casualità. Tuttavia, non c'è dubbio che molto dipenda dalla quantità di risorse impiegate in essa. Gli studi economici, come quello di Mario Coccia del 2008 [12], hanno infatti dimostrato una correlazione tra gli investimenti in ricerca e sviluppo e la crescita del prodotto e della produttività, anche se, come è stato esposto in precedenza, questo dato non ha cancellato le rilevanti differenze che esistono tra Paesi su questo tema.

Dato il ruolo cruciale che la ricerca assume nell'esistenza di un Paese, diventa necessario l'intervento di un operatore pubblico. Da un lato, quest'ultimo, tramite la regolamentazione dei mercati, deve favorire il flusso dei capitali verso i progetti più promettenti, anche se ad alto rischio, in maniera tale che costituiscano la base del sostegno finanziario nel medio e lungo periodo. Negli Stati Uniti, per esempio, è stato verificato che le imprese finanziate da fondi di *venture capitalists* siano riuscite a performare meglio rispetto alle concorrenti che operano nello stesso settore ad alta tecnologia [13]. Dall'altro, il governo deve occuparsi della formazione del proprio capitale umano [14]. La capacità di innovare è condizionata in maniera decisiva dalla qualità e dalla preparazione dei lavoratori, ricercatori e scienziati.

Collegandosi a questo secondo punto, acquista una notevole rilevanza il ruolo dell'istruzione e, in particolare, delle università. Tali istituti, infatti, rappresentano

l'eccellenza nell'educazione e nella formazione dei ricercatori e, per tale motivo, sono fondamentali per la ricerca scientifica [15]. Le università offrono un ambiente ideale per lo sviluppo e la diffusione delle conoscenze scientifiche attraverso la loro attività di ricerca e di insegnamento. Inoltre, gli istituti accademici sono spesso il luogo dove si incontrano e collaborano scienziati di diversi campi, favorendo, così, la nascita di nuove idee e progetti di ricerca. La ricerca universitaria ha un impatto significativo sulla società, contribuendo questa allo sviluppo di nuove tecnologie, cure mediche e soluzioni per i problemi globali [16]. Tramite vari canali, come l'insegnamento, la pubblicazione di ricerche, conferenze e *workshop*, le conoscenze sviluppate nelle università vengono trasmesse all'esterno. Gli studenti che le frequentano hanno l'opportunità di acquisire conoscenze tecniche da esperti del settore, e queste conoscenze possono poi essere trasferite nel mondo del lavoro. In sintesi, le università sono un pilastro fondamentale nella ricerca scientifica e un punto di incontro tra quest'ultima e la società, promuovendo la trasmissione della conoscenza.

Inoltre, le università sono anche un luogo dove la collaborazione tra istituti è incoraggiata e favorita. La collaborazione tra università, centri di ricerca, industrie e altre organizzazioni è fondamentale per raggiungere risultati scientifici di alto livello. Queste collaborazioni permettono di unire le risorse e le competenze di diverse istituzioni, permettendo loro di lavorare insieme su progetti ambiziosi e di interesse comune [17]. In definitiva, la collaborazione tra istituti è un elemento cruciale nella ricerca scientifica e nell'avanzamento delle conoscenze.

1.4. Collaborazione nella ricerca scientifica

Proprio la collaborazione è un aspetto che ha assunto una notevole rilevanza negli ultimi decenni e che ha favorito lo sviluppo e l'espansione della ricerca scientifica nel mondo. Vari fattori possono essere giudicati come responsabili di questo cambiamento. La crescente complessità e interdisciplinarietà della scienza, l'aumento dei costi dei fattori di produzione nei progetti di ricerca, la riduzione della spesa per il trasporto e le innovazioni nelle tecnologie dell'informazione e della comunicazione sono tutti elementi che hanno richiesto o permesso una sempre maggiore collaborazione nell'ambito della ricerca scientifica [18]. In quanto tale, la collaborazione di ricerca è considerata un approccio importante per far progredire la scienza e la tecnologia in tutto il mondo [19], oltre a

sviluppare l'impatto scientifico dei partecipanti, che vanno dai singoli scienziati ai sistemi nazionali [20]. Di conseguenza, la collaborazione nel campo della ricerca è stata uno degli argomenti di tendenza negli studi bibliometrici e scientometrici. Questi studi hanno come scopo principale quello di analizzare le collezioni bibliografiche e i servizi bibliografici tramite strumenti matematici e statistici. L'approccio di tipo scientifico offerto da questa metodologia, che ha come obiettivo quello di estrapolare tra i documenti e gli elementi che li compongono (parole, citazioni, autori, istruzione etc.) indicazioni quantitative, ha rappresentato una garanzia di oggettività dei risultati ottenuti. Con l'accesso a banche dati *online* che hanno facilitato la categorizzazione e la raccolta di tutte le informazioni riguardanti le pubblicazioni scientifiche, i dati conseguiti dall'analisi bibliometrica hanno permesso di valutare e misurare il contributo al progresso scientifico e tecnologico apportato da università, istituti di ricerca e nazioni.

Numerosi lavori [21, 22, 23] hanno studiato ed evidenziato come sussista una relazione positiva tra l'impatto scientifico di un articolo e il numero dei suoi autori. Oltre a questo, sono stati analizzati altri possibili benefici, come il contributo alla visibilità e al riconoscimento dei singoli scienziati, nonché alla capacità di ricerca e alla competitività dei Paesi [19, 20].

La collaborazione scientifica viene solitamente analizzata tramite tecniche di *network analysis* capaci di indagare sui legami esistenti tra i diversi co-autori [24, 25]. Esistono vari tipi di collaborazioni ed è stato dimostrato che diversi fattori scientifici, economici, geopolitici e culturali stanno modellando tale struttura [18].

In questo contesto, le università rappresentano le unità di base di accoglienza degli scienziati e, di conseguenza, il loro operato risulta di notevole importanza per lo sviluppo dell'istruzione accademica e della scienza. In particolare, molti dei progetti di ricerca sono supportati da uno o più istituti e all'interno di essi avviene il primo contatto tra l'istruzione e la ricerca scientifica. Inoltre, poiché ha acquisito sempre maggiore interesse la *performance*, il prestigio e la rete di contatti esterni di una università [26, 27], nella ricerca scientifica si stanno coinvolgendo sempre più attori. Per tale motivo la cooperazione è diventata un elemento chiave per lo sviluppo universitario.

L'ambito su cui si è soffermato il livello più alto di interesse e che ha rappresentato l'oggetto di numerose pubblicazioni è stato il rapporto intercorrente tra università, industria e governo. In merito a questo tema, nella letteratura è stato proposto un modello,

definito come “*Triple Helix model*”, che permette di interpretare l’innovazione come il risultato del processo creativo e collaborativo che coinvolge questi tre principali attori. All’interno di questa teoria, ogni interprete è rappresentato da un cerchio (*helix*) e le sovrapposizioni tra di essi costituiscono le interazioni. Un esempio viene mostrato in Figura 4.

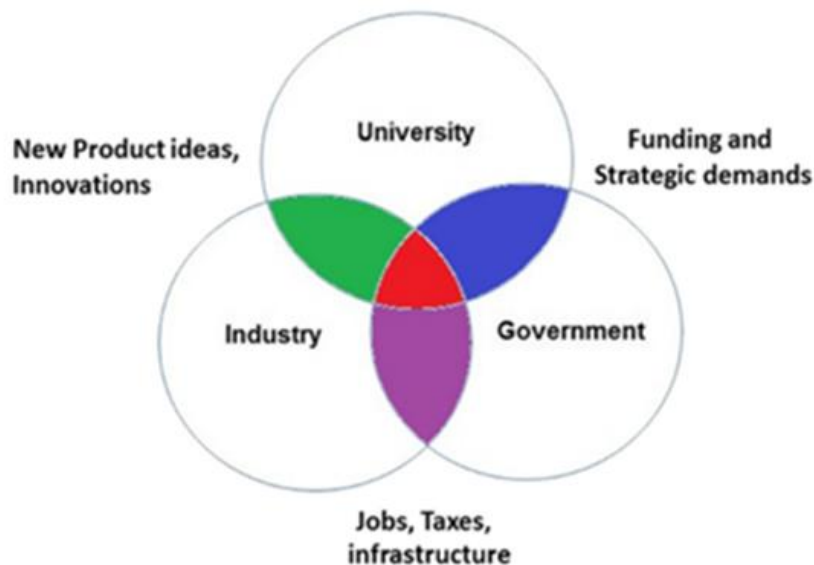


Figura 4 – Interazione strategica Triple Helix in un Paese sviluppato

Il modello iniziale è stato per la prima volta teorizzato da Henry Etzkowitz e Loet Leydesdorff nel 1995 [28] e, nel corso degli anni, è stato oggetto di numerosi studi ed è stato ampiamente adottato dalle autorità per la trasformazione di molti settori.

Secondo il modello, le relazioni esistenti tra questi tre agenti creano una base favorevole alla generazione di sinergie e innovazioni. Il concetto di sinergia si riferisce al raggiungimento, tramite l’azione congiunta di più elementi, di un risultato migliore comparato a quello che sarebbe ottenibile se ognuno lavorasse singolarmente. La cooperazione tra diversi fattori permette di raggiungere benefici superiori rispetto alla somma dei risultati che ognuno otterrebbe separatamente. La modalità più utilizzata per sfruttare le potenziali sinergie è la condivisione di attività o il trasferimento di competenze chiave.

Il modello *Triple Helix* rappresenta uno step successivo ai precedenti modelli lineari caratterizzati da un legame monodirezionale tra accademia e industria, in cui ad ogni attore veniva assegnato un compito specifico, come poteva essere per le università la ricerca di base, mentre per l’industria la produzione di beni commerciali. La nuova teoria

assegna anche un peso maggiore all'azione dell'autorità pubblica rispetto ai modelli tradizionali. Quest'ultimi, infatti, non erano più sufficienti a spiegare l'innovazione in un contesto economico in continuo mutamento e dove il rapporto tra università, industria e governo stava cambiando rapidamente [29].

Un esempio evidente può essere ripreso proprio dall'evoluzione del ruolo delle istituzioni accademiche. In un ambiente caratterizzato da un elevato tasso di specializzazione e da una crescente competizione economica, favorita in larga parte dall'espansione mondiale, la conoscenza diventava un fattore chiave per poter primeggiare. Si è passati da un'idea di università come luogo dove poter effettuare solamente analisi di *trend* economici ad una in cui è necessario il suo coinvolgimento diretto per la creazione di ricchezza [30]. Di conseguenza, il suo operato si è reso sempre più indispensabile nelle attività industriali e, in cooperazione con i centri di ricerca e sviluppo, è diventato il fulcro dell'innovazione e del progresso tecnico-scientifico della società moderna [31]. Tramite la ricerca universitaria di base vengono prodotte nuove conoscenze e idee, mentre attraverso la ricerca applicata si trovano risposte a specifici quesiti e problematiche [32].

In questo contesto ha assunto sempre maggiore rilievo il ruolo dell'autorità governativa come fornitore di risorse e facilitatore del rapporto tra università e industria. Infatti, da un lato, ha la facoltà di predisporre delle strutture che favoriscano la diffusione dell'innovazione e del flusso di conoscenze tra i diversi attori, dall'altro la capacità di fornire risorse finanziarie a sostegno delle migliori iniziative [33]. Proprio un passo successivo della collaborazione tra questi tre diversi elementi è stata la nascita di intermediari che costituissero delle ibridizzazioni delle tre entità [34]. Un esempio potrebbero essere i parchi tecnologici o gli incubatori di imprese. I primi svolgono il ruolo di aggregatori di imprese innovative che puntano allo sviluppo di nuove attività e favoriscono il dialogo tra aziende, università e centri di ricerca. Sono dei veri e propri ecosistemi a metà strada tra ricerca scientifica ed esigenze del mercato che hanno come obiettivo l'accelerazione del processo di introduzione di nuove tecnologie sul mercato. Solo in Italia vi sono oltre una trentina di parchi scientifici specializzati soprattutto nelle tecnologie dell'informazione e della comunicazione [35]. Gli incubatori, invece, differiscono dai centri di ricerca in quanto sono rivolti esclusivamente alle *start-up*, ovvero alle attività imprenditoriali appena nate. Essi infatti hanno l'obiettivo, attraverso servizi di assistenza e consulenza, di diventare il luogo ideale dove un'impresa può nascere e trovare le opportunità per compiere i primi passi del suo ciclo di vita. I primi

incubatori universitari in Italia sono nati negli anni '80, ma sono divenuti subito di grande rilevanza internazionale [36]. Infatti, PoliHub, incubatore del Politecnico di Milano, è nella Top 5 dei migliori incubatori universitari al mondo e I3P, incubatore del Politecnico di Torino, è sul gradino più alto del podio nella lista dei più importanti *Public business incubator*, secondo la classifica stilata dall'associazione svedese UBI Global [37].

Quanto detto conferma l'idea che la collaborazione di attori che possano apportare capacità, risorse e competenze differenti tra loro favorisca la creazione di sinergie nel processo di innovazione. Inoltre, è possibile identificare una relazione positiva tra numero di *partner* coinvolti e impatto nelle attività innovative [38].

La struttura *Triple Helix* ha contribuito, quindi, ad elevare la funzione dell'università nel processo di innovazione tecnologica e sociale e nella transizione da una società di stampo industriale ad una fondata sulla conoscenza. Inoltre, la rete di collaborazione di un'università ha assunto un ruolo fondamentale nel suo rendimento accademico [30]. Tuttavia, oltre al rapporto con industria e governo, anche la collaborazione interuniversitaria è ugualmente vitale per la generazione di nuovo sapere. Eppure su questo aspetto esiste una scarsa letteratura che indagli quali fattori ne favoriscano il successo e la frequenza. Lo scopo del presente elaborato è stato quello di indagare il grado di collaborazione tra università all'interno di un contesto come quello italiano nel periodo di tempo che intercorre tra il 2015 e il 2020, cercando di valutare se tale scelta possa essere impattata dall'ambito di ricerca considerato.

1.5. Materiali e dataset

Seguendo la metodica impiegata nella letteratura per lo studio della collaborazione in ambito internazionale e tra i singoli ricercatori, il presente elaborato cerca di adottare un approccio bibliometrico per analizzare l'aspetto collaborativo che influenza l'attività di ricerca delle università. Come punto di partenza del lavoro sono state considerate le pubblicazioni scientifiche delle università italiane presenti su *Web of Science* (WOS), una delle più grandi e utilizzate banche dati bibliografiche citazionali. WOS è un servizio di indicizzazione online che fornisce dati di riferimento e di citazione da riviste accademiche, atti di conferenze e altri documenti in numerose discipline tramite l'accesso a vari database. Impiegando delle evoluzioni degli indici di citazioni di Eugene Garfield,

che per primo li ha utilizzati nella letteratura accademica, esso permette di raggiungere oltre 170 milioni di *records* tramite quasi 1,9 miliardi di riferimenti bibliografici [39].

Poiché l'attendibilità dell'elaborazioni e delle considerazioni che verranno condotte in seguito, dipendono in larga misura dai dati impiegati, diventa opportuno riportare alcune considerazioni sul grado di copertura di questo strumento bibliografico, inteso come il rapporto tra il numero di pubblicazioni presenti all'interno del database rispetto alla quantità totale relativa ad un determinato settore scientifico. Infatti per ogni archivio citazionale esistono dei criteri che determinano la selezione o l'esclusione di un determinato materiale. Solitamente vengono considerati fattori come la qualità della pubblicazione, la sua natura e la tipologia di citazione. Tuttavia, se per il primo componente diventa rilevante il giudizio di un individuo incaricato nella valutazione del materiale, per i restanti assumono importanza le tipologie di comunicazione peculiari dell'area scientifica considerata. Per quanto riguarda il WOS, la tipologia documentaria prevalente è l'articolo pubblicato su rivista scientifica in lingua inglese, tipologia che costituisce fonte primaria di espressione per le ricerche in discipline ingegneristiche e di scienza della natura. È stato dimostrato da uno studio condotto da Moed nel 2005 [40] come le percentuali di riferimento ad articoli di rivista varino molto da settore a settore, superando abbondantemente il 90% per le scienze naturali, ma facendo fatica ad arrivare al 40% per quelle sociali e umanistiche. Per tale motivo il grado di copertura per i primi sarà molto elevato e superiore rispetto a quello dei secondi, dove vengono preferite altre forme di comunicazione come la monografia o i contributi pubblicati su riviste, giornali o libri di taglio generalista destinate a lettori non accademici. Inoltre, il rapporto di copertura non riguarda solo la tipologia documentaria ma anche le citazioni bibliografiche contenute al suo interno. Infatti, per un documento, maggiore è il numero di volte che viene citato da altri testi, maggiore sarà la sua possibilità di essere presente all'interno del database. Per tale motivo, dato che l'articolo di rivista rappresenta la tipologia prediletta dall'archivio citazionale preso in considerazione, anche la sua presenza all'interno delle fonti bibliografiche sarà preponderante. Questo favorirà ulteriormente quelle materie, come l'ingegneria e le scienze naturali, che prediligono questa forma di pubblicazione a discapito delle materie umanistiche, caratterizzate da forme di pubblicazione che rientrano in misura minore nell'indice bibliografico.

Dalle considerazioni precedentemente esposte, si può affermare come il grado di copertura per le scienze sociali e umane sia notevolmente inferiore e per tale motivo le

considerazioni che verranno esposte nel presente elaborato potrebbero essere affette da tale distorsione e andranno valutate con il beneficio del dubbio.

L'estrazione effettuata da *Web of Science* ha generato un *database* composta da 112 file csv (*comma separated values*) contenenti tutte le pubblicazioni effettuate dagli istituti presi in considerazione che fossero raggiungibili tramite l'archivio citazionale. I dati fanno riferimento ad un totale di 98 istituti italiani tra università statali, private, telematiche e cliniche associate lungo un intervallo temporale che può variare dai quaranta ai sei anni in base all'istituto.

All'interno di ogni file è presente, a corredo di ciascun articolo, una serie molto dettagliata di informazioni che possono essere utili per compiere diverse tipologie di analisi e studi. Tuttavia, ai fini del presente elaborato sono state considerate solamente le seguenti variabili:

- *Accession Number*: codice alfanumerico che identifica univocamente la pubblicazione
- *Research Area*: area di ricerca della pubblicazione che può essere singola o multipla
- *Document Type*: tipologia di documento (*Article, Proceeding Paper, Meeting Abstract, etc....*)
- *Publication Date*: anno di pubblicazione

La prima operazione compiuta è stata quella di restringere il campo di indagine ai soli articoli che, come si può notare dalla Figura 5, rappresentano la tipologia di documento dominante all'interno di questi file. Quasi il 70% delle ricerche qui presenti sono state divulgate tramite articolo su rivista di taglio scientifico.

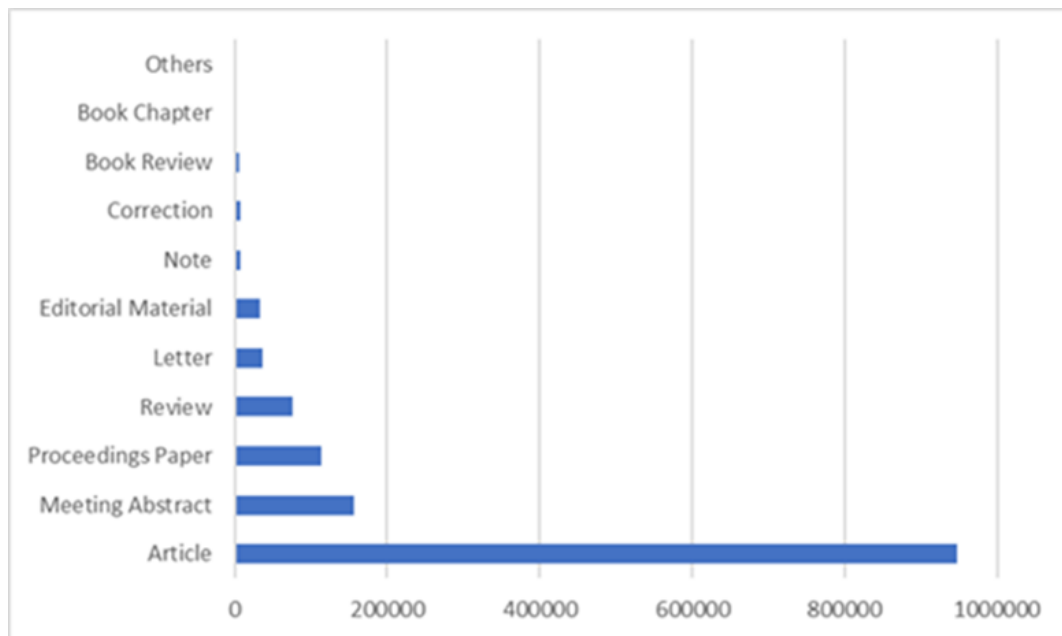


Figura 5 – Numero di occorrenze per tipologia di documento all'interno del database

Tale scelta influirà ulteriormente sulle distorsioni verso le scienze naturali e ingegneristiche di cui si è accennato in precedenza. Tuttavia, viene lasciato a lavori successivi il compito di condurre un'analisi più completa.

Lo scopo del presente elaborato è quello di esplorare il ruolo della specializzazione di un'università in determinati ambiti accademici e come essa possa influire sulla sua attitudine a collaborare. Seguendo tale esigenza, si è deciso di restringere anche l'intervallo di tempo, considerando solamente gli anni compresi tra il 2015 e il 2020. Questo ha permesso di mitigare il grado di cambiamento dei domini di ricerca intrinseco dell'istituzione accademica e di poter valutare in maniera più efficace il suo peso nell'aspetto collaborativo. La specializzazione può evolvere principalmente per due ragioni: la prima è legata al ricambio del personale all'interno dell'università e ad un suo cambio di priorità, la seconda, invece, è collegata alla nascita di nuove teorie e branche. Considerando una finestra temporale ristretta, queste variabilità non dovrebbero influenzare i risultati che sono stati ottenuti. Inoltre, tale scelta è stata guidata da una frammentazione e carenza di dati all'interno del dataset considerato rispetto agli anni di osservazione. Infatti, come viene mostrato dalla Figura 6, solamente negli ultimi 6 anni si ha a disposizione almeno un record per ogni ente considerato.

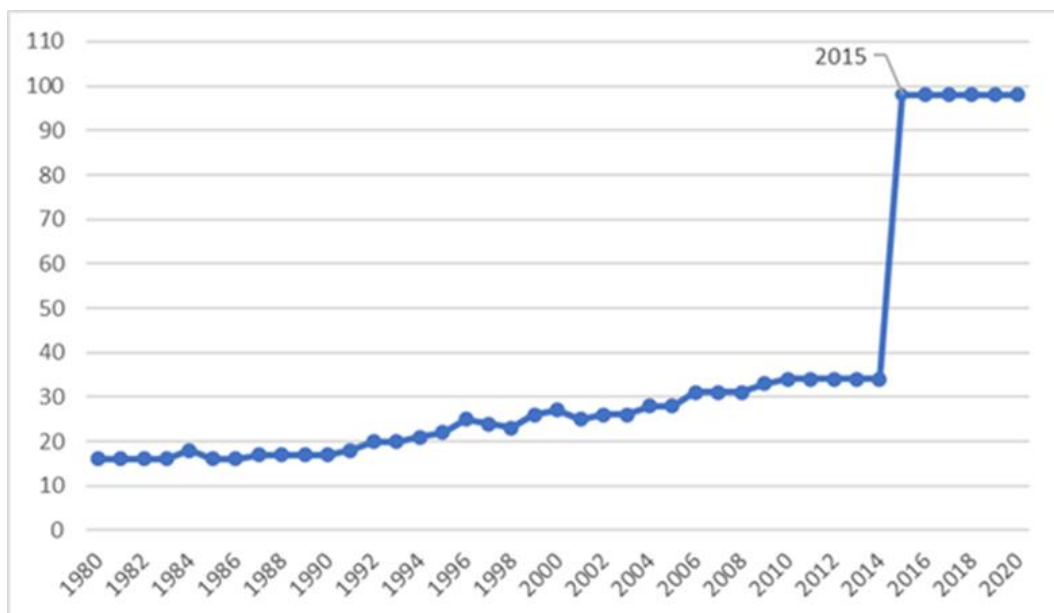


Figura 6 – Numero di istituti accademici con almeno una pubblicazione nell’anno considerato

Tuttavia, anche per l’ultimo anno, il 2020, bisognerebbe prestare un maggior livello di attenzione in quanto è stato influenzato dallo scoppio della pandemia di Covid-19 che, inevitabilmente, ha influenzato il mondo didattico e accademico causando rallentamenti e sospensione in alcuni ambiti di ricerca.

Alla fine di questi lavori di filtraggio, il database, che è stato effettivamente impiegato per le successive analisi, è risultato essere composto da 485.480 osservazioni effettuate nell’intervallo di tempo compreso tra il 2015 e il 2020 e suddiviso su 112 file, ciascuno dei quali associato ad uno dei 98 istituti italiani presi in considerazione.

All’interno dei dati a disposizione è stato possibile trovare documenti riferiti a 254 differenti ambiti di ricerca. Considerando che l’intero archivio WOS ha una copertura di 256 discipline, si può affermare come la ricerca scientifica italiana abbia, seppur in misura difforme, investigato tutti i principali argomenti della scienza, a partire dagli ambiti sociali e umanistici fino ad arrivare a quelli prettamente scientifici e matematici.

Nel proseguo della trattazione, per una maggiore chiarezza di esposizione e di comprensione dei risultati, le differenti aree di ricerca sono state suddivise in 13 macro-categorie in relazione al tema trattato. La ripartizione è stata eseguita seguendo l’esempio presente in un lavoro di Abramo e D’Angelo del 2015 [41], partiti anche loro da un *dataset* estratto da *Web of Science*. In Tabella 1 vengono riportati i dati di questa distribuzione, mentre si rimanda il lettore alla visione dell’Appendice, dove vengono elencate tutte le aree di ricerca considerate per ogni singola categoria.

Category	N° aree di ricerca	N° università	N° articoli
Physics	18	92	51981
Clinical Medicine	43	97	81376
Biomedical Research	14	94	42853
Biology	31	97	56343
Chemistry	8	91	31286
Engineering	42	97	70455
Earth and Space Sciences	13	95	27702
Economics	7	91	12276
Multidisciplinary Sciences	2	88	1560
Law, Political and Social Sciences	33	96	14170
Psychology	11	91	7688
Mathematics	6	88	18868
Arts and Humanities	26	88	8478
tot	254	98	321702

Tabella 1 – Informazioni riguardanti la produzione scientifica italiana tramite la suddivisione in macro categorie (2015-2020)

Dalla tabella si può notare come la somma delle righe relative al numero di articoli è superiore al numero totale riportato in basso. Questa distorsione è dovuta all'approccio *double count* utilizzato per la suddivisione in macro categorie, con cui ogni articolo che abbia trattato argomenti inerenti a categorie differenti è stato contato più volte. Un metodo identico è stato applicato anche per quanto riguarda la classificazione nelle singole aree di ricerca. Per quanto riguarda la colonna relativa al numero di università, è stato riportato per ogni macro-categoria il conteggio degli istituti che nel periodo preso in considerazione hanno pubblicato almeno un articolo di ricerca relativo ad esso.

Dopo aver esposto e descritto brevemente i dati a disposizione e l'obiettivo del presente elaborato, nel prossimo capitolo verranno esposti i principali risultati ottenuti tramite una serie di operazioni e confronti. In particolare verrà data evidenza ai comportamenti assunti dal Politecnico di Torino, di Milano e di Bari. Infine, nel terzo e ultimo capitolo verrà riportata una breve analisi di regressione compiuta per verificare sommariamente se le conclusioni tratte in precedenza possano trovare validità statistica.

Capitolo II

2.1. Ambiti di ricerca in Italia

La prima analisi compiuta è stata quella di valutare quali siano gli ambiti di studio di maggior interesse per la ricerca italiana. Per raggiungere tale scopo tutte le pubblicazioni a disposizione sono state raggruppate attraverso l'*Accession Number*, il codice alfanumerico univoco, e sono state conteggiate per ogni area di ricerca più volte nel caso di articoli multi settoriali. Data l'elevata quantità di campi scientifici presenti nel *database* e per ragioni di chiarezza espositiva, in Figura 7 è stata data evidenza solamente alle trenta aree di ricerca che presentavano il maggior numero di pubblicazioni.

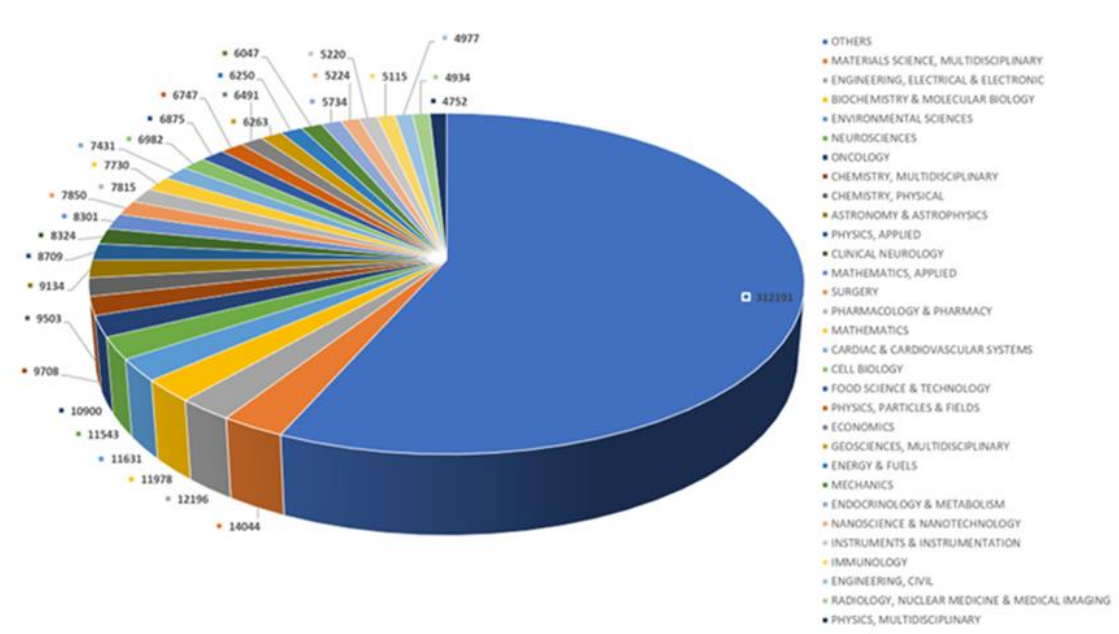


Figura 7 – Distribuzione del numero di articoli per le 30 aree di ricerca con più pubblicazioni tra il 2015 e 2020

Come si può notare dal grafico, su 254 differenti aree di ricerca che è stato possibile identificare all'interno del *database*, le trenta più analizzate coprono circa il 45% delle pubblicazioni totali effettuate da istituti italiani negli anni compresi tra il 2015 e il 2020, mentre le prime dieci arrivano a sfiorare il 20% di copertura. È possibile affermare, quindi, che circa la metà degli articoli pubblicati si concentrano in meno di un ottavo degli ambiti che sono stati investigati. In questa speciale classifica le materie scientifiche, che spaziano dalla matematica pura alla biochimica e biologia molecolare, dall'astronomia e astrofisica all'ingegneria elettrica ed elettronica, rappresentano i leader indiscussi. Per poter individuare un argomento di carattere sociale bisogna scendere fino alla 77-esima

posizione, dove trova posto la psicologia con appena 2.514 pubblicazioni. Questa notevole distorsione verso tematiche scientifiche e tecnologiche è sicuramente da imputare alla tipologia di documento analizzata, l'articolo scientifico, e alle caratteristiche dell'archivio citazionale dal quale sono stati estratti i dati a disposizione, come riportato nel primo capitolo. Tuttavia non è da escludere che vi sia una reale preferenza da parte degli istituti italiani in quanto l'articolo scientifico rappresenta la forma più popolare di pubblicazione in ambito scientifico.

Subito dopo ci si è concentrati sull'aspetto collaborativo, ovvero su quelle ricerche che per la loro stesura si è assistito ad una collaborazione tra due o più istituti italiani. Conducendo un'analisi analoga a quella precedente, sono stati ricercati quegli ambiti che presentavano il maggior numero di articoli "collaborativi". Gli esiti sono riportati in Figura 8.

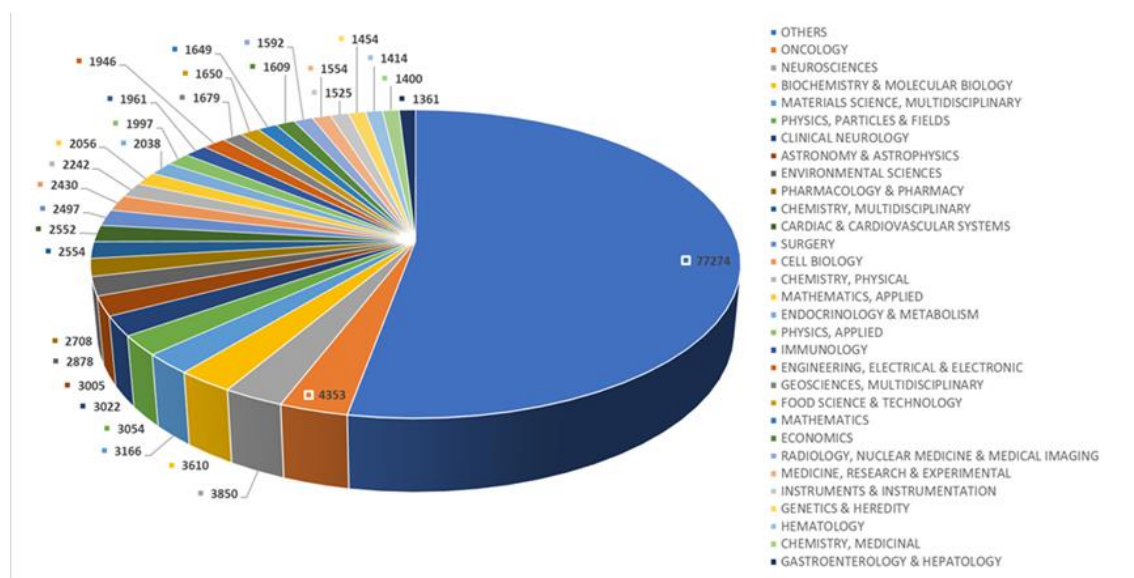


Figura 8 – Distribuzione del numero di articoli cooperativi per le 30 aree di ricerca più collaborative (2015-2020)

In conformità a quanto ricavato in precedenza, quasi il 50% delle pubblicazioni che hanno visto il coinvolgimento di due o più università è concentrato in sole 30 tematiche e tra di esse si possono trovare soltanto materie scientifiche o tecnologiche. Tuttavia questi risultati sono fortemente influenzati dal numero totale di documenti per un dato ambito, più articoli si hanno a disposizione, maggiore sarà la probabilità di trovarne qualcuno collaborativo. Infatti, se al posto del numero assoluto viene considerato il rapporto rispetto alla quantità totale di pubblicazioni, la situazione cambia e non di poco. In Tabella 2 vengono riportare le prime 10 aree scientifiche per la percentuale di articoli collaborativi

su quelli totali e per ciascuna di essa viene riportata la posizione in classifica che avrebbe se si considerasse solamente il numero assoluto di articoli collaborativi.

Rank	Area di ricerca	N° articoli collaborativi	N° articoli totali	% articoli collaborativi
35	PHYSICS, NUCLEAR	1280	2578	49,65%
129	ALLERGY	278	612	45,42%
5	PHYSICS, PARTICLES & FIELDS	3054	6747	45,26%
203	ANDROLOGY	70	161	43,48%
1	ONCOLOGY	4353	10900	39,94%
42	UROLOGY & NEPHROLOGY	1056	2730	38,68%
167	MEDICINE, GENERAL & INTERNAL	143	372	38,44%
18	IMMUNOLOGY	1961	5115	38,34%
46	RESPIRATORY SYSTEM	934	2479	37,68%
28	HEMATOLOGY	1414	3766	37,55%
tot		87601	321702	27,23%

Tabella 2 – Prime dieci aree di ricerca per rapporto articoli collaborativi su articoli totali

La tabella mostra come cambiando la prospettiva dalla quale si vuole analizzare un contesto si possono ottenere dei risultati molto discordanti tra di loro. Infatti, come accennato in precedenza, è possibile notare una classifica capovolta rispetto a quella precedente. Soltanto “*Oncology*” e “*Physics, particles & fields*” riescono a mantenere una posizione di vertice.

Un controllo più approfondito mostra come tra i primi dieci classificati sia possibile trovare solamente argomenti inerenti l’ambito medico e fisico. In particolare “*Oncology*”, oltre ad essere una delle aree con più pubblicazioni in assoluto, risulta essere anche una di quelle che necessita e coinvolge il maggior numero di istituti per le sue ricerche. Questo sottolinea come vi sia un’elevata attenzione da parte delle università italiane e quindi della ricerca scientifica italiana in generale verso la salute dell’uomo. Inoltre, considerando alcune caratteristiche intrinseche al settore della medicina e della fisica, è possibile trarre alcune riflessioni sull’aspetto della collaborazione. Per il primo, data l’importanza dell’argomento trattato, si richiede di raggiungere risultati validi nel minor tempo possibile, mentre per il secondo è la complessità della materia che richiede una moltitudine di conoscenze specifiche. Di conseguenza, se per raggiungere tali obiettivi, il mezzo più utilizzato dagli istituti italiani è stato quello della cooperazione, allora si può concludere come la collaborazione, tramite la condivisione di esperienze, conoscenze ed investimenti, ricopra un ruolo fondamentale in questo processo.

2.2. Ricerca scientifica negli istituti italiani

Dopo aver esaminato a livello generale i dati presenti all'interno del *database*, facendo emergere caratteristiche e peculiarità della produzione scientifica in Italia, con particolare attenzione agli ambiti di divulgazione, si è passati ad uno studio più specifico e meticoloso. Si è cercato di analizzare il comportamento di ogni singolo istituto e come esso possa influenzare o essere influenzato dal contesto complessivo.

Il primo passo è stato quello di individuare gli istituti che con la loro attività hanno contribuito maggiormente alla ricerca italiana in termini di produzione di articoli scientifici. In Figura 9 è stato riportato il numero di pubblicazioni per tutti i 98 enti dando evidenza soltanto di quelli più produttivi e dei tre Politecnici.

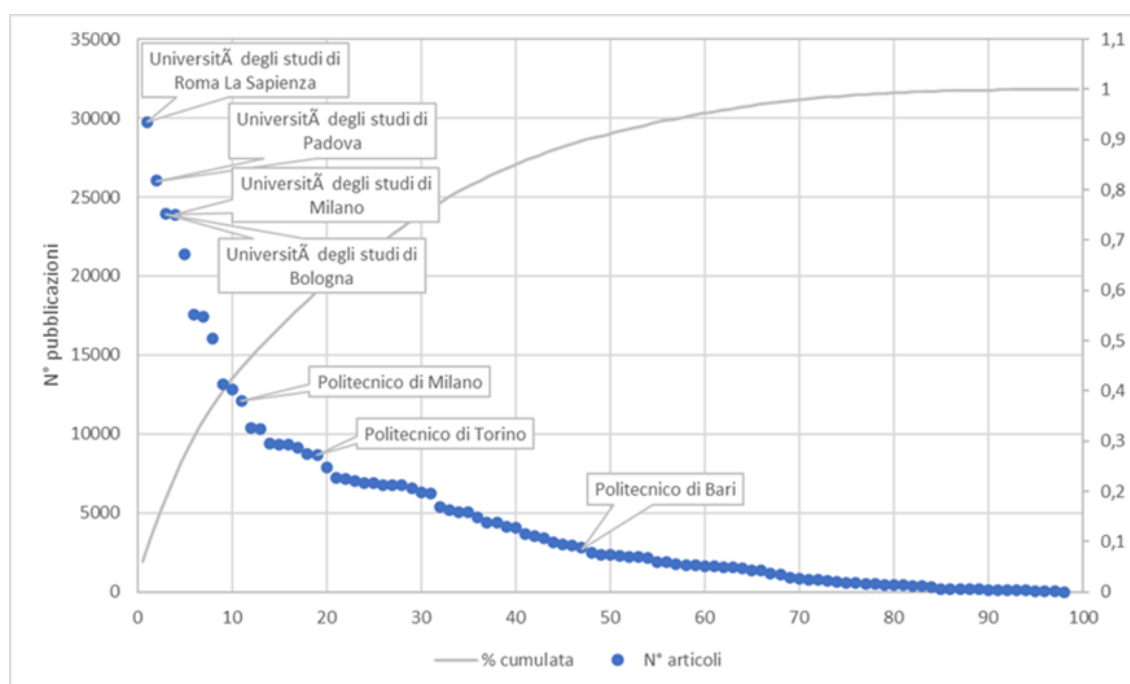


Figura 9 – Distribuzione del numero di articoli scientifici pubblicati per ciascun istituto analizzato (2015-2020)

Tra le università più prolifiche, sul podio si hanno alcune delle più longeve in Italia: al primo posto si trova l'Università La Sapienza di Roma con 29.740 articoli scientifici, seguita dall'Università degli studi di Padova con 26.036, mentre sul terzo gradino del podio si trovano quasi a pari merito l'Università degli studi di Bologna e di Milano. Più staccati invece risultano il Politecnico di Milano, di Torino e di Bari con rispettivamente 12.065, 8.664 e 2.814 articoli pubblicati. Sempre dal grafico, considerando la linea cumulata, si può notare come le prime 10 università riescono a coprire oltre il 40% della

produzione scientifica italiana. Questo valore testimonia la presenza di poche università “*leader*” nel panorama scientifico italiano.

Lo step successivo è stato lo studio della composizione di questi portafogli di articoli, di come le università tendano a distribuire i propri sforzi nei diversi ambiti scientifici. È stato valutato se vi fosse una propensione da parte degli istituti a focalizzarsi su un numero ristretto di argomenti oppure se vi fosse una ripartizione omogenea sulle diverse materie trattate. Questa indagine, in particolare, sarà utile per verificare se le aree di ricerca che sono state ricavate in precedenza con il maggior numero di articoli, siano anche quelle che hanno visto impegnati più atenei differenti. Sarà quindi valutata come la presenza di grandi università, leader nelle pubblicazioni, possa aver alterato la media dei risultati che sono stati ottenuti. L’esempio che segue potrebbe chiarire lo scopo di tale indagine: se 10 università pubblicano ciascuna 10 articoli in fisica nucleare, mentre l’11-esima ne pubblica 200 in filosofia, compiendo un’analisi in aggregato si avrà da parte degli istituti una preferenza verso la filosofia, quando nella realtà non vi è stata. Il risultato, infatti, è stato fortemente deviato dalla presenza dell’11-esima università che, comportandosi da *outliers*, ne ha alterato gli esiti.

In Figura 10 viene mostrato graficamente il legame sussistente tra il numero di articoli pubblicati e il numero di ambiti scientifici investigati considerando le informazioni a disposizione per ciascun istituto. Com’era lecito aspettarsi, esiste una chiara dipendenza tra questi due fattori. Le università, che nel periodo compreso tra il 2015 e 2020 hanno prodotto il maggior numero di articoli scientifici, sono anche quelle che hanno investigato più aree tematiche. Infatti maggiore è il volume di produzione scientifica, maggiore sarà la probabilità di trovare materiali che si collegano ad argomenti differenti. Questa relazione trova un riscontro anche dal punto di vista pratico. Gli istituti più rinomati e importanti sono anche quelli che dispongono di ingenti fondi e che possono contare sulla presenza di un maggior numero di scienziati e di ricercatori. Per tale motivo avranno maggiori opportunità rispetto ad un piccolo ateneo di investire in ambiti nuovi e lontani dalle loro competenze chiave.

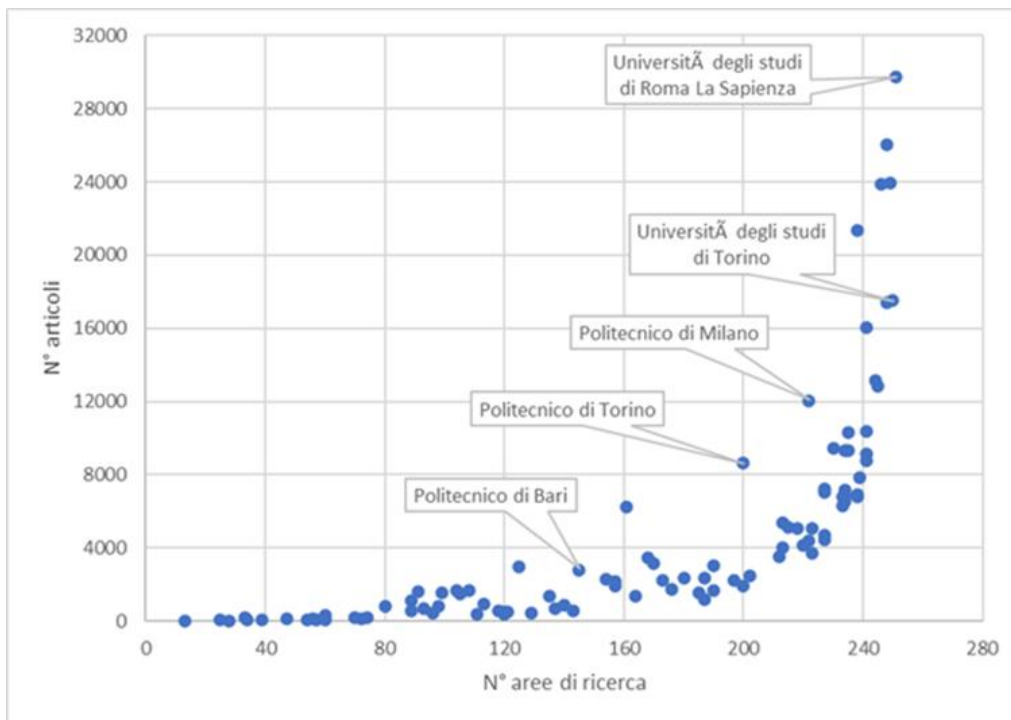


Figura 10 – Distribuzione del numero di aree di ricerca trattate rispetto al numero di articoli scientifici pubblicati per ciascuna università (2015-2020)

Da questo grafico, tuttavia, non è possibile intuire come ciascuna università abbia distribuito le varie pubblicazioni all'interno dei differenti ambiti scientifici. Si può solamente sapere quanti argomenti sono stati trattati, ma non con quale intensità, se sia stato dedicato uno studio occasionale o se sia stato investigato approfonditamente. Per queste e altre ragioni, è stato necessario compiere un ulteriore step in tale direzione.

Il problema precedentemente esposto ricade nell'ambito di studio della concentrazione e diversità presente all'interno di un dato oggetto o insieme. Per compiere tale analisi è stato preso in prestito un indice molto impiegato in campo economico: l'indice di Herfindahl-Hirschman o HHI. Esso è una misura della dimensione delle imprese in relazione all'industria in cui operano e rappresenta un indicatore della concentrazione del potere economico. Essendo utile per misurare il livello di competizione presente all'interno di un mercato è stato ampiamente adoperato a livello economico e in particolar modo dalle autorità *antitrust* che l'hanno utilizzato per valutare e comprendere come le operazioni di acquisizioni e/o fusioni, in inglese *Mergers and Acquisitions* o M&A, potessero impattare il loro mercato di riferimento [42]. L'indice di Herfindahl-Hirschman è calcolato elevando al quadrato la quota di mercato di ciascuna impresa che compete in un determinato settore e sommando i valori così ottenuti. Il suo valore è compreso tra 0 e 1. Più un mercato si avvicina ad essere un monopolio, più alta è la concentrazione al suo

interno e di conseguenza più elevato è il valore che assume l'indice HHI. Se, per esempio, fosse presente in un settore una sola impresa, essa avrebbe il 100% delle quote di mercato e l'indice HHI sarebbe uguale a 1. Se ci fossero migliaia di imprese, ognuna con una quota identica di mercato, allora l'indice HHI sarebbe molto prossimo a 0 ed indicherebbe una condizione di concorrenza perfetta. In generale valori inferiori a 0,15 rappresentano situazioni a bassa concentrazione, tra 0,15 e 0,25 situazioni a media concentrazione e, infine, superiori a 0,25 situazioni ad elevata concentrazione. La fortuna di questo indice, che ne ha determinato il successo, è stata la semplicità di calcolo e la ridotta quantità di dati, per altro di facile reperimento, necessari per il suo calcolo [43].

Nel presente elaborato è stata riprodotta la formulazione dell'indice di Herfindahl-Hirschman, adattata ai dati a disposizione e utilizzata per calcolare il livello di concentrazione presente all'interno dei portafogli di pubblicazioni di ciascuna delle 98 università analizzate. Di conseguenza l'indice di concentrazione, indicato per semplicità con *HHI*, associato al singolo istituto *i* è stato calcolato come la sommatoria dei quadrati del rapporto tra il numero di pubblicazioni nell'area di ricerca *a* e il numero totale di pubblicazioni:

$$HHI_i = \sum_a \left(\frac{n^{\circ}pubblicazioni_{a,i}}{\sum_a n^{\circ}pubblicazioni_{a,i}} \right)^2$$

In Figura 11 viene mostrata la distribuzione dell'indice appena misurato, qualora fosse approssimabile ad una distribuzione normale.

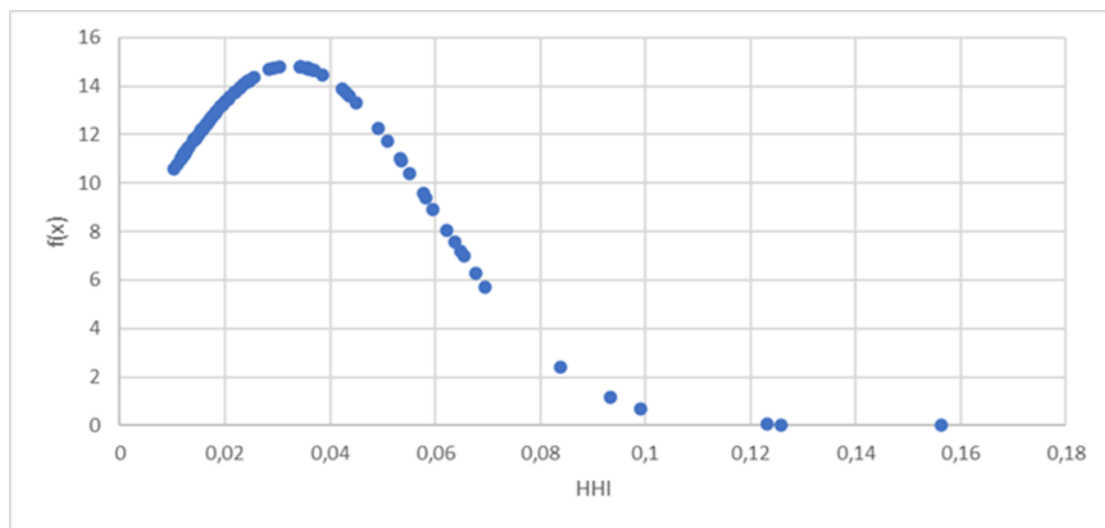


Figura 11 – Rappresentazione della distribuzione dell'indice di concentrazione HHI

I dati ottenuti permettono di affermare come le università italiane, durante la loro attività di ricerca, tendono a spaziare su differenti ambiti, non focalizzando gli sforzi su due o tre argomenti. Seppur esistano dei casi che presentano una concentrazione decisamente maggiore rispetto alla media, arrivando a sfiorare valori di 0,16, si rimane comunque lontani da quel valore di 0,25 che rappresenta una sorta di soglia per situazioni ad elevata concentrazione.

A questo punto, in Figura 12 e Figura 13, l'indice *HHI* viene relazionato rispettivamente al numero totale di articoli pubblicati e al numero di aree di ricerca trattate.

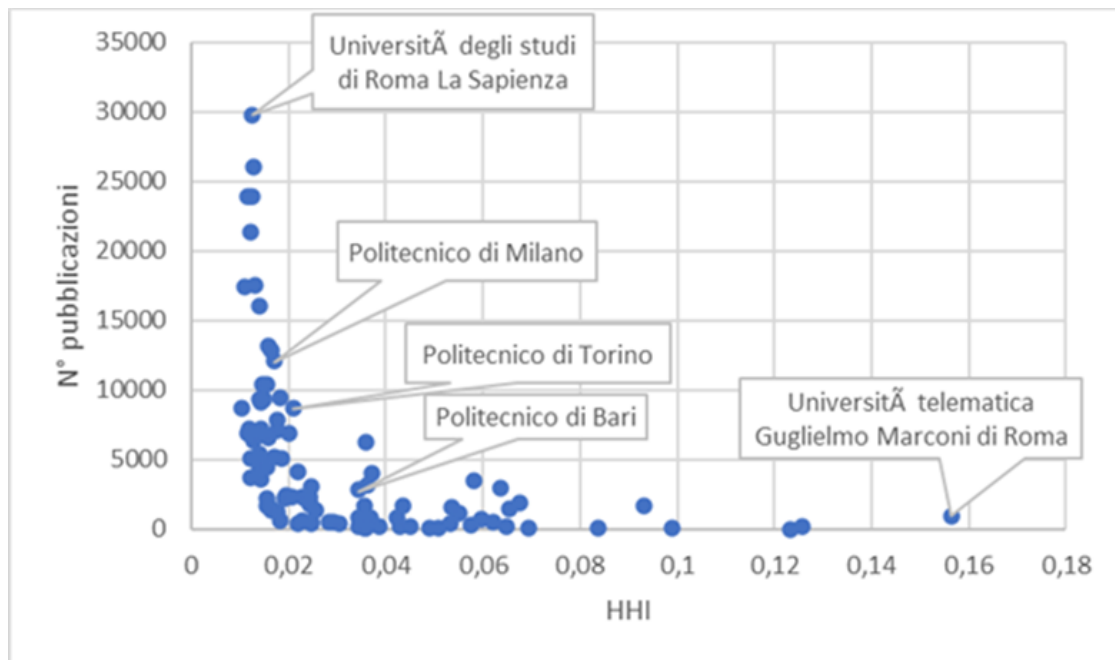


Figura 12 – Distribuzione indice di concentrazione HHI rispetto al numero di pubblicazioni totali per ciascuna università (2015-2020)

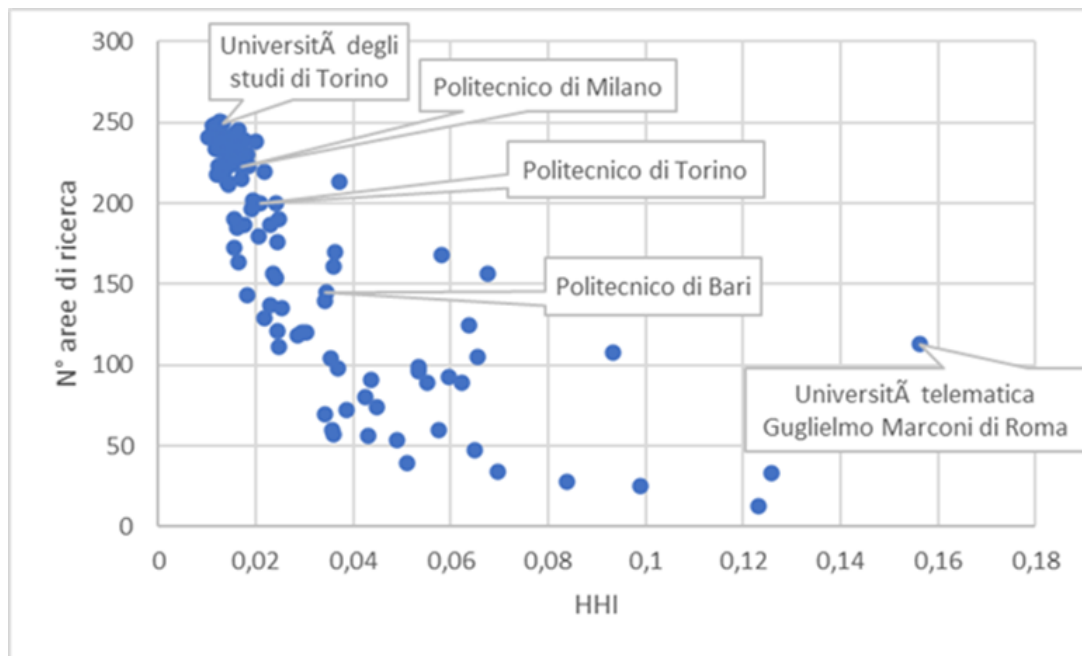


Figura 13 – Distribuzione indice di concentrazione HHI rispetto al numero di aree di ricerca per ciascuna università (2015-2020)

Questi due grafici mostrano in maniera inequivocabile come all’aumentare del numero di pubblicazioni o del numero di aree di ricerca, l’indice di concentrazione diminuisca. Vi sussiste quindi una chiara relazione inversamente proporzionale tra i primi due fattori e il terzo. Tale risultato permette di affermare come i più grandi istituti, ovvero coloro il cui lavoro potrebbe alterare gli esiti fino a questo punto ottenuti, tendano a distribuire in maniera equa l’impegno che offrono alle varie discipline.

La presenza di alcuni *outliers*, tra cui l’Università telematica Guglielmo Marconi di Roma che mostra il più alto livello di concentrazione nonostante un discreto numero di argomenti trattati, non possono rappresentare un elemento di disturbo per le analisi che sono state condotte. Si può, quindi, ribadire, con un certo grado di sicurezza, che le aree di ricerca, identificate per aver ottenuto il maggior numero di pubblicazioni a riguardo, siano anche quelle che hanno beneficiato di un impegno consistente da parte di numerose università italiane.

Definita per ogni istituto la composizione e la distribuzione delle loro pubblicazioni, si è passati ad analizzare il fattore relativo alla cooperazione. Seguendo un approccio simile a quello utilizzato per le singole aree di ricerca, l’aspetto collaborativo è stato studiato prima in termini assoluti e successivamente in termini relativi. I risultati sono stati riassunti nelle due seguenti figure.

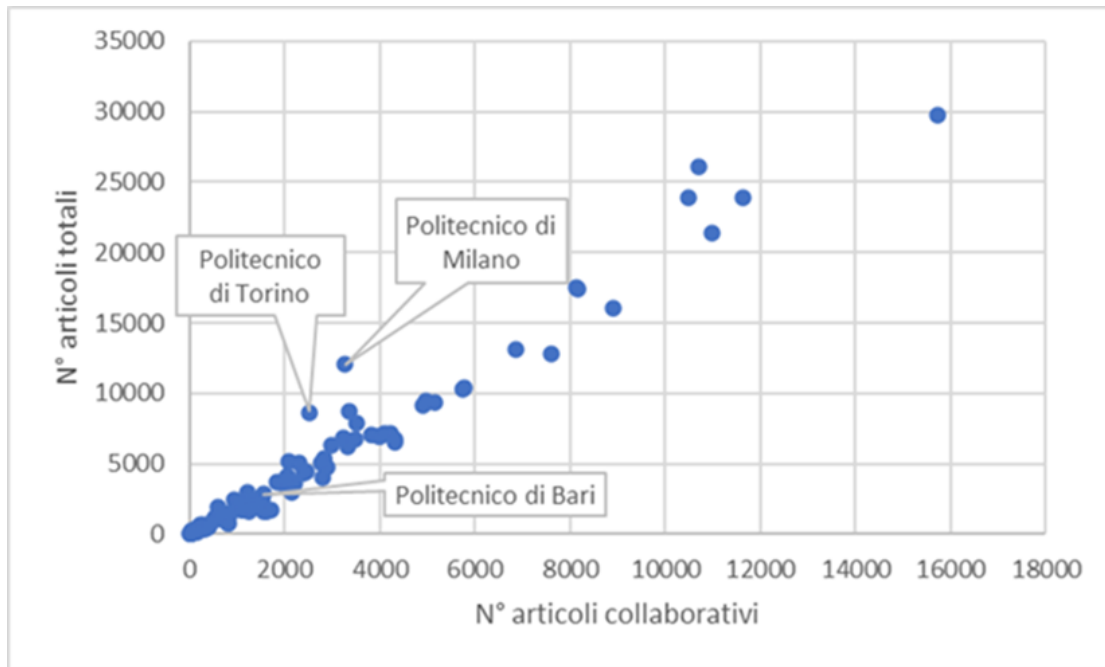


Figura 14 – Distribuzione numero di articoli collaborativi rispetto al numero di pubblicazioni totali per ciascuna università (2015-2020)

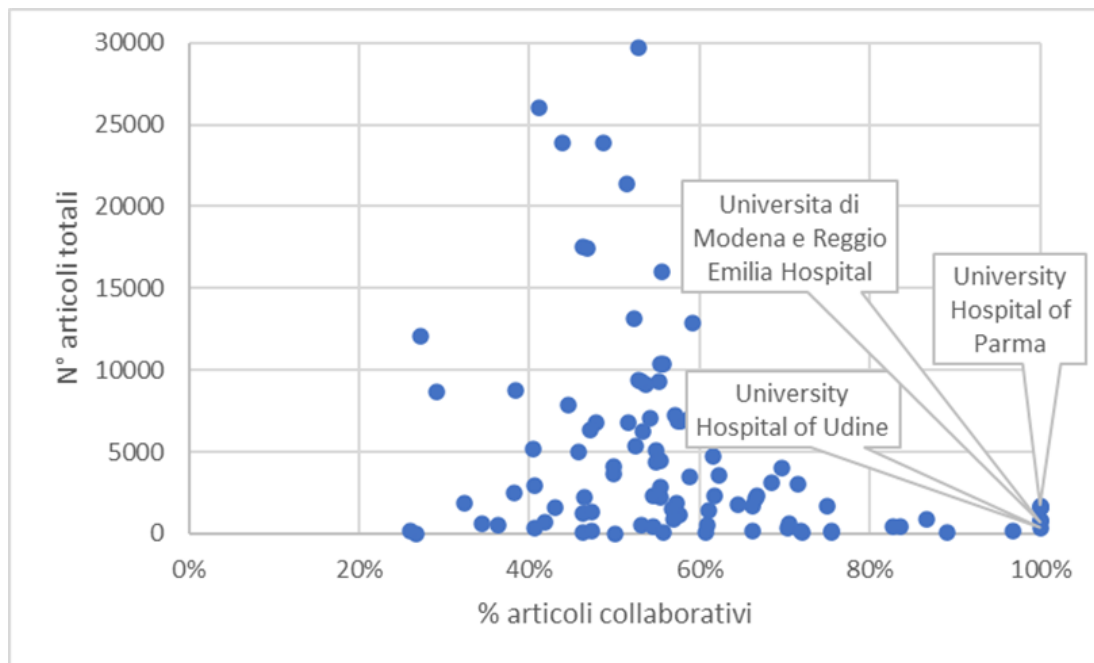


Figura 15 – Distribuzione % articoli collaborativi rispetto al numero di pubblicazioni totali per ciascuna università (2015-2020)

La Figura 14 mostra come vi sia una perfetta relazione lineare tra il numero di articoli collaborati e il numero di pubblicazioni totali. I punti, che rappresentano ciascuna delle 98 università analizzate, sembrano poter disegnare una retta nel piano. Solamente il politecnico di Torino e di Milano si discostano leggermente da questo orientamento

mostrando una minore attitudine verso la cooperazione, mentre il Politecnico di Bari rimane in linea con gli altri istituti.

Differente, invece, è la situazione rappresentata in Figura 15, dove si trova una maggiore varietà di risultati. Tramite uno studio per valori relativi è stato possibile individuare facilmente degli istituti che hanno avuto tra il 2015 e il 2020 una percentuale di collaborazione pari all'unità, ovvero dove ogni ricerca è stata condotta con il supporto di almeno un altro ente. Su 6 università identificate, 3 sono evidenziate in figura. Una caratteristica comune tra tutti e sei gli enti è che sono delle aziende ospedaliere universitarie, ovvero presidi ospedalieri nei quali vengono svolti corsi di laurea in medicina e chirurgia. Questo risultato combacia perfettamente con quanto segnalato in precedenza. Infatti, le aree di ricerca legate all'ambito medico sono quelle che presentano i più alti tassi di collaborazione.

Sempre in Figura 15 si può notare una forte concentrazione tra il 40% e il 60% di articoli collaborativi. Infatti, se si calcola la percentuale di collaborazione media tra tutti gli enti si trova un valore di poco superiore al 50%. Esso implica che per le università italiane una ricerca su due è condotta con la partecipazione almeno di un altro ente. Considerando che la percentuale tra articoli collaborativi e il numero totale di articoli pubblicati dalla ricerca scientifica italiana è prossima al 28%, molte delle cooperazioni avvengono tra tre o più istituti.

Conclusa la prima parte di analisi, che ha estrapolato le principali caratteristiche e inclinazioni della ricerca scientifica italiana, nel prosieguo del presente capitolo si è deciso di considerare una nuova variabile, quella temporale, per valutare come tali aspetti siano evoluti nell'intervallo di tempo in esame. Inoltre, sono state costruite delle metriche per poter confrontare in maniera più oggettiva e strutturata la sovra specializzazione di un ente in un ambito, la sua tendenza a collaborare e, qualora esistesse, far emergere una loro correlazione. Tali metriche saranno anche di grande rilevanza per la costruzione del modello di regressione che verrà presentato nel terzo capitolo.

2.3. Indice di specializzazione

Poiché molte delle aree di ricerca trattano temi peculiari tra di loro e per evitare di dover considerare molte combinazioni con un numero troppo modesto di osservazioni che, in

particolar modo per l'aspetto della collaborazione, potrebbero alterare i risultati come verrà mostrato in seguito, si è deciso di effettuare un raggruppamento in macro-categorie secondo il modello proposto in un lavoro del 2015 da Abramo e D'Angelo [41] e che è possibile visionare in Appendice. Inoltre, tale suddivisione permetterà di facilitare e di rendere maggiormente comprensibili l'esposizione dei risultati che sono stati ottenuti.

Completata questa operazione e suddivise le pubblicazioni di ciascun istituto tra le macro-categorie di appartenenza e gli anni di pubblicazione, si è passati alla costruzione di due indicatori tramite i quali poter assegnare un determinato livello di specializzazione e di collaborazione a ciascuna università. Per il primo elemento l'indice che è stato strutturato ha permesso di indicare se un'università possa essere definita maggiormente specializzata in una materia rispetto agli altri enti. Per il secondo, invece, la metrica ha permesso di individuare in che ambiti gli istituti tendono a prediligere un supporto esterno e, in un secondo momento, è stata sfruttata per valutare come questo aspetto possa influenzare o essere influenzato dal grado di specializzazione dell'università su di un argomento.

Per il calcolo della specializzazione è stato preso come riferimento un indice molto impiegato nella letteratura [44, 45] e che spesso viene indicato con la denominazione *Relative Scientific Advantage*, o abbreviato *RSA*. L'*RSA* rappresenta il grado di specializzazione di un'università in una determinata categoria. Il suo valore viene calcolato come la percentuale del numero di pubblicazioni nella categoria c per la singola università i , diviso per la media delle percentuali nella medesima categoria di tutte le università considerate:

$$RSA_{c,i} = \frac{n^{\circ}pubblicazioni_{c,i} / \sum_c n^{\circ}pubblicazioni_{c,i}}{media_i \left(n^{\circ}pubblicazioni_{c,i} / \sum_c n^{\circ}pubblicazioni_{c,i} \right)}$$

Nel presente lavoro l'indice è stato quantificato tenendo in considerazione anche l'anno di pubblicazione dell'articolo. Questo ulteriore livello di dettaglio è stato inserito per valutare anche come l'indice possa evolvere nel tempo, ovvero se le università nel corso di un numero ristretto di anni tendano a cambiare ambito di specializzazione o se in assenza di eventi esterni particolari preferiscano dare continuità al loro lavoro. La formula precedente è stata, quindi, aggiornata inserendo anche la variabile temporale t come mostrato di seguito:

$$RSA_{c,i,t} = \frac{n^{\circ}pubblicazioni_{c,i,t} / \sum_c n^{\circ}pubblicazioni_{c,i,t}}{media_i \left(n^{\circ}pubblicazioni_{c,i,t} / \sum_c n^{\circ}pubblicazioni_{c,i,t} \right)}$$

Quando $RSA_{c,i,t}$ è uguale 1, l'università i ha una percentuale di pubblicazione nella categoria c e nell'anno t che è uguale alla percentuale media di tutte le altre università per quell'anno. Per tale motivo, segue che $RSA_{c,i,t} = 1$ rappresenta una soglia di specializzazione: quando $RSA_{c,i,t} > 1$, l'università i è considerata specializzata nella categoria c e nell'anno t .

L'indice così costruito ha permesso di analizzare come gli articoli, presenti all'interno del portafoglio di pubblicazioni di ciascun istituto, fossero distribuiti tra le varie tematiche scientifiche e di confrontare tali ripartizioni tra di loro. Inoltre ha reso possibile il paragone tra istituti di dimensioni e caratteristiche notevolmente differenti.

Nelle seguenti figure vengono presentati alcuni dei valori ottenuti. Nel primo gruppo di immagini si è scelto di riportare l'evoluzione dell'indice RSA avvenuta durante l'intervallo di tempo analizzato per quanto riguarda la macro categoria “*Engineering*”. Nel secondo gruppo, invece, sono stati esposti i risultati ottenuti per alcune macro categorie in un anno di riferimento: il 2019. In entrambi i casi sono stati evidenziati in particolar modo i valori associati al Politecnico di Torino, di Milano e di Bari.



Figura 16 – Distribuzione indice di specializzazione RSA rispetto al numero di articoli pubblicati per la categoria “Engineering” per ogni anno compreso tra il 2015 e il 2020

A conferma di quanto si potesse prevedere, per la categoria “Engineering” i tre Politecnici dominano la classifica sia per quanto riguarda il grado di specializzazione che il numero di articoli. Tuttavia, come era stato preventivato, non vi è una relazione tra numero di pubblicazioni in quell’ambito e indice di specializzazione. A parità di ricerche pubblicate è possibile individuare una notevole varietà nel tasso di specializzazione in relazione all’istituto considerato. Infine, confrontando i risultati ottenuti nel corso degli anni, si può notare una certa regolarità nella dispersione dei punti, a dimostrazione del fatto che in un periodo relativamente breve, come possono essere 6 anni, le università italiane tendono a dare una certa continuità al loro lavoro.

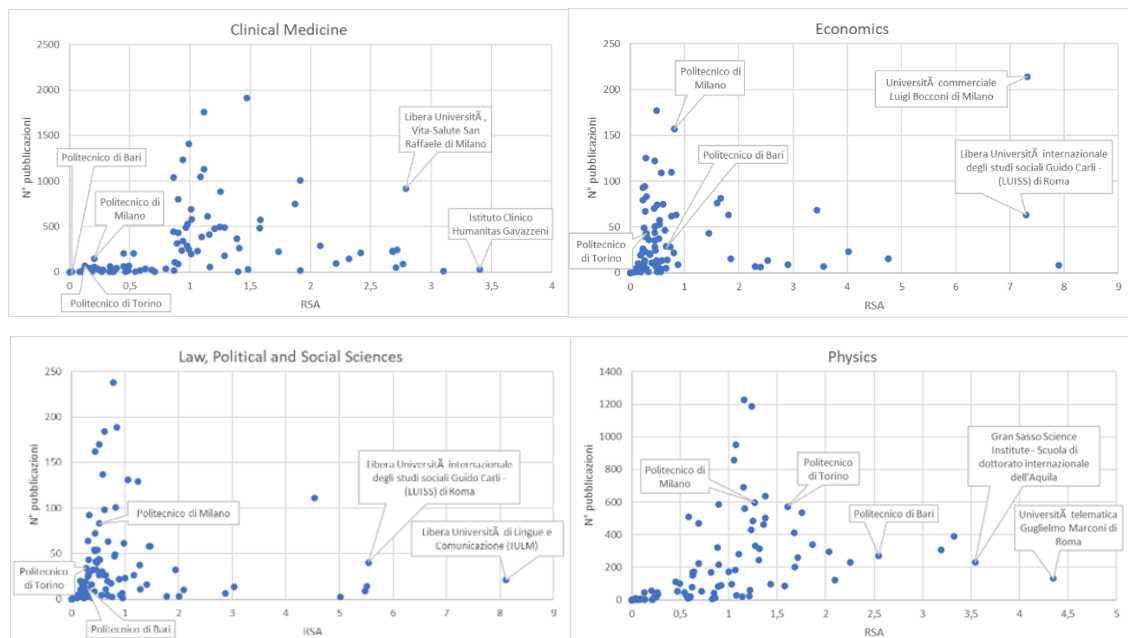


Figura 17 – Distribuzione indice di specializzazione RSA rispetto al numero di articoli pubblicati nell'anno 2019 per le categorie: “Clinical Medicine”, “Economics”, “Law, Political and Social Sciences” e “Physics”

Considerando invece la Figura 17 è possibile osservare per ciascuna macro categoria una differente distribuzione di punti, indicatore dell'eterogeneità presente all'interno del panorama scientifico italiano. I tre Politecnici mostrano un comportamento simile nei vari ambiti e coerente con le aspettative: bassa specializzazione per quanto riguarda gli ambiti inerenti al campo medico, economico e sociale, mentre una discreta competenza e attenzione verso la fisica, in particolar modo per il Politecnico di Bari. Vi è congruenza anche con i risultati ottenuti per gli altri istituti. Per esempio, l'Università Bocconi di Milano, famosa per i percorsi di laurea in economia che offre, oltre a detenere nella categoria “*Economics*” il secondo più alto valore di specializzazione, è anche quella che ha effettuato il maggior numero di pubblicazioni. In ambito di medicina clinica, invece, gli istituti ospedalieri sono quelli che presentano i più alti gradi di specializzazione.

Prima di proseguire con la trattazione, bisogna aggiungere un'ulteriore considerazione. A valori estremamente elevati dell'indice *RSA* non corrisponde un altrettanto livello di competenza e padronanza della materia. Per come è strutturato l'indice, le università con un modesto volume di pubblicazioni hanno un maggiore livello di concentrazione delle pubblicazioni e, di conseguenza, maggiori probabilità di ottenere valori di *RSA* elevati. Per tale motivo affermare che un ente sia maggiormente specializzato rispetto ad un altro su un certo argomento perché ha un valore dell'indice *RSA* più elevato non è corretto.

Infatti, tra livelli simili dell'indice è più plausibile che l'istituto con più pubblicazioni abbia una specializzazione e una conoscenza più approfondita della materia.

Inoltre, la presenza di alcuni *outliers* potrebbe alterare leggermente le analisi compiute in quanto renderebbe più difficile per un'università risultare specializzata in un certo ambito. Tuttavia, dato il loro numero esiguo e il livello di dettaglio condotto in questa sezione del lavoro, si è deciso di continuare a considerarli e di prestare maggiore attenzione alle conclusioni a cui si è giunti.

Il passo successivo è stata la realizzazione di una serie di “matrici di specializzazione”, MS_t , una per ogni anno di pubblicazione compreso tra il 2015 e il 2020. Ogni matrice è stata costruita in maniera tale che le righe riportassero le 98 università italiane in esame, le colonne rappresentassero le macro categorie che sono state introdotte e il generico elemento (i,c) fosse uguale a 1 se l'università i fosse stata specializzata nella categoria c al tempo t :

$$MS_t(i, c) = \begin{cases} 1 & \text{se } RSA_{i,c,t} > 1 \\ 0 & \text{se } RSA_{i,c,t} \leq 1 \end{cases}$$

L'impiego di queste matrici, oltre a facilitare alcune operazioni di analisi e di esposizione dei risultati, è stato di grande utilità per il confronto tra specializzazione e collaborazione che verrà esposto in seguito.

I dati presenti nelle matrici sono stati riassunti e analizzati sotto due punti di vista: per il primo si è cercato il numero di istituti sovra specializzati presenti in ciascuna categoria e come questo valore sia cambiato negli anni, mentre per il secondo sono state raggruppate le università per ogni anno in base al numero di categorie in cui sono sovra specializzate. In Figura 18 sono mostrati i risultati relativi al primo aspetto, mentre in Figura 19 gli esiti per il secondo.

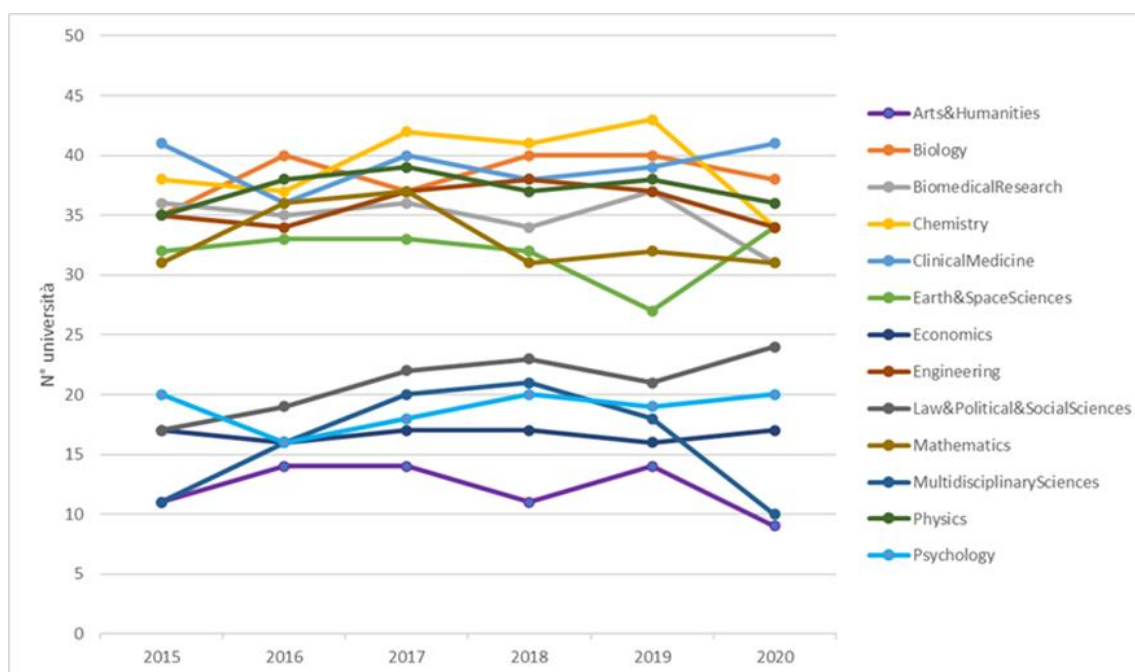


Figura 18 – Andamento del numero di università sovra specializzate in ciascuna macro categoria (2015-2020)

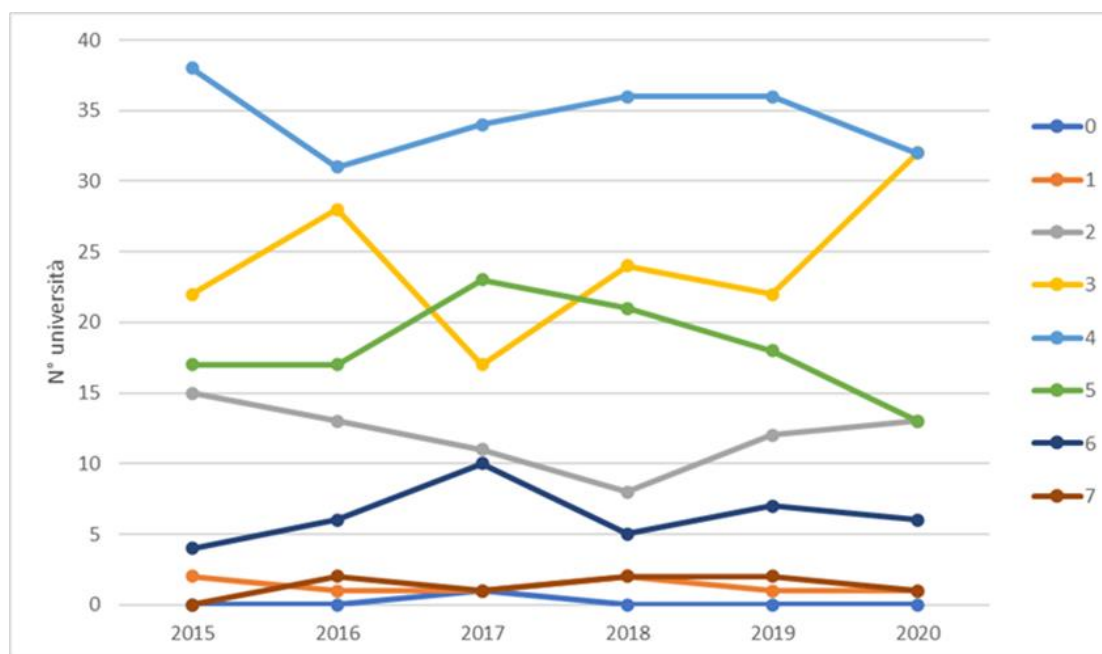


Figura 19 – Andamento del numero di università per numero di categorie in cui sono sovra specializzate (2015-2020)

Dal primo grafico sono facilmente individuabili due macro gruppi, uno in alto composto da “Chemistry”, “Clinical Medicine”, “Biology”, “Physics”, “Engineering”, “Earth and Space Sciences”, “Mathematics” e “Biomedical Research” e uno in basso contenente “Law, Political and Social Sciences”, “Psychology”, “Economics”, “Multidisciplinary Sciences” e “Arts and Humanities”. Tale suddivisione rispecchia esattamente le

conclusioni esposte in precedenza, ovvero che le università italiane, per quanto riguarda la tipologia dell'articolo scientifico, tendono a dedicarsi maggiormente alle attività tecnico-scientifiche piuttosto che a quelle umanistiche, determinando una sovra specializzazione in esse.

All'interno della stessa categoria non sono ravvisabili drastici cambiamenti da un anno all'altro, confermando in altra maniera la continuità e stabilità della ricerca scientifica italiana. Si può, quindi, sostenere che la maggior parte della variabilità presente nel grafico sia dovuto a quelle università che presentano un indice di specializzazione *RSA* vicino al valore di soglia e quindi, di anno in anno, possono raggiungerlo oppure no.

Guardando la Figura 19, invece, si può constatare come poco più di un terzo degli istituti italiani analizzati tende a sovra specializzarsi in quattro macro categorie durante il corso di un anno. Considerando anche quelli con tre o cinque si arriva ad una copertura di quasi l'80%. Molto rari sono i casi estremi di università sovra specializzate in una sola o in sette ambiti differenti, tipicamente istituti di dimensioni modeste che si concentrano su ambiti prevalentemente umanistici, mentre l'unico caso con zero categorie è dovuto alla mancanza di pubblicazioni di quell'ente per l'anno 2017. Per quanto riguarda la situazione dei Politecnici, sia quello di Torino che di Milano rimangono specializzati in 4 categoria durante l'intervallo temporale considerato, mentre un po' più variegata è la condizione per quello di Bari che spazia da tre a cinque ambiti di sovra specializzazione in base all'anno osservato.

Per i valori estremi occorre fare una precisazione. La variabilità caratterizzante il campione analizzato è principalmente dovuta alla presenza di istituti con una quantità modesta di pubblicazioni ed al loro impatto sulla struttura impiegata per la costruzione della metrica. Mostrando graficamente la relazione esistente tra il numero totale di pubblicazioni effettuate da ciascuna università durante il periodo compreso tra il 2015 e il 2020 e il numero medio di categorie in cui si sovra specializzano, come avviene in Figura 20, è possibile notare come i valori più rari si trovino proprio in corrispondenza degli enti con meno pubblicazioni.

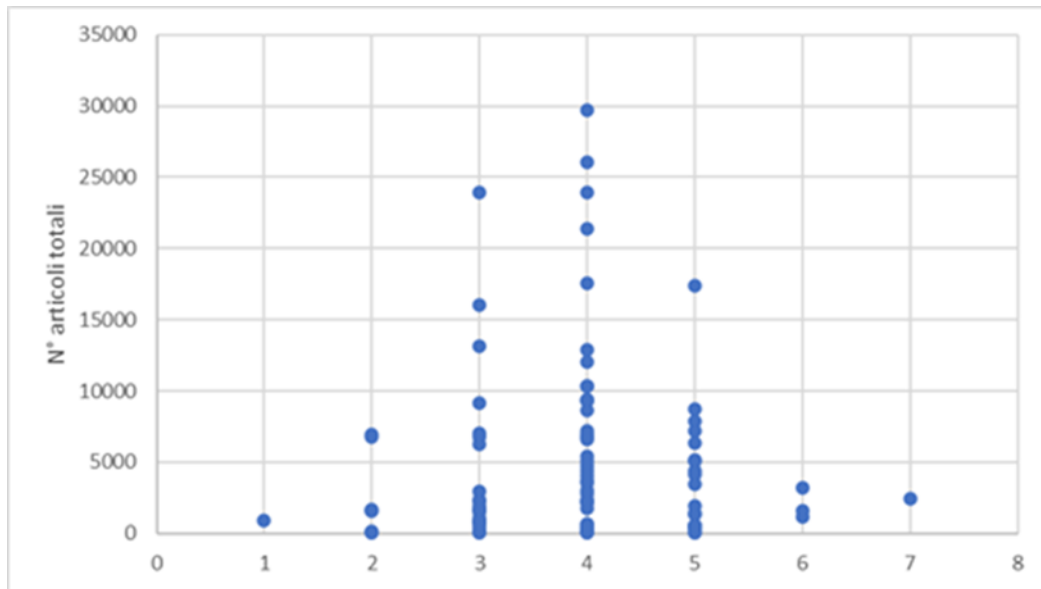


Figura 20 – Distribuzione del numero medio di categorie in cui si è specializzata rispetto al numero totale di articoli pubblicati per ciascuna università (2015-2020)

2.4. Indice di collaborazione

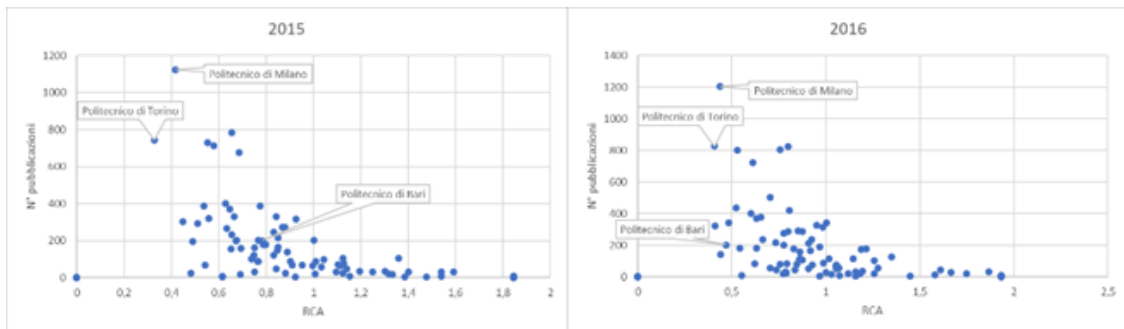
A questo punto, dopo aver introdotto il concetto di specializzazione e aver mostrato i principali trend che sono emersi dagli studi effettuati, si è passati all'analisi dell'aspetto della collaborazione interuniversitaria. Per dare continuità con quanto fatto in precedenza si è optato per un procedimento e un ragionamento simile a quello impiegato per la specializzazione: è stato ricavato un indice di collaborazione, definito come *RCA* (*Relative Collaborative Advantage*), e in seguito è stata realizzata una matrice di collaborazione per ogni anno di pubblicazione. Lo studio di tale fattore è volto a identificare l'esistenza di una sua relazione con la specializzazione. Si è cercato, quindi, di comprendere se le università italiane tendono a cooperare con un altro istituto per mancanza di conoscenza in un certo ambito o se il loro contributo è maggiormente richiesto in presenza di un elevato grado di competenza.

Per la costruzione dell'indice è stato confrontando il rapporto tra le pubblicazioni che sono state condotte in collaborazione con un altro ente dall'università i nella categoria c e il numero di pubblicazioni totali effettuate nella medesima categoria con la media di tale rapporto tra tutte le università analizzate. Per compiere tale confronto si è anche tenuto conto dell'anno di pubblicazione t . La formula viene riportata di seguito:

$$RCA_{c,i,t} = \frac{n^{\circ}art.collaborativi_{c,i,t} / n^{\circ}pubblicazioni_{c,i,t}}{media_i \left(n^{\circ}art.collaborativi_{c,i,t} / n^{\circ}pubblicazioni_{c,i,t} \right)}$$

L'indice rappresenta la tendenza di un'università a condurre ricerche in collaborazione con uno o più enti a seconda dell'ambito scientifico esaminato. Poiché il confronto avviene prendendo in considerazione solamente i dati riferiti ad una determinata categoria, non sarà possibile effettuare dei paragoni tra categorie diverse e, quindi, non si potrà avere evidenza in quali settori vi sia in valore assoluto una maggiore cooperazione. Inoltre, all'interno dello stesso istituto la presenza di elevati valori di RCA non implica una maggiore predisposizione dell'ente a cooperare in quella materia rispetto alle altre come, invece, si potrebbe ipotizzare sull'aspetto della specializzazione. Idealmente un'università potrebbe essere maggiormente collaborativa in tutte le categorie se presenta per ciascuna un indice di collaborazione superiore alla media. Quando $RCA_{c,i,t}$ è uguale 1, l'università i ha una percentuale di collaborazione, nella categoria c e al tempo t , che è uguale alla percentuale media tra tutte le altre università. Per tale motivo quando $RCA_{c,i,t} > 1$, si può solamente affermare che l'università i preferisca condurre, nella categoria c , ricerche in collaborazione rispetto a quanto avviene per una buona parte degli altri istituti.

Seguendo quanto fatto in precedenza, nelle successive figure vengono riportati alcuni dei risultati ottenuti. In particolare, partendo dalle medesime situazioni rappresentate in Figura 16 e Figura 17 ed evidenziando la posizione degli stessi istituti, si è cercato di abbozzare un primo confronto con l'aspetto della specializzazione.



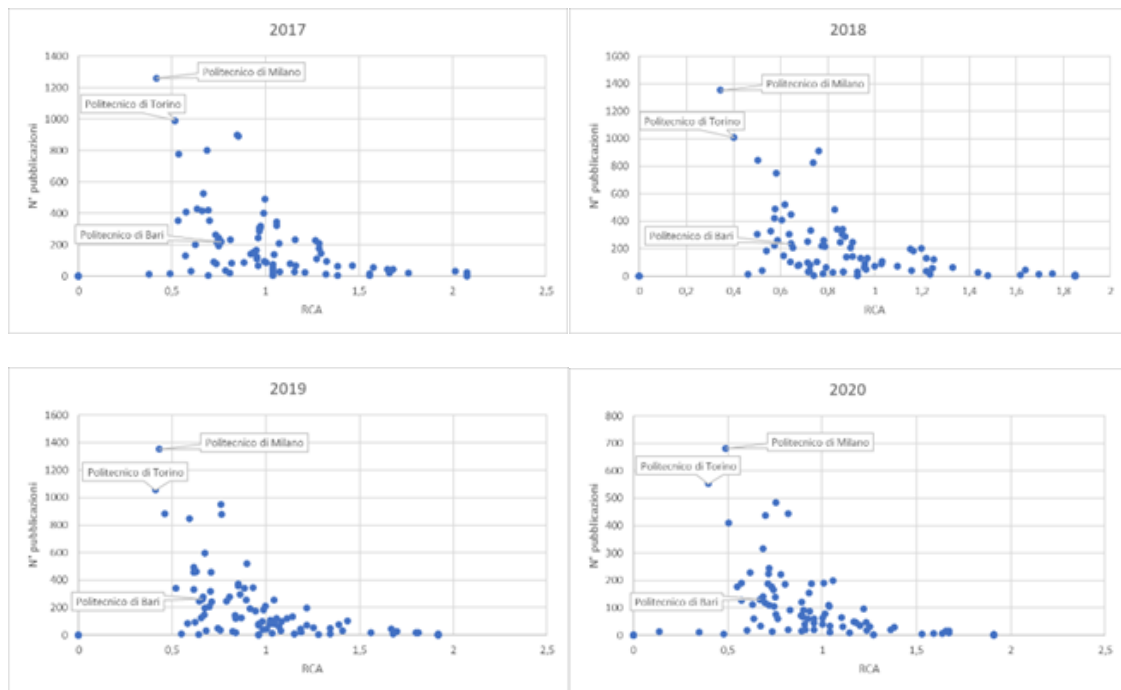


Figura 21 – Distribuzione indice di collaborazione RCA rispetto al numero di articoli pubblicati per la categoria “Engineering” per ogni anno compreso tra il 2015 e il 2020

Osservando i differenti grafici presenti nella Figura 21, si può notare come, per quanto riguarda la macro categoria “Engineering”, sia associato un indice di collaborazione relativamente basso alle università con il maggior numero di articoli pubblicati. Un esempio sono proprio il Politecnico di Torino e di Milano ai quali vengono accostati alcuni dei valori più bassi della categoria. Questa caratteristica, oltre a rispecchiare una superiore capacità e tendenza da parte dei grandi istituti a condurre ricerche per conto proprio, risente del modesto numero di pubblicazioni offerte da alcuni enti. Infatti, minore risulta essere tale quantità, maggiore è la probabilità di avere una percentuale di collaborazione elevata e di conseguenza un indice *RCA* superiore alla media. Tuttavia, tale distorsione è ravvisabile anche in senso opposto. Si può osservare come al diminuire delle pubblicazioni vi sia un aumento proprio della variabilità. Questo effetto diventa più chiaro se si guardano le categorie rappresentate in Figura 22, in particolare “Clinical Medicine” ed “Economics”. Al loro interno è possibile notare, per le università con più articoli, una concentrazione di punti tra i valori 0,8 e 1 di *RCA*. Tale range viene ampliato scendendo lungo l’asse delle ordinate fino a raggiungere la sua massima estensione alla base. Tale situazione conferma come la struttura dell’indice di collaborazione altera leggermente il valore numerico finale ma non influisce sull’orientamento alla cooperazione che continua a dipendere da istituto a istituto. Infine, considerando l’evoluzione negli anni della distribuzione di punti, è stata osservata una certa regolarità,

a conferma di come i comportamenti interni alle università e i loro legami esterni siano persistenti nel tempo.

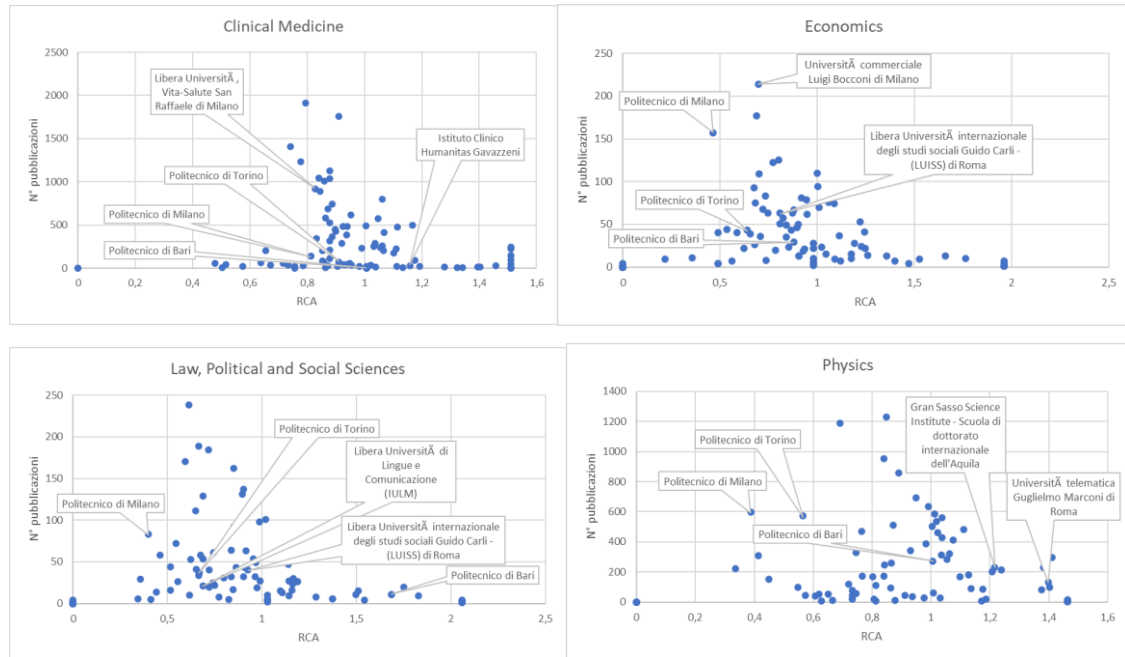


Figura 22 – Distribuzione indice di collaborazione RCA rispetto al numero di articoli pubblicati nell'anno 2019 per le categorie: “Clinical Medicine”, “Economics”, “Law, Political and Social Sciences” e “Physics”

A questo punto, dopo aver esposto qualche breve considerazione anche sull'indice di collaborazione RCA , si è deciso di unire i valori illustrati rispettivamente dalle Figura 16 e Figura 21 e dalle Figura 17 e Figura 22 per poter introdurre i primi risultati in merito al rapporto tra il fattore di specializzazione e di collaborazione.

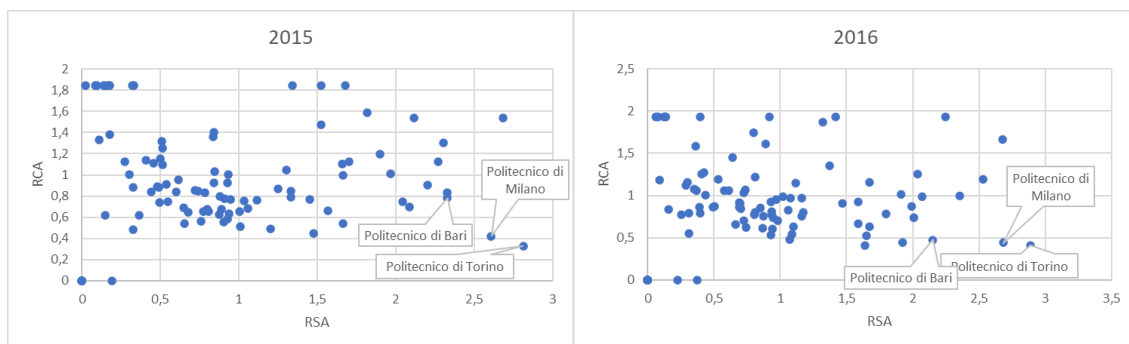




Figura 23 – Distribuzione indice di specializzazione RSA rispetto all'indice di collaborazione RCA per la categoria “Engineering” per ogni anno compreso tra il 2015 e il 2020

In Figura 23 sono stati rappresentati gli indici calcolati per la categoria “Engineering” per i diversi anni di pubblicazione. Per quanto riguarda i Politecnici, si può notare che tutti e tre sono accomunati da un’elevata specializzazione, ma da un basso tasso di collaborazione. Considerando, invece, gli altri istituti risulta difficile, ad una prima occhiata, evidenziare una relazione: la distribuzione dei punti non sembra avere un chiaro orientamento. Concentrandosi sulla parte alta di ciascun grafico, dove sono situati gli enti più collaborativi, è possibile osservare una maggiore concentrazione per gli istituti con RSA inferiori ad 1. Questo potrebbe far pensare che vi sia una relazione inversa, seppur debole, tra specializzazione e collaborazione nel campo dell’ingegneria. Ovvero che le università con meno competenze tendono a colmare il gap di conoscenza tramite l’aiuto e il supporto da parte di altri enti. Per esempio, considerando le macro categorie “Economics” e “Law, Political and Social Sciences” mostrate in Figura 24, alle università con un indice di specializzazione RSA maggiore di 1 è associato un grado di collaborazione relativamente basso, mentre la situazione sembra leggermente ribaltata per gli istituti con un indice RSA minore. Questo risultato potrebbe confermare che, al diminuire del livello di competenza, gli istituti sembrano essere più disponibili a coinvolgere o essere coinvolti in ricerche di gruppo. Tuttavia, se si osserva la distribuzione per le due categorie rimanenti, l’esito precedente non viene ritrovato. Infatti,

in questa circostanza, alle università più specializzate è associato anche un elevato indice di collaborazione RCA . Da questi esempi diventa difficile evidenziare una relazione generale tra i due aspetti e, per tale motivo, è stato necessario uno studio più approfondito. In particolare si è dovuto porre maggiore attenzione al comportamento degli enti non specializzati che costituiscono la casistica più numerosa e incerta.

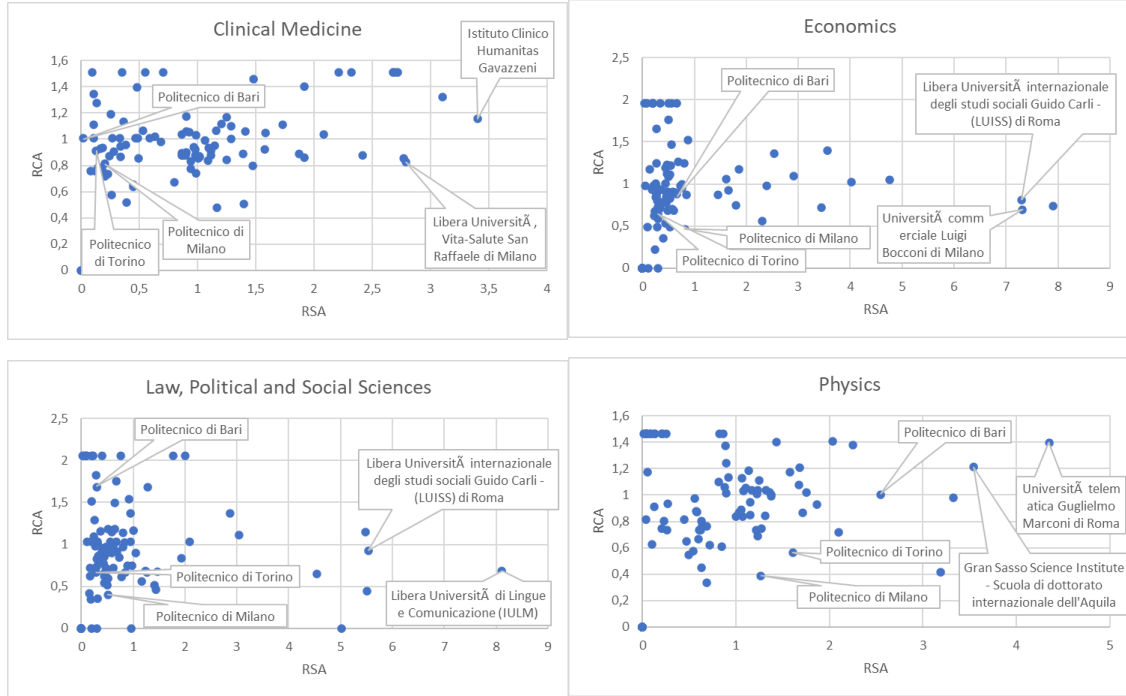


Figura 24 – Distribuzione indice di specializzazione RSA rispetto all'indice di collaborazione RCA nell'anno di pubblicazione 2019 per le categorie: "Clinical Medicine", "Economics", "Law, Political and Social Sciences" e "Physics"

A questo punto, diventa chiaro come nello studio del rapporto esistente tra collaborazione e specializzazione risulta opportuno considerare, oltre al comportamento della singola università, anche l'ambito scientifico oggetto della pubblicazione. Tuttavia, a causa della complessità di un confronto visivo tra questi aspetti, un'analisi più approfondita è lasciata al prossimo capitolo, dove saranno tutti compresi in un modello matematico.

Dopo aver introdotto un primo confronto a livello dettagliato tra questi due fattori e aver presentato i problemi relativi a tale approccio, il passo successivo è stata la realizzazione delle matrici di collaborazione, MC_t , identiche a quelle di specializzazione, con il generico elemento (c, i, t) uguale a 1 se l'università i ha un tasso di collaborazione più elevato della media nella categoria c nell'anno t :

$$MC_t(c, i) = \begin{cases} 1 & \text{se } RSA_{c,i,t} > 1 \\ 0 & \text{se } RSA_{c,i,t} \leq 1 \end{cases}$$

L'introduzione di tali matrici ha permesso di eliminare la grande variabilità presente all'interno dei valori calcolati e di operare un paragone più semplificato e di maggiore chiarezza e comprensione per il lettore.

Prima, però, di passare al vero e proprio confronto si è deciso di esporre i risultati contenuti nelle matrici di collaborazione in maniera del tutto simile a quanto fatto per quelle di specializzazione.

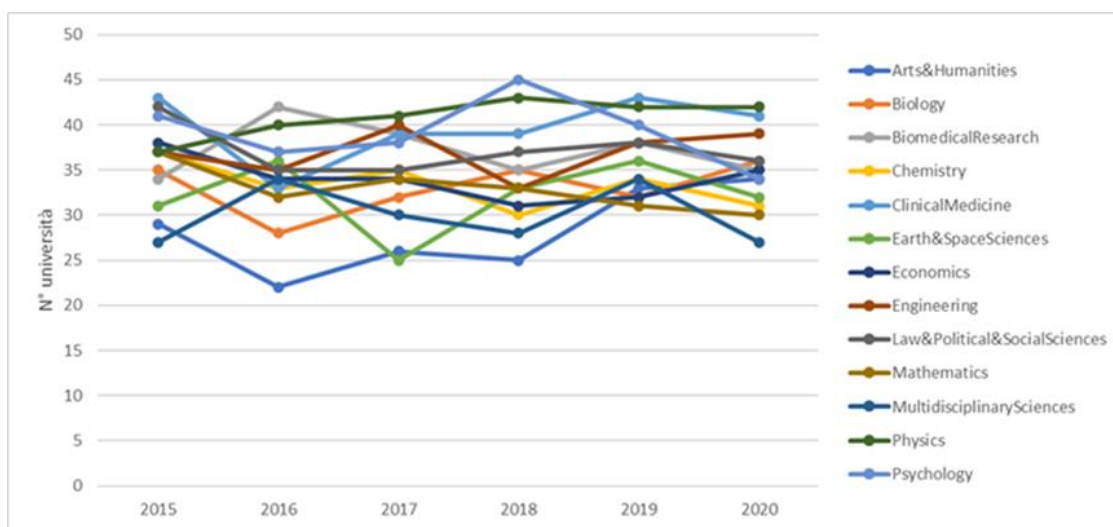


Figura 25 – Andamento del numero di università maggiormente collaborative in ciascuna macro categoria (2015-2020)

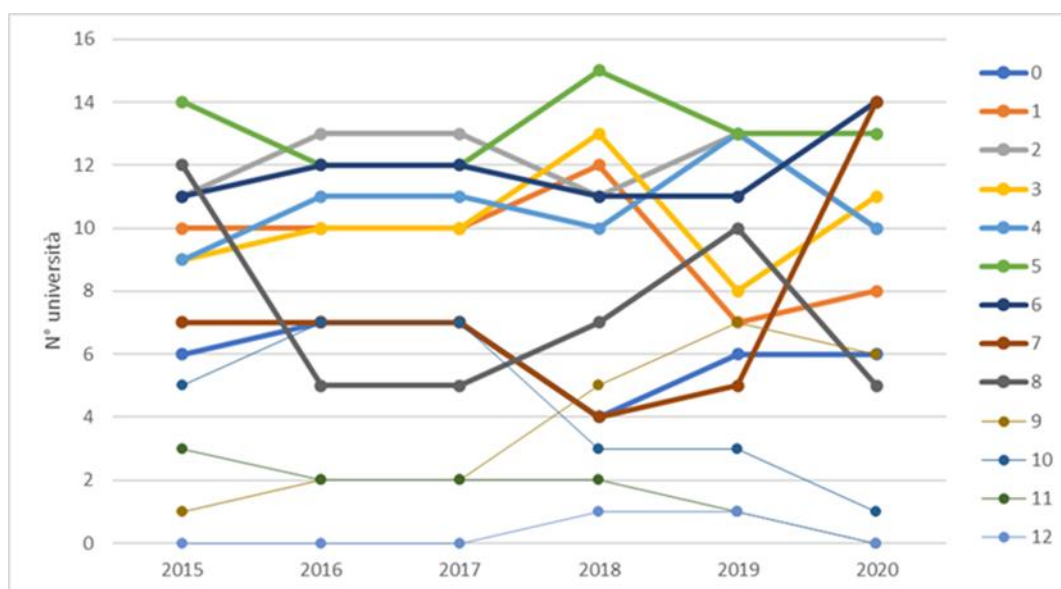


Figura 26 – Andamento del numero di università per numero di categorie in cui sono maggiormente collaborative (2015-2020)

Per quanto riguarda la Figura 25, è corretto che non vi sia quasi differenza tra una categoria e l'altra. Infatti, per come è stato costruito l'indice di collaborazione, è normale

che ciascun ambito scientifico presenti, nel corso degli anni, un numero simile di università con un tasso di collaborazione superiore alla media. Considerando, invece, il numero di categorie in cui ciascun istituto tende ad avere una maggiore cooperazione, come mostrato in Figura 26, si osserva una situazione più variegata. Mediamente le università risultano essere maggiormente collaborative su 4 o 5 differenti macro categorie. Fino al 2018 la suddivisione è rimasta piuttosto stabile con addirittura nessun cambiamento tra il 2016 e il 2017. Negli ultimi due anni, invece, la situazione è cambiata drasticamente, in particolare nel 2020 che, a differenza di quanto si potesse pensare data l'arrivo della pandemia, ha visto un aumento del numero di università collaborative su più ambiti. Infine, è da sottolineare che la presenza di istituti che sono risultati collaborativi in più di 7 o 8 macro categorie differenti è dovuto principalmente alla distorsione, accennata in precedenza, per quelle casistiche che presentano un numero di osservazioni molto modesto. Per esempio, se il Politecnico di Torino avesse una sola pubblicazione in ambito sociale nel 2018 e fosse stata condotta con il coinvolgimento di un altro istituto, allora si avrebbe per quella categoria e per quell'anno una percentuale di collaborazione del 100% e quindi avrebbe un indice *RCA* sicuramente superiore a 1. Tuttavia, seguendo la decisione presa per un problema simile riscontrato nell'indice di specializzazione, si è optato per non prendere nessun rimedio a tale distorsione nel presente capitolo.

2.5. Confronto specializzazione e collaborazione

Dopo aver presentato i principali trend e problemi connessi alle metriche utilizzate per la descrizione dell'aspetto della specializzazione e della collaborazione si è passati al confronto diretto tra le matrici precedentemente ottenute.

Data la stretta somiglianza tra le matrici di specializzazione e di collaborazione per come sono state pensate e realizzate, si è optato di valutare il loro grado di sovrapposizione. Infatti, valutando per ogni singolo elemento (c,i,t) se vi fosse coincidenza tra le due matrici, è possibile ottenere un'indicazione generale di orientamento della ricerca scientifica. Qualora vi fosse un cospicuo livello di somiglianza tra le matrici si potrebbe sostenere che tra la specializzazione e la collaborazione vi sia una sorta di correlazione positiva, poiché all'aumentare dell'una si otterrebbe un cambiamento nello stesso verso anche dell'altra. Inoltre, un tale risultato avrebbe valenza generale perché, avendo

confrontato contemporaneamente tutti gli ambiti scientifici, non dipenderebbe dall'argomento specifico trattato nella ricerca.

Per il confronto è stata cercata in letteratura una misura che potesse restituire, in maniera sintetica e credibile, un'indicazione del loro livello di somiglianza. Tale ricerca ha avuto esito nell'indice di Jaccard. L'indice di Jaccard, noto anche come coefficiente di somiglianza di Jaccard, è una statistica utilizzata per misurare la somiglianza e la diversità di insiemi di campioni finiti [46]. Il coefficiente di Jaccard è definito come la dimensione dell'intersezione divisa per la dimensione dell'unione degli insiemi considerati:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

Si noti che in base alla progettazione, se A intersezione B è l'insieme vuoto, allora $J(A, B) = 0$. Il coefficiente di Jaccard è ampiamente utilizzato in informatica, ecologia, genomica e altre scienze, dove vengono utilizzati dati binari o resi tali, come in questo contesto. Data la sua struttura, l'indice di Jaccard sarà sempre un valore compreso tra 0 e 1.

$$0 \leq J(A, B) \leq 1$$

Collegato ad esso, vi è la distanza di Jaccard, una misura della dissimulazione tra insiemi campionari, che è complementare al coefficiente di Jaccard e si ottiene sottraendo il coefficiente di Jaccard da 1, o, in maniera equivalente, dividendo la differenza delle dimensioni dell'unione e dell'intersezione di due insiemi per la dimensione dell'unione:

$$d_J(A, B) = 1 - J(A, B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}$$

Come detto in precedenza, questo indice si addice particolarmente al confronto tra oggetti binari perché è un'utile misura della sovrapposizione tra i loro attributi. Per il suo calcolo, poiché gli attributi delle matrici di specializzazione e di collaborazione possono assumere solamente il valore 0 o 1, si è partiti raggruppando le varie combinazioni di attributi nella seguente maniera:

- M_{11} rappresenta il numero totale di attributi in cui entrambe le matrici hanno valore di 1.
- M_{01} rappresenta il numero totale di attributi in cui l'attributo della matrice di specializzazione è 0 e l'attributo della matrice di collaborazione è 1.
- M_{10} rappresenta il numero totale di attributi in cui l'attributo della matrice di specializzazione è 1 e l'attributo della matrice di collaborazione è 0.

- M_{00} rappresenta il numero totale di attributi in cui entrambe le matrici hanno valore pari a 0.

Ogni attributo deve rientrare in una di queste quattro categorie, il che significa che la loro somma deve essere uguale al numero di elementi, n , presenti in una matrice:

$$M_{11} + M_{01} + M_{10} + M_{00} = n$$

A questo punto l'indice di similarità di Jaccard, J , è dato come:

$$J = \frac{M_{11}}{M_{01} + M_{10} + M_{11}}$$

E la distanza di Jaccard, d_J , come:

$$d_J = \frac{M_{01} + M_{10}}{M_{01} + M_{10} + M_{11}} = 1 - J$$

Poiché ciascun anno di pubblicazione compreso tra il 2015 e il 2020 è caratterizzato da una propria matrice di specializzazione e di collaborazione, per la realizzazione dell'analisi sono stati calcolati sei differenti valori dell'indice di Jaccard. Facendo la media tra di essi è stato ottenuto un valore pari a circa 0,214 e quindi un d_J pari a 0,786. Questo risultato smentisce l'esistenza di una relazione positiva tra la specializzazione e la collaborazione. A seguito dei valori ottenuti, si deduce che i due tipi di matrice siano piuttosto dissimili tra di loro e, quindi, risulta improbabile che un'università specializzata in un determinato ambito scientifico sia anche maggiormente cooperativa. Tuttavia, data l'elevata distanza esistente, non si può escludere che vi sussista una relazione inversa. Le variabili M_{01} e M_{10} rappresentano il numero di istituti che possiedono solo uno dei due indici, *RSA* o *RCA*, superiore ad 1 e, di conseguenza, università che, in una certa categoria, sono specializzate, ma non godono di un elevato tasso di collaborazione o viceversa. Per poter ottenere certi valori dell'indice di Jaccard è necessario che la condizione rappresentata dai due sottoinsiemi sia preponderante nel contesto analizzato. Di conseguenza si può constatare come le università italiane tendano a cooperare maggiormente quando il loro livello di conoscenza e di competenza risulta carente. Tale risultato, oltre a far emergere una probabile relazione negativa tra questi due aspetti, avrebbe un valore di carattere generale in quanto per il suo raggiungimento non si è operata alcuna distinzione a livello di singolo ambito scientifico. Inoltre, potrebbe

confermare il ruolo centrale della collaborazione come strumento per l'acquisizione di sapere e di esperienza nel minor tempo possibile.

Dopo aver affrontato l'argomento a livello generico, per valutare se le conclusioni a cui si è giunti in precedenza trovassero riscontro anche a un livello di dettaglio maggiore, si è passati ad un'analisi più minuziosa andando ad analizzare la situazione settore per settore. Per procedere con questo tipo di analisi si è deciso di suddividere le varie università in 4 categorie, come mostrato in Figura 27, a seconda dell'ambito scientifico e dell'anno di pubblicazione considerato: (1) “*Open & Specialized*” se presenta sia l'indice di specializzazione che l'indice di collaborazione superiore a 1; (2) “*Specialized*” se presenta solo l'indice di specializzazione superiore a 1; (3) “*Collaborative*” se presenta solo l'indice di collaborazione superiore a 1; (4) “*Ordinary*” se entrambi gli indici sono pari o inferiori a 1.

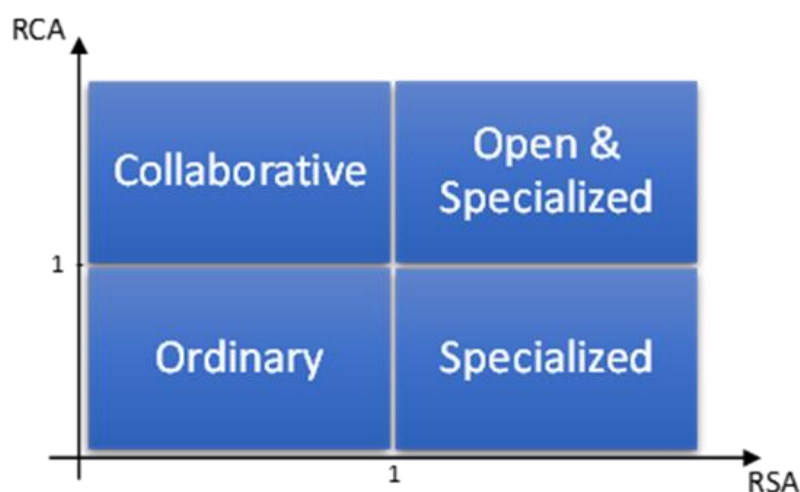


Figura 27 – Classificazione delle università sulla base dei valori associati di RSA e RCA

Nelle successive tabelle vengono mostrati i risultati del raggruppamento per alcune macro categorie. Per avere una visione un po' più completa si è deciso di non mostrare gli ambiti già in parte trattati.

Tabella 3 – Classificazione per le macro categorie “Biology”, “Chemistry”, “Earth&SpaceSciences” e “Psychology” (2015-2020)

Biology	2015	2016	2017	2018	2019	2020	Chemistry	2015	2016	2017	2018	2019	2020
Open & Specialized	6	6	7	9	8	11	Open & Specialized	7	6	8	9	10	5
Specialized	29	34	30	31	32	27	Specialized	31	31	34	32	33	29
Collaborative	29	22	25	26	24	25	Collaborative	30	27	27	21	24	26
Ordinary	24	27	29	30	28	24	Ordinary	13	17	15	20	16	18

Earth&SpaceSciences	2015	2016	2017	2018	2019	2020
Open & Specialized	9	16	8	10	11	13
Specialized	23	17	25	22	16	21
Collaborative	22	20	17	23	25	19
Ordinary	35	34	37	36	37	39

Psychology	2015	2016	2017	2018	2019	2020
Open & Specialized	11	11	11	12	10	12
Specialized	9	5	7	8	9	8
Collaborative	30	26	27	33	30	22
Ordinary	28	37	38	28	32	30

Prima di passare al confronto vero e proprio tra collaborazione e specializzazione, si è preferito esporre qualche informazione a corredo dei dati che non è possibile dedurre dalla rappresentazione in tabella. La suddivisione degli istituti nelle quattro classi è molto più variegata di quel che sembra. Infatti per ogni categoria sono differenti le università che compongono i vari gruppi. Se prendiamo, ad esempio, il Politecnico di Torino, di Milano e di Bari, essi rientrano quasi sempre nella classe degli enti “*Ordinary*” per l’ambito “*Biology*” e “*Earth & Space Sciences*”. Solo quello di Bari nel 2018 e 2020 finisce tra le “*Collaborative*” per il primo, mentre tra il 2015 e il 2017 in “*Specialized*” per il secondo. Considerando la categoria “*Chemistry*” il Politecnico di Torino e di Milano fanno parte degli enti “*Specialized*”, mentre quello di Bari è partito dal gruppo “*Collaborative*” ed è finito in quello “*Specialized*” passando tra il 2018 e il 2019 in quello “*Ordinary*”. Infine per l’ultima macro categoria esposta i tre Politecnici si alternano negli anni tra la classe “*Collaborative*” e “*Ordinary*”. Se, invece, osserviamo un altro istituto, per esempio l’Istituto universitario di studi superiori di Pavia, esso rientra tra gli enti “*Open & Specialized*” per la categoria “*Psychology*”, “*Specialized*” per “*Earth & Space Sciences*” e tra “*Ordinary*” e “*Collaborative*” per i due ambiti rimanenti. Questo conferma come la situazione cambi da ente a ente a seconda dell’argomento scientifico che viene trattato.

Passando all’analisi vera e propria dei numeri presenti nelle tabelle, da un rapido sguardo si può notare come i valori di ciascun gruppo non cambino drasticamente da un anno all’altro. Questo permette di ribadire ancora una volta come il comportamento delle università italiane tende a rimanere costante da un anno all’altro. Di conseguenza anche gli spostamenti da una classe all’altra avvengono in maniera graduale. Se un’università viene classificata come “*Ordinary*” in un anno, risulta complicato che l’anno dopo faccia parte degli enti “*Open & Specialized*” e viceversa. Lo stesso avviene tra le classi “*Collaborative*” e “*Specialized*” e, per tale motivo, Il passaggio tra gruppi non adiacenti si rivela un evento raro.

A questo punto, in continuità con quanto ritrovato dallo studio delle matrici, si può scorgere che per alcuni ambiti scientifici gli enti classificati come “*Collaborative*” o “*Specialized*” siano più numerosi rispetto agli altri. Questa situazione è particolarmente

evidente nelle macro categorie “*Biology*” e “*Chemistry*”, mentre risulta trascurabile, se non nulla, in quelle rimanenti. Ad esempio, per “*Psychology*” si hanno più università “*Open & Specialized*” che “*Specialized*” e la somma delle prime con quelle “*Ordinary*” è spesso superiore alle seconde con le “*Collaborative*”. Considerando la classificazione anche per le altre macro categorie, non riportate nel presente lavoro, l’unico aspetto comune che emerge è la maggiore presenza di istituti classificati come “*Collaborative*” o “*Specialized*” rispetto a quelli “*Open & Specialized*”. A seguito di tale esito appare ancora più evidente come non si possa affermare che sussista una relazione positiva tra l’aspetto della specializzazione e della collaborazione. In presenza di un fattore è molto più probabile che l’altro sia trascurabile piuttosto che rilevante.

Il passo successivo è stato quello di verificare se la seconda ipotesi, emersa durante il confronto delle matrici, possa essere confermata: l’esistenza di una dipendenza negativa. Tuttavia, a causa di una quantità non trascurabile di università classificate come “*Ordinary*” per ciascuna macro categoria, che andrebbero ad influenzare il calcolo di un indice di correlazione, dalle analisi finora proposte non è possibile dare un esito a tale quesito. Sarà necessario l’introduzione di un modello matematico, in particolare quello della regressione, per poter valutare l’esistenza di una relazione tra l’aspetto della collaborazione e della specializzazione.

Prima di passare al prossimo capitolo, che presenterà l’analisi di regressione compiuta, sono state ripetute le precedenti analisi cambiando una delle due metriche utilizzate. Infatti, a differenza di quanto accaduto con la specializzazione, per l’indice di collaborazione non si ha avuto un riscontro dalla letteratura, ma è stato costruito secondo il giudizio dell’autore. Data la soggettività che ha contraddistinto il processo di realizzazione dell’indice, si è cercato di valutare come i risultati precedentemente ottenuti fossero sensibili a un cambiamento di questo tipo.

Per il calcolo del nuovo indice di collaborazione, indicato come $RCA2$, è stato mantenuto a numeratore il rapporto tra pubblicazioni che sono state effettuate in collaborazione e pubblicazioni totali in una determinata categoria C , mentre a denominatore non si trova più la media di tale valore tra i 98 istituti, ma bensì la media di tale valore tra le varie categorie della medesima università i . Di seguito viene riportata la formula:

$$RCA2_{c,i,t} = \frac{n^{\circ}art.collaborativi_{c,i,t} / n^{\circ}pubblicazioni_{c,i,t}}{media_c \left(n^{\circ}art.collaborativi_{c,i,t} / n^{\circ}pubblicazioni_{c,i,t} \right)}$$

L'indice, così formulato, permette di evidenziare in quali categorie ciascuna università sia più propensa a condurre studi in collaborazione con altri enti. Pertanto, quando $RCA2_{c,i,t} > 1$, allora è possibile affermare che l'università i nell'anno t abbia collaborato maggiormente nella categoria c rispetto ad altre. A differenza dell'indicatore precedente, il confronto viene effettuato internamente al singolo istituto e non rispetto agli altri enti. Di conseguenza il paragone tra due università su chi sia maggiormente collaborativa in un determinato ambito non è più ammissibile.

Anche per questa metrica sono state realizzate le matrici di collaborazione per ciascun anno di pubblicazione studiato. Nelle due seguenti figure vengono riportati i principali risultati.

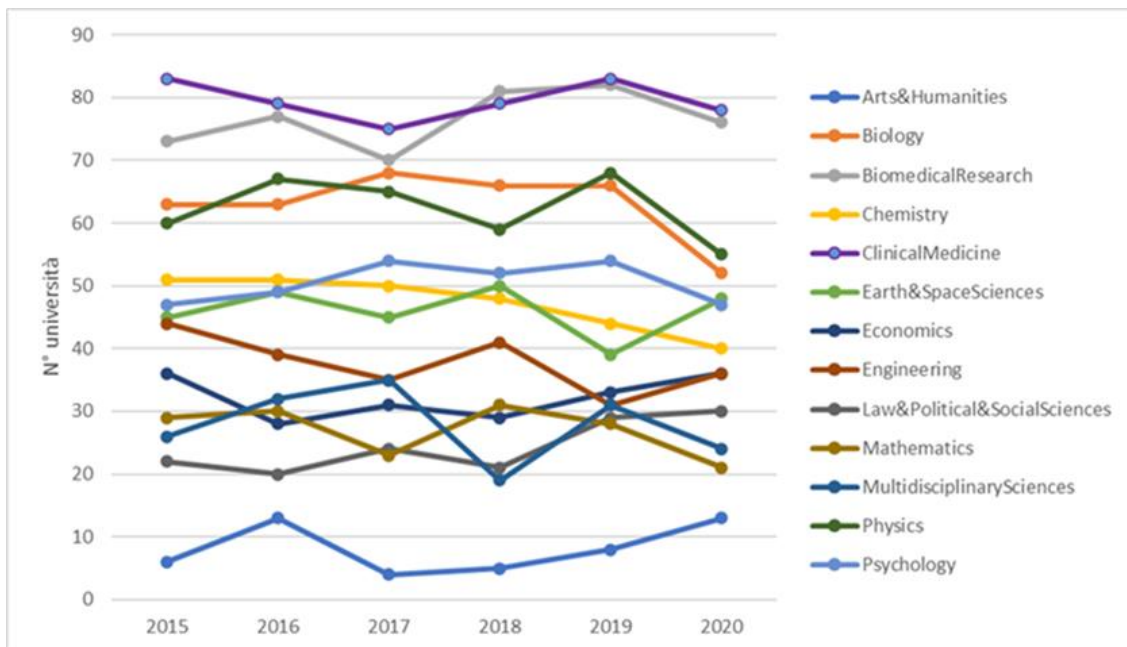


Figura 28 – Andamento del numero di università maggiormente collaborative per l'indice RCA2 in ciascuna macro categoria (2015-2020)

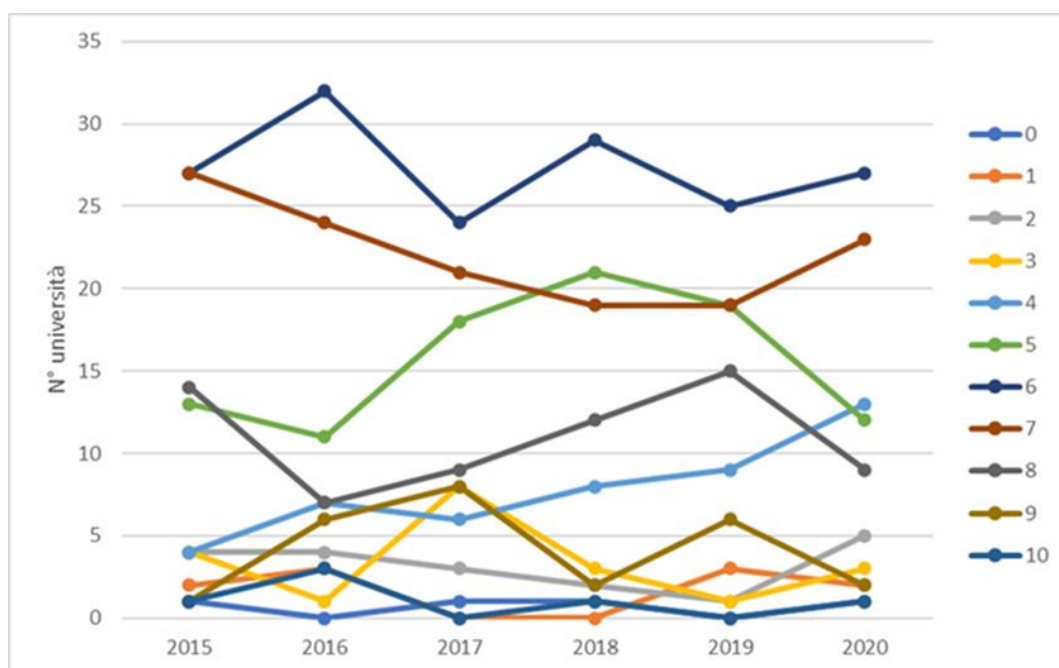


Figura 29 – Andamento del numero di università per numero di categorie in cui sono maggiormente collaborative secondo l'indice RCA2 (2015-2020)

A differenza di quanto avvenuto con il primo indice di collaborazione, in Figura 28 si può avere evidenza di quali siano gli ambiti scientifici in cui le università italiane hanno cooperato in maniera più frequente. In continuità con quanto ricavato dall'analisi delle singole aree di ricerca riportata all'inizio del capitolo, le categorie con il maggiore tasso di collaborazione sono inerenti all'ambito medico e fisico. Infatti, tra di esse si osservano: “*Clinical Medicine*”, “*Biomedical Research*”, “*Physics*” e “*Biology*”. In particolare, se si considerano le prime due è sorprendente come su 98 istituti analizzati in media quasi 80 possiedono per questi ambiti un tasso di collaborazione superiore a quello che possiedono per gli altri argomenti. Dall'altra parte della classifica, invece, si trovano materie inerenti al campo umanistico e sociale come “*Psychology*” e “*Law, Political and Social Sciences*”. Dal precedente risultato si può affermare con un opportuno livello di sicurezza che la cooperazione interuniversitaria dipenda in larga misura dal campo scientifico studiato. Tuttavia, quello che rimane da capire è se questo fattore possa essere influenzato anche dal grado di specializzazione in quel determinato campo.

Osservando, invece, la Figura 29 sono ravvisabili informazioni riguardanti la distribuzione delle percentuali degli articoli collaborativi all'interno di ciascuna università. Per come è stato formulato l'indice è corretto che una buona parte degli istituti siano maggiormente collaborativi in 5, 6 o 7 categorie. La presenza di enti cooperativi in più o meno ambiti permette di avere indicazioni sulla differenza esistente tra i valori delle

percentuali. Infatti nel caso di più categorie è probabile che all'interno del portafoglio di pubblicazioni di un istituto siano presenti una o due materie con un tasso di collaborazione molto basso se non nullo. Tale situazione ha come risultato un abbassamento della media complessiva e, quindi, una maggiore probabilità di essere cooperativo nelle altre categorie. Nel caso contrario, invece, il numero ridotto potrebbe essere dovuto o ad una mancanza di pubblicazioni su alcune categorie o la presenza di valori di collaborazione molto alti, tendenti all'unità, che hanno come conseguenza un effetto opposto al caso precedente.

Dopo queste considerazioni iniziali si è passati al confronto con le matrici di specializzazione. È stato calcolato, come in precedenza, l'indice di Jaccard per ciascun anno di pubblicazione e facendone la media è stato ottenuto un valore di 0,266 che non si discosta molto da quello ottenuto in precedenza. Dati i valori simili anche le considerazioni che possono essere riportate sono le medesime.

A questo punto, poiché il paragone a livello generale ha dato gli stessi risultati ottenuti con il primo indice di collaborazione, si è reso necessario valutare se un'analoga situazione si ritrovasse anche a livello delle singole categorie. Utilizzando la stessa classificazione effettuata precedentemente, vengono riportati di seguito i risultati ottenuti per alcune macro categorie.

Tabella 4 – Classificazione per le macro categorie “Biology”, “Chemistry”, “Earth&SpaceSciences” e “Psychology” rispetto all'indice RCA2 (2015-2020)

Biology	2015	2016	2017	2018	2019	2020
Open & Specialized	22	23	22	21	24	16
Specialized	13	17	15	19	16	22
Collaborative	35	33	39	39	36	30
Ordinary	18	16	15	17	16	19

Chemistry	2015	2016	2017	2018	2019	2020
Open & Specialized	17	17	18	20	18	9
Specialized	21	20	24	21	25	25
Collaborative	32	28	28	24	21	26
Ordinary	11	16	14	17	19	18

Earth&SpaceSciences	2015	2016	2017	2018	2019	2020
Open & Specialized	15	21	13	14	9	17
Specialized	17	12	20	18	18	17
Collaborative	24	25	27	32	28	25
Ordinary	33	29	27	27	34	33

Psychology	2015	2016	2017	2018	2019	2020
Open & Specialized	11	13	11	11	11	12
Specialized	9	3	7	9	8	8
Collaborative	31	34	39	38	39	32
Ordinary	27	29	26	23	23	20

Confrontando i numeri dei raggruppamenti presenti in Tabella 4 rispetto a quelli mostrati in Tabella 3, si può notare come per le categorie “Biology” e “Chemistry” vi sia stato un notevole aumento degli istituti classificati come “Open & Specialized” e una leggera diminuzione di quelli “Ordinary”. Questo cambiamento, in linea con quanto mostrato in Figura 28, è dovuto alla crescita drastica del numero di università considerate

collaborative in quei determinati ambiti. Tuttavia, i cambiamenti apportati dal nuovo indice di collaborazione sono rimasti circoscritti alle materie in campo medico e fisico. È possibile, infatti, ritrovare una suddivisione pressoché identica per le altre due categorie mostrate in figura: “*Earth & Space Sciences*” e “*Psychology*”.

In generale, l’aumento del numero di università classificate come “*Open & Specialized*” o come “*Collaborative*” per alcuni ambiti scientifici non è sufficiente per poter dedurre qualche tipo di relazione tra l’aspetto della collaborazione e della specializzazione. Anzi, il cambiamento avvenuto per le categorie “*Biology*” e “*Chemistry*” ha reso ancora meno evidente l’esistenza di una relazione inversa. Con il nuovo indice di collaborazione, infatti, la distribuzione degli istituti nei vari gruppi è diventata più omogenea rendendo ancora più complicato trarre delle conclusioni.

Dopo aver presentato i risultati conseguiti con l’utilizzo dell’indice *RCA2*, si può confermare che il cambio di una variabile non ha apportato alcun beneficio nel raggiungimento dell’obiettivo del presente elaborato. Infatti, è risultato chiaro che difficilmente possa sussistere una relazione positiva tra l’aspetto della specializzazione e quello della collaborazione, ma non è stato possibile aggiungere qualche considerazione interessante sulla seconda supposizione che è stata derivata, ovvero che, se un istituto risulta specializzato in un determinato settore, appare più improbabile che sia anche maggiormente collaborativo e viceversa. L’utilizzo di due diversi indici di collaborazione, quindi, ha prodotto esiti simili. A questo punto viene lasciato al prossimo e ultimo capitolo la presentazione di un modello matematico che possa restituire qualche risultato più strutturato e statisticamente valido.

Capitolo III

3.1. Analisi di regressione

L'analisi di regressione è un metodo statistico, particolarmente utilizzato dall'econometria, che permette di studiare la relazione tra due o più variabili. In particolare, tramite modelli matematici tenta di stimare l'effetto prodotto su una di queste, denominata variabile dipendente, dalla variazione delle altre, definite variabili indipendenti. I risultati ottenuti vengono impiegati per quantificare grandezze rilevanti, effettuare inferenza statistica, prevedere l'andamento futuro di alcuni indicatori, testare ipotesi o definire relazioni di causalità. Nel caso oggetto della presente tesi gli strumenti derivati dall'analisi di regressione sono stati adoperati per verificare l'esistenza di una relazione tra l'aspetto della specializzazione e della collaborazione universitaria nell'ambito della ricerca scientifica italiana.

Per poter derivare la presenza di un effetto causale tra due variabili bisognerebbe condurre un esperimento controllato casualizzato che consiste nel paragonare una medesima caratteristica tra due gruppi che differiscono semplicemente per aver ricevuto o meno un trattamento e che, di conseguenza, vengono indicati come gruppo di trattamento e gruppo di controllo. La peculiarità di tale esperimento risiede nel poter assegnare il trattamento in maniera casuale. Tramite questo espediente è ragionevole aspettarsi che venga eliminata la presenza di una relazione sistematica tra la ricezione del trattamento ed altre caratteristiche intrinseche del gruppo interessato che andrebbero ad influire sul risultato finale. Sotto tale ipotesi, quindi, si può sostenere che la differenza dei risultati ottenuti per i due gruppi sia imputabile al trattamento e che non vi sia distorsione da variabile omessa.

Tuttavia, gli esperimenti controllati casualizzati rappresentano un riferimento teorico ideale in quanto difficili o impossibili da realizzare, spesso dovuti ad una mancanza di tecnica, risorse o addirittura per motivi legati all'etica. In un contesto variegato e complesso come il mondo reale, quasi la totalità di queste analisi vengono effettuate con il supporto di dati osservazionali, ovvero derivati dall'osservazione del comportamento naturale di fenomeni al di fuori di un contesto sperimentale. Di conseguenza, anche per il compimento di quest'ultimo studio si è partiti dal medesimo *database* impiegato nelle analisi precedenti.

3.2. Modello impiegato

Dalle osservazioni effettuate nel secondo capitolo si è potuto constatare che l'effetto temporale abbia avuto un impatto trascurabile sulle metriche utilizzate e che il contesto rappresentato sia rimasto piuttosto stabile negli anni. Per tale motivo si è deciso di effettuare un'analisi statica, ovvero un'analisi in cui non è stata considerata la variabile temporale rappresentata dall'anno di pubblicazione degli articoli. Inoltre, poiché la tipologia di analisi lo permetteva e per avere un maggiore livello di dettaglio, si è preferito compiere il confronto tra specializzazione e collaborazione suddividendo i dati posseduti per ciascuna università a livello delle singole aree di ricerca e non più nelle macro categorie che erano state introdotte fittiziamente. Tale scelta ha permesso di tenere in considerazione anche la parte di variabilità che era venuta meno con il raggruppamento.

Per la definizione dell'aspetto della specializzazione e della collaborazione sono state utilizzate le medesime metriche, definite durante la trattazione del presente elaborato, per dare continuità al lavoro svolto e rendere comparabili i risultati ottenuti in precedenza con quelli che saranno estratti dal modello econometrico impiegato. In particolare, per quanto riguarda il fattore della cooperazione interuniversitaria è stato preso in considerazione il primo indice calcolato, *RCA*, in quanto si è potuto verificare che gli esiti conseguiti con una seconda metrica non si discostassero molto da quelli precedenti. L'unico adattamento operato è stato quello di adeguare la formulazione degli indici alla suddivisione nelle singole aree di ricerca.

Sempre da quanto è emerso nel secondo capitolo, le metriche impiegate sono afflitte da alcune problematiche che potrebbero alterare i risultati del modello di regressione. L'indice di specializzazione, per esempio, non definisce in maniera inequivocabile e oggettiva il grado di specializzazione di un ente su un determinato argomento. Infatti, qualora venissero paragonati due istituti con un indice di specializzazione simile, non sempre si può affermare che l'università a cui è associato il valore più elevato sia anche quella maggiormente specializzata. Solitamente all'istituto con un elevato volume di pubblicazioni è associata una minore concentrazione nelle tematiche studiate e di conseguenza un livello di specializzazione inferiore rispetto a quello che può essere definito per un ente più piccolo che, tuttavia, non si riflette nella realtà dei fatti. Questa distorsione è legata alla formulazione stessa dell'indice e si manifesta principalmente in presenza di enti con un modesto numero di pubblicazioni. Un discorso analogo è

applicabile anche per l'indice di collaborazione. Per evitare che i differenti gradi di specializzazione e di collaborazione potessero influire sull'esito dell'analisi si è deciso di considerare soltanto due possibili situazioni per ciascun fattore: specializzato o no e collaborativo o no. Si è optato, quindi, di racchiudere i possibili valori degli indici attraverso una variabile binaria, o *dummy*, come avvenuto per la costruzione delle matrici di specializzazione e di collaborazione, nella seguente maniera:

$$rsa_d_{a,i} = \begin{cases} 1 & \text{se } RSA_{a,i} > 1 \\ 0 & \text{se } RSA_{a,i} \leq 1 \end{cases} \quad rca_d_{a,i} = \begin{cases} 1 & \text{se } RCA_{a,i} > 1 \\ 0 & \text{se } RCA_{a,i} \leq 1 \end{cases}$$

Il pedice a rappresenta l'area di ricerca. Dal momento che l'obiettivo del presente elaborato è circoscritto ad individuare una relazione tra l'aspetto della specializzazione e della collaborazione universitaria a livello complessivo di ricerca scientifica italiana, la semplificazione introdotta ha permesso di considerare contemporaneamente e senza particolari considerazioni i valori ottenuti per le differenti aree di ricerca. Infatti, per alcuni ambiti, principalmente quelli oggetto di un minor numero di articoli scientifici, sono stati ottenuti dei valori di RSA estremamente elevati che non rispecchiano una situazione realistica, ma sono influenzati dalle distorsioni che le università con una modesta quantità di pubblicazioni possono provocare nel calcolo dell'indice. Poiché tali valori, in un modello di regressione lineare, si comporterebbero da *outliers* andando ad influenzare i risultati finali, la scelta di un operatore binario ha mitigato questo problema.

A questo punto, l'introduzione delle variabili *dummy* ha reso necessario l'utilizzo di un modello di regressione non lineare. Tra le opzioni ammissibili si è scelto il modello definito come “*Logit*” che modella la probabilità che la variabile dipendente sia uguale 1 utilizzando la funzione di ripartizione logistica *standard*.

Il modello di regressione finale impiegato per analizzare la relazione tra specializzazione e collaborazione universitaria è il seguente:

$$Pr(rsa_d_{a,i} = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 rca_d_{a,i} + \beta_2 Tot_{a,i} + \beta_3 HHI_i + \beta_j Area_i + \varepsilon_{a,i})}}$$

dove i rappresenta l' i -esima università e a le differenti aree di ricerca. Inoltre:

- $rsa_d_{a,i}$ è la variabile dipendente *dummy* che indica se l'università i è considerata specializzata nell'area di ricerca a ;
- $rca_d_{a,i}$ è la variabile *dummy* di interesse che indica se l'università i è considerata maggiormente collaborativa nell'area di ricerca a ;

- $Tot_{a,i}$, HHI_i e $iArea_i$ sono delle variabili di controllo per alcuni fattori che possono influire sulla specializzazione di un istituto. In particolare $Tot_{a,i}$ rappresenta il numero di articoli che l'università i ha pubblicato nell'area di ricerca a nel periodo considerato, HHI_i rappresenta l'indice di concentrazione che è stato calcolato nel secondo capitolo per ciascun istituto, infine $iArea_i$ rappresenta una serie di variabili *dummy* che catturano l'area geografica in cui è localizzata ciascuna università;
- $\beta_0, \beta_1, \dots$ rappresentano i parametri da stimare per valutare l'impatto di ciascuna variabile.

Nel modello finale è stato preso in considerazione anche l'aspetto geografico per valutare se la localizzazione di un istituto possa influire sul suo comportamento. Le università studiate, in base alla loro posizione geografica, sono state suddivise in cinque aree. In Figura 30 viene mostrata la loro distribuzione.

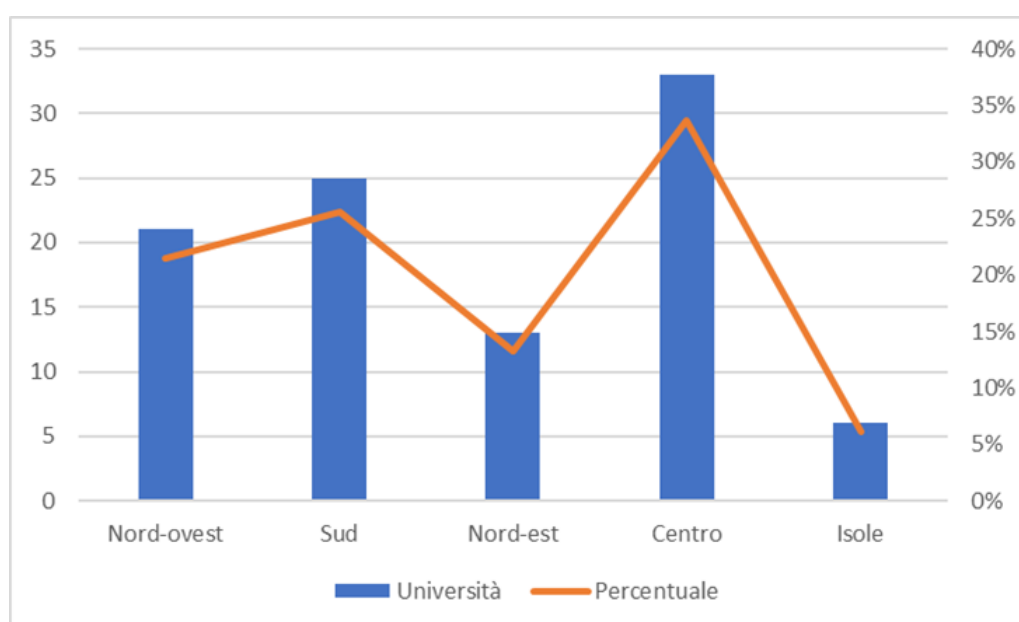


Figura 30 – Suddivisione delle università del database per macro regione

Per poter rappresentare tale suddivisione nel modello è stato necessario generare per ogni macro regione una variabile binaria che ne indicasse l'appartenenza. Per evitare la trappola della variabile *dummy* e, di conseguenza, il problema della collinearità perfetta, in quanto variabili mutuamente esclusive ed esaustive, si è deciso di omettere uno dei gruppi dal modello. In particolare è stato scelto il “Centro” poiché rappresenta la macro regione maggiormente numerosa e, per tal motivo, quella più adatta ad essere impiegata come termine di paragone per stabilire se sussiste un divario significativo con gli enti che appartengono ad un'altra zona geografica.

Prima di passare all'applicazione dei metodi di stima sui dati a disposizione, avendo impostato un modello di analisi multivariato, è stato controllato che non vi fosse collinearità imperfetta tra i regressori specificati. Questo fenomeno si verifica quando due o più variabili indipendenti sono altamente correlate tra di loro. Per tale motivo è stata costruita una matrice di correlazione e i valori ottenuti sono mostrati in Tabella 5.

Tabella 5 – Matrice di correlazione tra i regressori impiegati nel modello di regressione

Correlation Matrix							
Variables	rca_d	Tot	HHI	Isole_d	Nord-Est_d	Nord-Ovest_d	Sud_d
rca_d	1						
Tot	-0.178***	1					
HHI	0.190***	-0.148***	1				
Isole_d	-0.0862***	-0.0142	-0.168***	1			
Nord-Est_d	-0.0556***	0.0291***	-0.0613***	-0.149***	1		
Nord-Ovest_d	-0.0445***	0.0509***	0.0446***	-0.159***	-0.268***	1	
Sud_d	0.0569***	-0.0558***	-0.0144	-0.160***	-0.270***	-0.288***	1

p < 0.05, ** *p* < 0.01, *** *p* < 0.001

Per identificare se vi sia il rischio che uno o più dei coefficienti di regressione siano stimati in maniera imprecisa, si è deciso di impostare come soglia di sicurezza:

$$|\rho_{AB}| \geq 0,35$$

Come si può evincere dalla Tabella 5 nessuna coppia di regressori presenta un coefficiente di correlazione maggiore della soglia e, di conseguenza, non sono stati necessari ulteriori aggiustamenti al modello realizzato. Si è potuto, quindi, procedere con il calcolo dei coefficienti e l'esposizione dei risultati ottenuti.

3.3. I risultati del modello

L'introduzione delle variabili binarie *rsa_d* e *rca_d*, che hanno sostituito rispettivamente l'indice di specializzazione *RSA* e l'indice di collaborazione *RCA*, ha permesso di mitigare l'effetto di distorsione che si manifesta nel calcolo di quest'ultimi per quelle università che presentano un ridotto volume di pubblicazioni. Tuttavia, seppur attenuato, tale effetto continua a possedere un ruolo nel determinare le caratteristiche di un istituto. Per poter valutare in maniera più appropriata la relazione tra specializzazione e collaborazione e stabilire se tali alterazioni posso influire sui risultati dello studio, si è optato di applicare il modello di regressione su due differenti campioni: il primo

considerando tutte le informazioni a disposizione nel database di partenza, il secondo, invece, considerando per ciascuna università soltanto le aree di ricerca che sono state argomento di pubblicazione di almeno 10 articoli scientifici. Tale scelta è stata attuata in quanto si è ritenuto che l'aver condotto, nel periodo compreso tra il 2015 e il 2020, 10 studi in un determinato ambito rappresenti la soglia minima per poterla considerare un'attività continuativa nel tempo ed una quantità adeguata che possa fornire indicazioni su un ente in merito all'orientamento alla cooperazione o all'ambito di specializzazione senza provocare alterazioni eccessive del valore degli indici. Per distinguere le due specificazioni, le variabili interessate da tale cambiamento sono state indicate con *rsa_d* e *rca_d* per il primo gruppo, mentre con *rsa10_d* e *rca10_d* per il secondo. Infine, per rendere quanto più rigorosa possibile l'analisi, si è prestata attenzione a modificare le impostazioni per il calcolo degli errori standard in modo tale che fossero robusti all'eteroschedasticità, evitando quindi di ottenere dei valori più piccoli che porterebbero ad una statistica *t* più elevata con eventuale rifiuto errato dell'ipotesi nulla.

Nella seguente tabella vengono mostrati i risultati ottenuti.

Tabella 6 – Risultati analisi di regressione per i gruppi specificati

VARIABLES	(1) rsa_d	(2) rsa_d	(3) rsa10_d	(4) rsa10_d
rca_d	-0.00877 (0.0390)	-0.0214 (0.0391)		
Tot	0.00682*** (0.000425)	0.00685*** (0.000428)	0.00288*** (0.000252)	0.00299*** (0.000259)
HHI	16.76*** (1.079)	16.51*** (1.103)	56.62*** (3.770)	57.54*** (3.882)
Isole_d		-0.119 (0.0732)		0.0703 (0.0965)
Nord-Est_d		-0.133** (0.0545)		-0.0819 (0.0717)
Nord-Ovest_d		-0.0820 (0.0514)		-0.235*** (0.0712)
Sud_d		0.0223 (0.0507)		0.182*** (0.0698)
rca10_d			-0.101** (0.0514)	-0.128** (0.0521)
Constant	-1.610*** (0.0427)	-1.551*** (0.0535)	-2.154*** (0.0785)	-2.145*** (0.0929)
Observations	16,152	16,152	9,073	9,073
log pseudolikelihood	-9168	-9163	-5089	-5071
Prob > chi2	0	0	0	0
Pseudo R2	0.0697	0.0703	0.0888	0.0920

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

A conferma di quanto è stato evidenziato nel secondo capitolo, anche mediante l'impiego dell'analisi di regressione è stato possibile individuare una relazione negativa tra l'aspetto della specializzazione e della collaborazione universitaria. Infatti il coefficiente di segno negativo associato alle variabili binarie *rca_d* e *rca10_d* nei diversi modelli indica che se un'università, in una determinata area di ricerca, è considerata maggiormente cooperativa rispetto alla media, allora la probabilità che sia anche sovra specializzata in essa si riduce. Tuttavia, nelle prime due specificazioni, ovvero quelle in cui sono stati considerati tutti i dati a disposizione, i valori ottenuti non sono statisticamente significativi, mentre soltanto nelle due rimanenti acquisiscono un livello di significatività del 5%. L'aver, quindi, selezionato per ciascuna università solamente le aree tematiche per le quali, nel corso degli anni compresi tra il 2015 e il 2020, è rilevabile una discreta continuità nella produzione scientifica, ha permesso di ottenere dei risultati che da un punto di vista statistico sono accettabili. Tuttavia, bisogna considerare che tale scelta ha comportato una drastica riduzione nel numero di osservazioni disponibili per l'applicazione dell'analisi, che sono passate da oltre 16.000 a circa 9.000. Tale diminuzione ha provocato la perdita di molte delle informazioni presenti nel *database* soprattutto per quanto riguarda le università più piccole. Di conseguenza, data l'eliminazione di parte della variabilità apportata al campione da quest'ultime, i risultati dell'analisi di regressione sono stati principalmente influenzati dal comportamento degli istituti più importanti. Si può quindi affermare con un discreto grado di sicurezza che tale relazione negativa tra specializzazione e collaborazione è principalmente osservabile nelle università con un maggior volume di pubblicazioni.

Considerando invece le variabili *Tot* e *HHI*, entrambe possiedono in ciascuno dei modelli adoperati un effetto positivo e statisticamente significativo sulla probabilità di essere specializzati in una materia. Com'era lecito prevedere, più la produzione scientifica di un ente rispetto ad uno specifico ambito di ricerca è cospicua, maggiore è la possibilità che quest'ultimo diventi sovra specializzato in quel campo. In maniera analoga, più elevato è l'indice di concentrazione associato ad un ente, minore è il numero di aree tematiche su cui condensa la pubblicazione di articoli e di conseguenza maggiore è la probabilità che risulti specializzato in una di esse.

Infine, per quando riguarda l'aspetto geografico, nel secondo modello soltanto le università localizzate nel Nord-Est d'Italia si discostano significativamente da quelle del Centro, presentando una minore tendenza ad essere specializzate su un determinato

argomento. Piuttosto differente è, invece, la situazione presentata nel quarto modello dove sono gli istituti del Nord Ovest e del Sud d'Italia che si differenziano in maniera statisticamente rilevante da quelli del Centro: i primi presentando una minore tendenza a specializzarsi, mentre i secondi una maggiore. Soltanto per gli enti che risiedono sulle due Isole le discrepanze risultano non significative in entrambe le specificazioni.

In conclusione, si può constatare che le supposizioni avanzate nel corso del precedente capitolo hanno trovato un riscontro anche attraverso l'analisi di regressione. Infatti, come si era avanzata l'ipotesi nel secondo capitolo che la specializzazione e la collaborazione nella ricerca scientifica italiana fossero legate da una relazione negativa, allo stesso modo il segno meno davanti al coefficiente associato alla variabile *rca_d* denota una relazione analoga. Tuttavia, un problema riscontrato è che non sempre i valori ottenuti dalla regressione sono risultati statisticamente significativi. C'è stato bisogno di operare una notevole restrizione del campione per poter ottenere dei valori significativi. Per tale motivo appare complicato asserire che nel panorama scientifico italiano, in particolare per quanto riguarda il contesto della produzione di articoli scientifici tra il 2015 e 2020, vi sia stata una chiara tendenza da parte delle università a collaborare con altri istituti in quelle materie dove si è registrata una carenza di pubblicazioni e, quindi, di specializzazione. È stata identificata una piccola evidenza che, però, necessita di ulteriori investigazioni ed analisi per essere considerata rappresentativa. I risultati raggiunti dal presente lavoro possono essere considerati un punto di partenza per studi futuri che vadano ad esplorare in maniera più strutturata e approfondita tale relazione in un contesto come quello italiano, ma non solo.

Conclusioni

Le università rappresentano il luogo di formazione e accoglienza di ricercatori e scienziati e, per tal motivo, assumono un ruolo centrale nello sviluppo della scienza e della tecnologia. Per favorire tale aspetto, un fattore chiave che ha caratterizzato particolarmente l'attività di ricerca negli ultimi decenni è stata la collaborazione. La condivisione di idee, conoscenze ed esperienze, ha favorito e velocizzato il progresso e l'espansione della ricerca scientifica nel mondo. Mentre la maggior parte della letteratura si è concentrata sul rapporto di collaborazione tra università e industria, l'obiettivo del presente elaborato è stato quello di esplorare l'aspetto della collaborazione interuniversitaria nell'ambito della ricerca scientifica e di come tale elemento si possa relazionare all'ambito e al grado di specializzazione di ciascuna università. Per facilitare tale operazione, si è optato di restringere il campo di studio al panorama scientifico italiano nel periodo intercorrente tra il 2015 e il 2020. A partire dalle informazioni presenti su *Web of Science*, una delle più grandi banche dati bibliografiche citazionali, per quanto riguarda la tipologia documentale dell'articolo scientifico, si è tentato di indagare la relazione tra collaborazione e specializzazione tramite un approccio scientifico e l'impiego di strumenti derivati dall'analisi statistica ed econometrica. I risultati di tale studio hanno evidenziato che a livello generale, quindi, considerando contemporaneamente i valori ottenuti per differenti ambiti di ricerca, sussiste una sorta di relazione negativa tra questi due fattori. Infatti, è stata individuata una leggera tendenza da parte delle università italiane a collaborare con altri istituti quando il loro livello di specializzazione su un certo argomento fosse modesto.

Tuttavia, come è stato sottolineato nel corso della trattazione, gli esiti ottenuti non sempre potevano essere considerati statisticamente significativi e il loro raggiungimento poteva essere influenzato, se non perfino alterato, dalle metriche che sono state impiegate. Inoltre, la generalizzazione non ha permesso di identificare e confrontare in maniera struttura le differenze esistenti tra le varie aree di ricerca e, quindi, le dinamiche riscontrate risultano non essere valide per determinati ambiti. Di conseguenza, molte sono le limitazioni del presente studio che potrebbero essere indagate e affrontate in lavori futuri. Come prima azione, si potrebbe ampliare la base di dati sulla quale poter condurre l'analisi andando a considerare anche altre tipologie documentali utilizzate per la pubblicazione delle attività di ricerca. Infatti, l'aver ristretto il campo di indagine alla sola

tipologia degli articoli scientifici, ha indirizzato l'analisi verso quei settori relativi all'ambito ingegneristico e delle scienze naturali, che prediligono questa struttura di comunicazione, a discapito di altri, come quelli di carattere sociale e umanistico, che si orientano verso altre forme di divulgazione. Di conseguenza, i risultati così ricavati hanno risentito maggiormente dalle tendenze e peculiarità relative a quei settori scientifici. Una differente maniera per aumentare il numero di osservazioni a disposizione potrebbe essere quella di ampliare l'intervallo temporale considerato. Tale scelta permetterebbe, in aggiunta, di conferire maggior rilievo alla variabile tempo rispetto a quanto è stato effettuato nel presente lavoro. Per tale motivo sarebbe possibile compiere anche un'analisi dinamica che consenta di studiare l'evoluzione negli anni della relazione esistente tra specializzazione e collaborazione. Come secondo aspetto, si potrebbero sviluppare e implementare delle metriche differenti che possano meglio qualificare e definire i fattori oggetto dell'analisi: specializzazione e collaborazione. Gli indici impiegati nel presente lavoro, infatti, soffrono di alcune problematiche strutturali che ne alterano i valori in presenza di combinazioni tra università e aree di ricerca con un modesto numero di pubblicazioni e che, quindi, potrebbero causare delle distorsioni dei risultati finali. Un terzo punto, potrebbe essere quello di sviluppare e migliorare l'analisi di regressione. Si potrebbero identificare nuovi fattori da inserire nel modello che possano determinare o influenzare in qualche maniera i due aspetti della specializzazione e della collaborazione e che permettano di ridurre le distorsioni delle stime derivati dai problemi connessi alle variabili omesse. Infine, come ultimo suggerimento di cui viene fatto menzione, si potrebbe optare per investigare in maniera più approfondita la situazione presente all'interno dei vari ambiti scientifici, andando a studiarli singolarmente, e solo in un secondo momento decidere di sommare i risultati ottenuti per poter trarre delle conclusioni a livello generale. L'aver scelto, come avvenuto nel presente lavoro, di non prestare particolare attenzione alle distinzioni esistenti nelle differenti aree di ricerca, soprattutto durante l'applicazione dell'analisi di regressione, ha causato, oltre ad una perdita di variabilità, anche un effetto compensativo dove le tendenze di un particolare settore verso una certa forma di relazione tra specializzazione e collaborazione potevano essere bilanciate da tendenze opposte presenti in altri settori.

Appendice

Macro-categoria	Aree di ricerca
<i>Arts & Humanities</i>	"ARCHAEOLOGY", "ARCHITECTURE", "ART", "HUMANITIES, MULTIDISCIPLINARY", "ASIAN STUDIES", "CLASSICS", "DANCE", "FILM, RADIO, TELEVISION", "FOLKLORE", "HISTORY", "LANGUAGE & LINGUISTICS", "LITERATURE", "LITERARY REVIEWS", "LITERARY THEORY & CRITICISM", "LITERATURE, AFRICAN, AUSTRALIAN, CANADIAN", "LITERATURE, AMERICAN", "LITERATURE, BRITISH ISLES", "LITERATURE, GERMAN, DUTCH, SCANDINAVIAN", "LITERATURE, ROMANCE", "LITERATURE, SLAVIC", "MEDIEVAL & RENAISSANCE STUDIES", "MUSIC", "PHILOSOPHY", "POETRY", "RELIGION", "THEATER"
<i>Biology</i>	"AGRICULTURE, MULTIDISCIPLINARY", "AGRICULTURAL ECONOMICS & POLICY", "AGRICULTURAL ENGINEERING", "AGRICULTURE, DAIRY & ANIMAL SCIENCE", "AGRONOMY", "BIOCHEMISTRY & MOLECULAR BIOLOGY", "BIODIVERSITY CONSERVATION", "BIOCHEMICAL RESEARCH METHODS", "BIOLOGY", "BIOPHYSICS", "BIOTECHNOLOGY & APPLIED MICROBIOLOGY", "CELL BIOLOGY", "CELL & TISSUE ENGINEERING", "DEVELOPMENTAL BIOLOGY", "ENTOMOLOGY", "ECOLOGY", "EVOLUTIONARY BIOLOGY", "FISHERIES", "FOOD SCIENCE & TECHNOLOGY", "FORESTRY", "HORTICULTURE", "MARINE & FRESHWATER BIOLOGY", "MATHEMATICAL & COMPUTATIONAL BIOLOGY", "MICROBIOLOGY", "MYCOLOGY", "ORNITHOLOGY", "PLANT SCIENCES", "REPRODUCTIVE BIOLOGY", "SOIL SCIENCE", "VETERINARY SCIENCES", "ZOOLOGY"
<i>Biomedical Research</i>	"ALLERGY", "ANATOMY & MORPHOLOGY", "CHEMISTRY, MEDICINAL", "HEMATOLOGY", "IMMUNOLOGY", "INFECTIOUS DISEASES", "MEDICAL LABORATORY TECHNOLOGY", "MEDICINE, RESEARCH & EXPERIMENTAL", "ONCOLOGY", "PATHOLOGY", "PHARMACOLOGY & PHARMACY", "RADIOLOGY, NUCLEAR MEDICINE & MEDICAL IMAGING", "TOXICOLOGY", "VIROLOGY"
<i>Chemistry</i>	"CHEMISTRY", "CHEMISTRY, ANALYTICAL", "CHEMISTRY, APPLIED", "CHEMISTRY, INORGANIC & NUCLEAR", "CHEMISTRY, MEDICINAL", "CHEMISTRY, MULTIDISCIPLINARY", "CHEMISTRY, ORGANIC", "CHEMISTRY, PHYSICAL", "ELECTROCHEMISTRY", "POLYMER SCIENCE"
<i>Clinical Medicine</i>	"AGRONOMY", "ALLERGY", "ANATOMY & MORPHOLOGY", "ANDROLOGY", "ANESTHESIOLOGY", "ANTHROPOLOGY", "AUDIOLOGY & SPEECH-LANGUAGE PATHOLOGY", "BEHAVIORAL SCIENCES", "CARDIAC & CARDIOVASCULAR SYSTEMS", "CLINICAL NEUROLOGY", "CRITICAL CARE MEDICINE", "DENTISTRY, ORAL SURGERY & MEDICINE", "DERMATOLOGY", "EMERGENCY MEDICINE", "ENDOCRINOLOGY & METABOLISM", "ERGONOMICS", "GASTROENTEROLOGY & HEPATOLOGY", "MEDICINE, GENERAL & INTERNAL", "GENETICS & HEREDITY", "GERIATRICS & GERONTOLOGY", "HEALTH CARE SCIENCES & SERVICES", "INTEGRATIVE & COMPLEMENTARY MEDICINE", "MEDICINE, LEGAL", "NEUROIMAGING",

	"NEUROSCIENCES", "NUTRITION & DIETETICS", "OBSTETRICS & GYNECOLOGY", "OPHTHALMOLOGY", "ORTHOPEDICS", "OTORHINOLARYNGOLOGY", "PARASITOLOGY", "PEDIATRICS", "PERIPHERAL VASCULAR DISEASE", "PHYSIOLOGY", "PRIMARY HEALTH CARE", "PSYCHIATRY", "PUBLIC, ENVIRONMENTAL & OCCUPATIONAL HEALTH", "REHABILITATION", "RESPIRATORY SYSTEM", "RHEUMATOLOGY", "SPORT SCIENCES", "SUBSTANCE ABUSE", "SURGERY", "TRANSPLANTATION", "TROPICAL MEDICINE", "UROLOGY & NEPHROLOGY"
<i>Earth & Space Sciences</i>	"ENVIRONMENTAL SCIENCES", "ENVIRONMENTAL STUDIES", "GEOCHEMISTRY & GEOPHYSICS", "GEOLOGY", "GEOGRAPHY, PHYSICAL", "GEOSCIENCES, MULTIDISCIPLINARY", "LIMNOLOGY", "METEOROLOGY & ATMOSPHERIC SCIENCES", "MINERALOGY", "MINING & MINERAL PROCESSING", "OCEANOGRAPHY", "PALEONTOLOGY", "WATER RESOURCES"
<i>Economics</i>	"BUSINESS", "BUSINESS, FINANCE", "ECONOMICS", "INFORMATION SCIENCE & LIBRARY SCIENCE", "MANAGEMENT", "PLANNING & DEVELOPMENT", "PUBLIC ADMINISTRATION", "TRANSPORTATION"
<i>Engineering</i>	"AUTOMATION & CONTROL SYSTEMS", "COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS", "COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE", "COMPUTER SCIENCE, CYBERNETICS", "COMPUTER SCIENCE, HARDWARE & ARCHITECTURE", "COMPUTER SCIENCE, INFORMATION SYSTEMS", "COMPUTER SCIENCE, SOFTWARE ENGINEERING", "COMPUTER SCIENCE, THEORY & METHODS", "CONSTRUCTION & BUILDING TECHNOLOGY", "ENGINEERING, AEROSPACE", "ENGINEERING, BIOMEDICAL", "ENGINEERING, CHEMICAL", "ENGINEERING, CIVIL", "ENGINEERING, ELECTRICAL & ELECTRONIC", "ENGINEERING, ENVIRONMENTAL", "ENGINEERING, GEOLOGICAL", "ENGINEERING, INDUSTRIAL", "ENGINEERING, MANUFACTURING", "ENGINEERING, MARINE", "GREEN & SUSTAINABLE SCIENCE & TECHNOLOGY", "ENGINEERING, MECHANICAL", "ENGINEERING, MULTIDISCIPLINARY", "ENGINEERING, OCEAN", "ENGINEERING, PETROLEUM", "INSTRUMENTS & INSTRUMENTATION", "MATERIALS SCIENCE, BIOMATERIALS", "MATERIALS SCIENCE, CERAMICS", "MATERIALS SCIENCE, CHARACTERIZATION & TESTING", "MATERIALS SCIENCE, COATINGS & FILMS", "MATERIALS SCIENCE, COMPOSITES", "MATERIALS SCIENCE, MULTIDISCIPLINARY", "MATERIALS SCIENCE, PAPER & WOOD", "MATERIALS SCIENCE, TEXTILES", "MEDICAL INFORMATICS", "METALLURGY & METALLURGICAL ENGINEERING", "NANOSCIENCE & NANOTECHNOLOGY", "NUCLEAR SCIENCE & TECHNOLOGY", "QUANTUM SCIENCE & TECHNOLOGY", "REMOTE SENSING", "ROBOTICS", "TELECOMMUNICATIONS", "TRANSPORTATION SCIENCE & TECHNOLOGY"
<i>Law, Political & Social Sciences</i>	"ANTHROPOLOGY", "AREA STUDIES", "COMMUNICATION", "CRIMINOLOGY & PENOLOGY", "CULTURAL STUDIES", "DEMOGRAPHY", "DEVELOPMENT STUDIES", "EDUCATION & EDUCATIONAL RESEARCH", "EDUCATION, SCIENTIFIC DISCIPLINES", "EDUCATION, SPECIAL", "ETHNIC STUDIES", "ETHICS", "HEALTH POLICY & SERVICES", "FAMILY STUDIES", "GEOGRAPHY", "GERONTOLOGY", "HOSPITALITY, LEISURE, SPORT &

	TOURISM", "INDUSTRIAL RELATIONS & LABOR", "LAW", "HISTORY OF SOCIAL SCIENCES", "INTERNATIONAL RELATIONS", "LINGUISTICS", "MEDICAL ETHICS", "NURSING", "REGIONAL & URBAN PLANNING", "SOCIAL SCIENCES, BIOMEDICAL", "SOCIAL SCIENCES, MATHEMATICAL METHODS", "POLITICAL SCIENCE", "SOCIAL ISSUES", "SOCIAL SCIENCES, INTERDISCIPLINARY", "SOCIAL WORK", "SOCIOLOGY", "URBAN STUDIES", "Women's Studies"
<i>Mathematics</i>	"LOGIC", "MATHEMATICS", "MATHEMATICS, APPLIED", "MATHEMATICS, INTERDISCIPLINARY APPLICATIONS", "OPERATIONS RESEARCH & MANAGEMENT SCIENCE", "STATISTICS & PROBABILITY"
<i>Multidisciplinary Sciences</i>	"EDUCATION, SCIENTIFIC DISCIPLINES", "HISTORY & PHILOSOPHY OF SCIENCE", "MULTIDISCIPLINARY SCIENCES"
<i>Physics</i>	"ACOUSTICS", "ASTRONOMY & ASTROPHYSICS", "CRYSTALLOGRAPHY", "ENERGY & FUELS", "IMAGING SCIENCE & PHOTOGRAPHIC TECHNOLOGY", "MECHANICS", "MICROSCOPY", "OPTICS", "PHYSICS, APPLIED", "PHYSICS, ATOMIC, MOLECULAR & CHEMICAL", "PHYSICS, CONDENSED MATTER", "PHYSICS, FLUIDS & PLASMAS", "PHYSICS, MATHEMATICAL", "PHYSICS, MULTIDISCIPLINARY", "PHYSICS, NUCLEAR", "PHYSICS, PARTICLES & FIELDS", "SPECTROSCOPY", "THERMODYNAMICS"
<i>Psychology</i>	"PSYCHOLOGY", "PSYCHOLOGY, APPLIED", "PSYCHOLOGY, BIOLOGICAL", "PSYCHOLOGY, CLINICAL", "PSYCHOLOGY, DEVELOPMENTAL", "PSYCHOLOGY, EDUCATIONAL", "PSYCHOLOGY, EXPERIMENTAL", "PSYCHOLOGY, MATHEMATICAL", "PSYCHOLOGY, MULTIDISCIPLINARY", "PSYCHOLOGY, PSYCHOANALYSIS", "PSYCHOLOGY, SOCIAL"

Bibliografia

- [1] – P. Bisogno, *Ricerca scientifica*, in https://www.treccani.it/enciclopedia/ricerca-scientifica_%28Enciclopedia-Italiana%29/
- [2] – The World Bank, *Spesa per ricerca e sviluppo (% del PIL) - Stati Uniti, Regno Unito, Unione Europea, Giappone, Cina*, in <https://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS?locations=US-GB-EU-JP-CN>
- [3] – OECD.Stat, *Main Science and Technology Indicators database*, in <https://stats.oecd.org>
- [4] – Commissione Europea, *EUROPA 2020 Una strategia per una crescita intelligente, sostenibile e inclusiva*, in http://publications.europa.eu/resource/cellar/6a915e39-0aab-491c-8881-147ec91fe88a.0008.02/DOC_1
- [5] – Openpolis, *Ricerca e sviluppo: quanto investono l'Italia e i paesi Ue in questo settore*, in <https://www.openpolis.it/ricerca-e-sviluppo-quanto-investono-litalia-e-i-paesi-ue-in-questo-settore/>
- [6] – Ministero dell'Università e della Ricerca, *Programma nazionale per la ricerca*, in <https://www.mur.gov.it/it/aree-tematiche/ricerca/programmazione/programma-nazionale-la-ricerca>
- [7] – Ministero dell'Istruzione, dell'Università e della Ricerca, *Programma Nazionale della Ricerca 2015-2020*, in <https://www.miur.gov.it/documents/20182/71637/Il+Programma.pdf/2faf018a-b63d-454b-96a4-6b55fd3a6955?version=1.0&t=1495038747623>
- [8] – Ministero dell'Istruzione, dell'Università e della Ricerca, *Programma Nazionale della Ricerca 2021-2027*, in <https://www.mur.gov.it/sites/default/files/2021-01/Pnr2021-27.pdf>
- [9] – F. M. Scheter, *Innovation and Growth: Schumpeterian Perspectives*, MIT Press Book, 1986
- [10] – P. Aghion, P. Howitt, *Technical Progress in the Theory of Economic Growth*, in «Economics in a Changing World», 1995, Volume 5, pp. 101-122

- [11] – L. Aliu Mulaj, B. Dedaj, *Knowledge-Based Society: R&D Investments in New Economic Transformation*, in «Accounting, Finance, Sustainability, Governance and Fraud», 2023, Volume 4, pp. 49-67
- [12] – M. Coccia, *Science, Founding and economic growth: Analysis and science policy implications*, in «World Review of Science, Technology and Sustainable Development», 2008, Volume 5, pp. 1-27
- [13] – J. R. Brown, *Venture Capital and Firm Performance Over the Long-Run: Evidence from HighTech IPOs in the United States*, in «Journal of Entrepreneurial Finance and Business Ventures», 2005, Volume 10, pp. 1-33
- [14] – M. Draghi, *Dalla ricerca all'innovazione per la crescita economica*, in https://www.bancaditalia.it/pubblicazioni/interventi-governatore/integov2007/draghi_24_07_07.pdf
- [15] – O. Kolomytseva, A. Pavlovska, *The Role Of Universities In The National Innovation System*, in «Baltic Journal of Economic Studies», 2020, Volume 6
- [16] - B. Godin, Y. Gingras, *The place of universities in the system of knowledge production*, in «Research Policy», 2000, Volume 29, pp. 273-278
- [17] – G. W. Prigge, *University–industry partnerships: what do they mean to universities?: a Review of the Literature*, in «Industry and Higher Education», 2005, Volume 19, pp. 221-229
- [18] – L. Hou, Y. Pan, J. J. H. Zhu, *Impact of scientific, economic, geopolitical, and cultural factors on international research collaboration*, in «Journal of Informetrics», 2021, Volume 15, Articolo n. 101194
- [19] – R. C. Leung, *Networks as sponges: International collaboration for developing nanomedicine in China*, in «Research Policy», 2013, Volume 42, pp. 211–219
- [20] – X. Kong, M. Mao, H. Jiang, S. Yu, L. Wan, *How does collaboration affect researchers' positions in co-authorship networks?*, in «Journal of Informetrics», 2019, Volume 13, pp. 887–900
- [21] – S. C. J. Sin, *International coauthorship and citation impact: A bibliometric study of six LIS journals, 1980–2008*, in «Journal of the American Society for Information Science and Technology», 2011, Volume 62, pp. 1770–1783

- [22] – S. Wuchty, B. F. Jones, B. Uzzi, *The increasing dominance of teams in production of knowledge*, in «Science», 2007, Volume 316, pp. 1036–1039
- [23] – V. Larivière, Y. Gingras, C. R. Sugimoto, A. Tsou, *Team size matters: Collaboration and scientific impact since 1900*, in «Journal of the Association for Information Science and Technology», 2014, Volume 66, pp. 1323-1332
- [24] – M. E. Newman, *Coauthorship networks and patterns of scientific collaboration*, in «Proceedings of the National Academy of Sciences of the United States of America», 2004, Volume 101, pp. 5200-5205
- [25] – L. Wu, D. Wang, J. A. Evans, *Large teams develop and small teams disrupt science and technology*, in «Nature», 2019, Volume 566, pp. 378–382
- [26] – F. Franceschini, D. Maisano, L. Mastrogiacomo, *Evaluating research institutions: The potential of the success-index*, in «Scientometrics», 2013, Volume 96, pp. 85–101
- [27] – P. Taylor, R. Braddock, *International university ranking systems and the idea of university excellence*, in «Journal of Higher Education Policy and Management», 2007, Volume 29, pp. 245–260
- [28] – H. Etzkowitz, L. Leydesdorff, *The Triple Helix, University-Industry-Government Relations: A laboratory for Knowledge-Based Economic Development*, in «EASST Review», 1995, Volume 14, pp. 14-19
- [29] – H. Etzkowitz, L. Leydesdorff, *The dynamics of innovation: from National Systems and “Mode 2” to a Triple Helix of university–industry–government relations*, in «Research Policy», 2000, Volume 29, pp. 109–123
- [30] – E. W. Mainardes, H. Alves, M. Raposo, *The process of change in university management: from the “ivory tower” to entrepreneurship*, in «Transylvanian Review of Administrative Sciences», 2011, Volume 7, pp. 124–149
- [31] – B. Godin, Y. Gingras, *The place of universities in the system of knowledge production*, in «Research Policy», 2000, Volume 29, pp. 273–278
- [32] – K. Miller, M. McAdam, R. McAdam, *The changing university business model: a stakeholder perspective*, in «R&D Management», 2014, Volume 44, pp. 265–287

- [33] – I. M. B. Freitas, N. Von Tunzelmann, *Mapping public support for innovation: a comparison of policy alignment in the UK and France*, in «Research Policy», 2008, Volume 37, pp. 1446–1464
- [34] – C. Champenois, H. Etzkowitz, *From boundary line to boundary space: the creation of hybrid organizations as a Triple Helix micro-foundation*, in «Technovation», 2018, Volume 76, pp. 28–39
- [35] – A. Albertin, *Parchi scientifici e tecnologici, a cosa servono e cosa fanno*, in <https://www.teknoring.com/news/internet-of-things/parchi-scientifici-e-tecnologici-a-cosa-servono-e-cosa-fanno/#:~:text=I%20parchi%20scientifici%20e%20tecnologici%20sono%20aggregatori%20di,le%20regioni%2C%20con%20la%20Lombardia%20capofila%20con%20sei.>
- [36] – F. Nitti, *Le università: centri di innovazione e sviluppo!*, in <https://www.minnovo.it/universita-innovazione-sviluppo/>
- [37] – H. Meyer, *UBI Global World Rankings of Business Incubators and Accelerators 2019-2020*, UBI Global, 2019
- [38] – E. G. Carayannis, D. F. Campbell, *'Mode 3' and 'Quadruple Helix': toward a 21st century fractal innovation ecosystem*, in «International Journal of Technology Management», 2019, Volume 46, pp. 201–234
- [39] – Clarivate, *Web of Science*, in <https://clarivate.com/webofsciencegroup/solutions/web-of-science/>
- [40] – H.F. Moed, *Citation analysis in research evaluation*, in «10th Biennial International Conference of the International Society for Scientometrics and Informetrics», 2005, Volume 2, pp. 437-441
- [41] – G. Abramo, C. A. D'Angelo, *The relationship between the number of authors of a publication, its citations and the impact factor of the publishing journal: Evidence from Italy*, in «Journal of Informetrics», 2015, Volume 9, pp. 746-761
- [42] – Conversable Economist, *The Herfindahl-Hirschman Index: Story, Primer, Alternatives*, in <https://conversableeconomist.blogspot.com/2020/02/the-herfindahl-hirschman-index-story.html>
- [43] – Investopedia, *Definizione, formula ed esempio dell'indice di Herfindahl-Hirschman (HHI)*, in <https://www.investopedia.com/terms/h/hhi.asp>

[44] – F. Caviggioli, P. Jensen, G. Scellato, *Highly skilled migrants and technological diversification in the US and Europe*, in «Technological Forecasting and Social Change», 2020, Volume 154, Articolo n. 119951

[45] – P. A. Balland, R. Boschma, *Do scientific capabilities in specific domains matter for technological diversification in European regions?*, in «Research Policy», 2022, Volume 51, Articolo n. 104594

[46] – Readitaliano, *Indice di Jaccard*, in https://readitaliano.com/wiki/it/Jaccard_index