

“Disseminazione, archiviazione e riuso dei dati: a che punto siamo?”

La *survey analysis*: dalla genesi del questionario alle statistiche finali.

Il presente lavoro di tesi presenta, nella sua parte più consistente, una serie di analisi effettuate sui risultati dell'indagine “Disseminazione, archiviazione e riuso dei dati: a che punto siamo?” per capire in che modo la condivisione dei dati viene percepita da chi, con i dati, ci lavora e quali sono i possibili ostacoli ad una piena apertura verso il *data sharing*, con particolare attenzione alle differenze tra gli intervistati italiani e quelli stranieri.

Il primo capitolo consiste in una dissertazione puramente teorica sulle metodologie di realizzazione di una indagine condotta a scopo di ricerca, a partire dalla definizione degli obiettivi, la realizzazione del questionario e la raccolta dei dati sino ad arrivare alle analisi e al report finale.

Il secondo capitolo riguarda l'indagine oggetto della tesi, condotta da Metis Ricerche su di un campione che, dopo le opportune fasi di pulizia dei dati, è stato ridotto a 81 soggetti in base ad un indice di bontà dell'intervista relativo al numero di risposte fornite. Tutte le fasi sono state dettagliatamente descritte, in particolar modo il processo di data quality, effettuato tramite il software SAS Base versione 9.4. Largo spazio è stato attribuito anche alle analisi statistiche: dapprima quelle monovariate – relativamente ad alcune caratteristiche sociometriche e alle domande più rilevanti – e successivamente quelle bivariate, con particolare attenzione sulle principali domande in batteria messe in relazione con la provenienza (Italia/estero) dei soggetti intervistati. È stato poi effettuato il test d'indipendenza del chi-quadrato sulle quattro batterie più importanti per stabilire se le risposte a tali domande dipendessero – e, se sì, a quale livello di significatività – dalla provenienza dei soggetti intervistati. Infine mediante un test di correlazione e la verifica del coefficiente α si è voluta quantificare la coerenza interna delle batterie. Queste ultime fasi sono corredate di grafici e tabelle, prodotti con SAS Base.

La terza ed ultima parte consiste in un confronto diacronico dei risultati dell'indagine relativamente alle batterie principali con quelli corrispondenti di una precedente ricerca effettuata in due fasi – nel 2009/2010 e nel 2013/2014 – nell'ambito del progetto DataONE, supportato dalla National Science Foundation e nato per la creazione di una infrastruttura informatica che permetta un accesso globale ma sicuro ai dati provenienti da diverse ricerche scientifiche.