

Nuovi parametri nelle analisi statistiche di accesso alle risorse elettroniche digitali

Domenico Dellisanti, Antonio Balducci

CILEA, Segrate

Abstract

Nell'ambito del consueto appuntamento annuale con Bibliocom, rassegna delle professioni, dei prodotti e dei servizi per la gestione dell'informazione e della conoscenza, svoltosi a Roma dal 27 al 29 ottobre, si è tenuto un seminario organizzato dal CASPUR, "La biblioteca digitale: statistiche d'uso. Studi ed esperienze a confronto". Il CILEA, all'interno dello stesso, ha presentato uno studio su nuovi parametri da utilizzarsi nelle analisi statistiche di accesso alle risorse elettroniche.

Keywords: Biblioteca digitale, risorse elettroniche, statistiche di accesso, parametri, punto di equilibrio.

Il CILEA offre mensilmente ai suoi sottoscrittori la possibilità di consultare le statistiche di accesso ai vari servizi d'editoria elettronica. Queste sono fruibili tramite username/password e, vengono messe in linea dalla pagina:

<http://www.cilea.it/cdl/report.htm>.

Qui è possibile visionare quelle relative a diversi editori tra i quali segnaliamo Elsevier, Kluwer, ACS, OUP, ISI, Springer, o a servizi, forniti dal CILEA, come SFX.

In questo studio l'attenzione si è focalizzata sulle statistiche relative ad Elsevier. Mensilmente il log degli accessi alle riviste Elsevier, messe in linea su di un proprio server (**mirror SDOS ScienceDirect OnSite**), viene analizzato ed elaborato per produrre le statistiche a livello consortile e, per ogni singolo ente sottoscrittore del servizio.

Il file di log, per ogni record, contiene le seguenti informazioni: IP completo della macchina che ha effettuato l'accesso, ISSN della rivista, Titolo della rivista, Sigla dell'ente e Nome dell'ente.

Vengono rilasciati dei file in formato .htm riassuntivi degli accessi mensili a livello consortile e di singolo ente.

A livello consortile è possibile consultare:

- scarichi del mese con Titolo della rivista e numero di scarichi;
- grafico riepilogativo degli accessi relativi agli ultimi dodici mesi.

A livello di singolo ente:

- scarichi del mese con Titolo della rivista e numero di scarichi;
- scarichi raggruppati per IP (fino alla Classe C).

Avendo a disposizione una mole di dati imponente (centinaia di migliaia di record) si è cercato di realizzare un sistema di valutazione del servizio fornito ai sottoscrittori.

Questo sistema si articola in tre fasi fondamentali: analisi, misurazione e valutazione dei dati basandosi su due concetti fondamentali quali **efficienza** ed **efficacia** [1].

La fruizione di questi servizi è possibile grazie ad ingenti investimenti finanziari da parte dei singoli sottoscrittori quindi, ricopre un ruolo rilevante l'analisi ed il monitoraggio del processo di erogazione-fruizione degli stessi servizi.

La misura dell'efficienza, intesa come l'attitudine ad utilizzare correttamente le risorse disponibili, può essere effettuata attraverso la misura dell'utilizzo delle risorse finalizzate al raggiungimento degli obiettivi.

Si può affermare che un servizio è efficiente se minimizza l'uso degli investimenti ossia le risorse, a cui si accede tramite sottoscrizione finanziaria, sono utilizzate nella giusta misura.

La valutazione dell'efficacia si ottiene attraverso il confronto fra il risultato conseguito e quello che si era prefissato di raggiungere.

Si è cercato di proporre nuovi indicatori orientati a migliorare l'efficienza dei servizi ba-

sati su risorse elettroniche e, in grado di quantificare i costi nell'ambito di un contratto consortile.

La disponibilità di questi indicatori può essere utilizzata a fini gestionali per valutare ed assumere decisioni nel management di un contratto.

Le finalità di questo studio riguardano sia il CILEA, come Consorzio proponente il contratto, che i singoli sottoscrittori:

- dal punto di vista del CILEA possono essere utilizzati per la valutazione dell'efficienza complessiva del servizio offerto;
- a livello di singolo sottoscrittore per formulare e pianificare il budget nella maniera migliore, per gestire le singole riviste (mantenimento, cancellazione, o attivazione di una nuova) e monitorare un particolare abbonamento in un determinato periodo.

Il primo di questi indicatori che si viene qui ad illustrare è l'indice costo/utilizzo o, più semplicemente, **indice di utilizzo**.

L'analisi costo/utilizzo fornisce un'indicazione significativa sul livello d'equilibrio raggiunto tra le quote d'investimento e le quote d'accessi.

Per "quota d'investimento" s'intende la somma che ciascun ente sottoscrittore versa al CILEA per poter usufruire del servizio in questione.

La "quota d'accessi", invece, è data dal conteggio degli accessi registrati sul log del server opportunamente filtrato per eliminare le operazioni spurie. Infatti per "accesso" in questa trattazione si intende il "download" del full-text di un articolo e non, ad esempio, operazioni di browsing che vengono ugualmente registrate sul server.

Il calcolo dell'indice di utilizzo si ottiene prendendo in esame, per ogni ente sottoscrittore, le percentuali d'investimento e le percentuali di accesso. La percentuale di investimento è la quota investita rapportata all'ammontare generale del contratto, a livello consortile, espressa in termini percentuali. La percentuale di accesso è il numero di accessi rapportato alla somma totale di accessi registrati a livello consortile, sempre espressa in termini percentuali.

Calcolati questi valori si definisce l'indice di utilizzo come il rapporto tra le percentuali di accesso e le percentuali di investimento.

Quest'indice, così calcolato, fornisce il punto di equilibrio raggiunto tra le quote di investimento e le quote di accessi.

Questi valori sono riportati su un piano cartesiano (vedi Fig. 1). Sull'asse delle ascisse sono

rappresentati i valori delle percentuali di investimento, mentre sull'asse delle ordinate quelli relativi alle percentuali di accesso. Il confronto è definito dalla diagonale che, nel grafico, è la retta di riferimento per valutare se c'è equilibrio tra investimenti ed accessi. Un ente che occupi un punto situato sulla retta di equilibrio o nella regione posta superiormente ad essa, ha una percentuale di accessi uguale o superiore alla percentuale d'investimenti sostenuti per l'acquisizione delle risorse attribuibili a quell'ente. Per contro, quegli enti situati nella regione inferiore alla retta di equilibrio sono caratterizzati da una percentuale di investimenti superiori alla percentuale di accessi effettuati. L'area di piano posta al di sotto della diagonale rappresenta l'area critica.

Quest'analisi grafica ha un valido complemento nell'indicatore indice di utilizzo il quale si calcola rapportando la percentuale di utilizzo delle risorse elettroniche alla percentuale di investimento.

Quando il valore dell'indicatore è 1 si è nella condizione di equilibrio tra investimenti ed accessi. Se il valore è minore di 1 c'è stata una percentuale di investimenti maggiore rispetto a quella relativa agli accessi. Infine, se il valore è maggiore di 1 c'è stata una percentuale di accessi maggiore rispetto a quella degli investimenti effettuati.

In base agli investimenti effettuati gli enti partecipanti al contratto sono stati suddivisi in tre fasce:

- **Large**
- **Medium**
- **Small**

Nella categoria Large sono stati compresi 7 enti, nella Medium 9 enti e, nella Small 23 enti.

Vediamo come questi enti risultano distribuiti all'interno di differenti intervalli di valori assunti dall'indice di utilizzo.

Indice di utilizzo	Numero di enti	Distribuzione
oltre 2,5	3	3S
tra 1 e 2,5	7	3L, 1M, 3S
tra 0,7 e 1	14	2L, 6M, 6S
tra 0 e 0,7	15	2L, 2M, 11S

Tab. 1 - Distribuzione degli enti rispetto al valore dell'indice di utilizzo

Valutiamo i risultati raccolti nella Tab. 1.

I tre enti **Small** con indice di utilizzo superiore a 2,5 sono casi particolari di Università di recente costituzione, che hanno utilizzato da subito le risorse elettroniche senza aumentare in maniera rilevante il proprio posseduto.

Gli enti con indice di utilizzo minore di 0,7 sono invece enti con interesse mirato non in ambito biomedico (il catalogo Elsevier è composto per circa il 70% di titoli in tale ambito).

Riportiamo i valori medi registrati nelle tre categorie di enti prese in esame:

- 0,95 **Large**
- 0,65 **Medium**
- 0,96 **Small**

Utilizzare questi metodi può avere alcuni limiti, legati a particolari situazioni, tuttavia si ritiene che il confronto possa fornire delle ottime indicazioni, generali e di tendenza, sul rapporto tra l'uso delle risorse elettroniche e gli investimenti.

Gli altri due indici che sono stati presentati includono tra i loro termini di valutazione il valore dell' FTE (Full Time Equivalent).

Per la classificazione sono stati utilizzati i dati statistici pubblicati dal MIUR o si è considerato il numero di ricercatori per gli enti non accademici.

Gli enti che hanno aderito nei primi sei mesi dell'anno al contratto in oggetto sono stati complessivamente 39, così suddivisi: 23 Università, il CNR (fino ad aprile era suddiviso in 11 Aree o Istituti distinti), 8 strutture sanitarie (ospedali, IRCCS, ASL) e 7 enti di ricerca.

In base al valore dell'FTE gli enti si possono classificare in questo modo:

- 1 oltre 30000
- 3 tra 20000 e 30000
- 4 tra 10000 e 20000
- 8 tra 5000 e 10000
- 6 tra 2000 e 5000
- 17 tra 0 e 2000

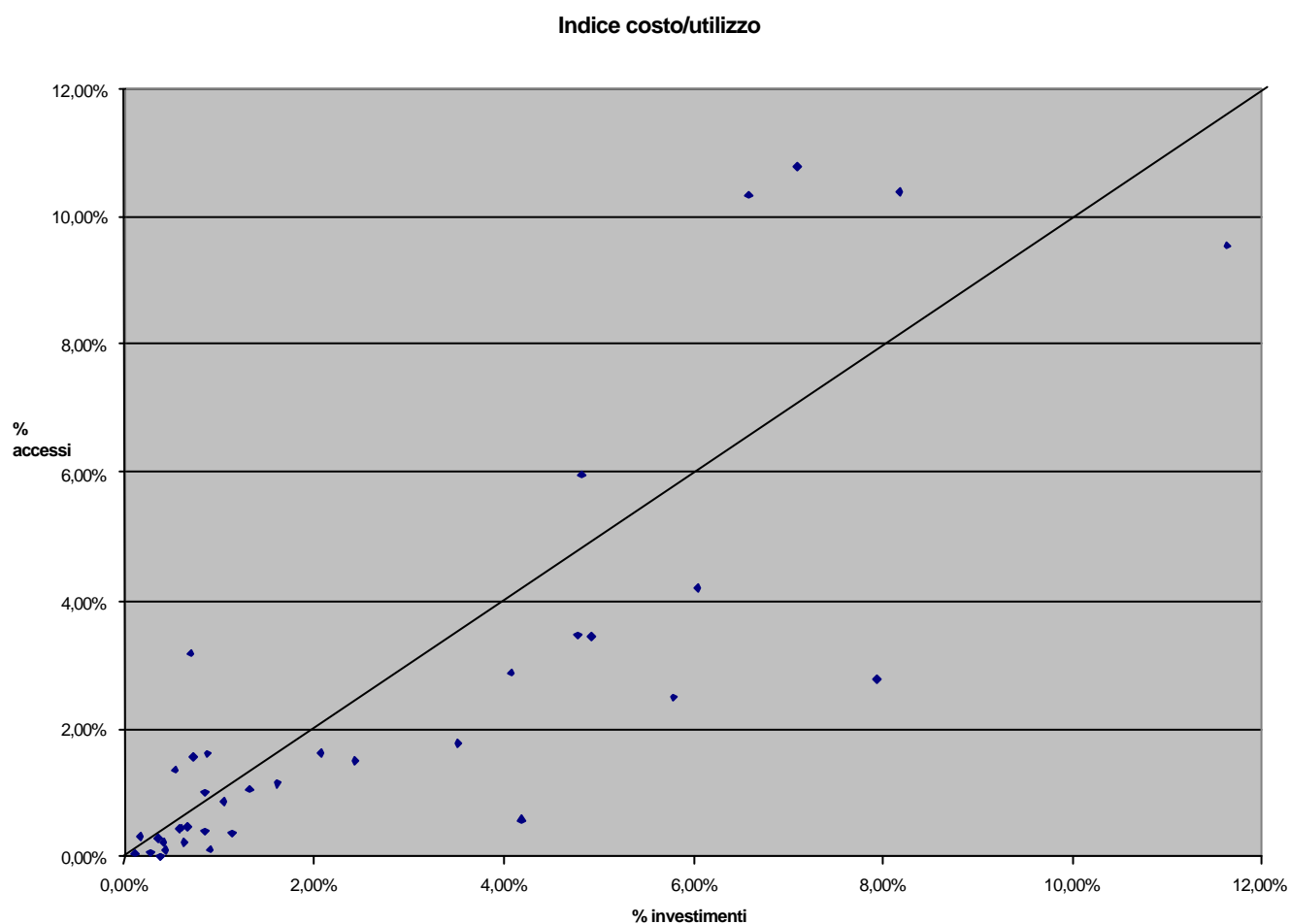


Fig. 1 - Grafico indice costo / utilizzo

Al fine di valutare il rapporto proporzionale tra gli accessi e l’FTE di ogni singolo ente sottoscrittore si è calcolato l’**indice accessi/FTE**.

Quest’indice si ottiene facendo il rapporto tra la percentuale degli accessi e la percentuale degli FTE, ottenuta rapportando l’FTE di ciascun ente alla somma degli FTE degli enti sottoscrittori.

Quest’indice ci permette di analizzare e valutare la quantità d’accessi di un ente rispetto al suo FTE.

Vediamo come questi enti risultano distribuiti all’interno di differenti intervalli di valori assunti dall’indice accessi/FTE.

Indice accessi/FTE	Numero di enti	Distribuzione
oltre 5	8	2L, 1M, 5S
tra 1 e 5	9	2M, 7S
tra 0,7 e 1	7	4L, 1M, 2S
tra 0 e 0,7	15	1L, 5M, 9S

Tab. 2 - Distribuzione degli enti rispetto al valore dell’indice accessi / FTE

Gli enti in cui si riscontra un indice accessi/FTE molto elevato, maggiore di 5, sono gli enti di ricerca o le strutture ospedaliere in cui i possibili fruitori del servizio, ricercatori o medici, hanno un ordine di grandezza (centinaia rispetto a migliaia) di tanto minore rispetto a quello degli enti accademici.

Questo è avvalorato dai dati che ci vengono forniti dal numero di accessi per persona che è di gran lunga superiore rispetto a quello che, in media, è attribuibile alle Università.

La media di accessi per persona per gli enti accademici è pari a 2,34 scarichi/persona, mentre la media di accessi per persona per gli enti di ricerca è uguale a 21,3 scarichi/persona.

L’ultimo indice che prendiamo in esame è l’indice **investimenti/FTE**, al fine di valutare il rapporto proporzionale tra gli investimenti e l’FTE di ogni singolo ente sottoscrittore.

Quest’indice si ottiene facendo il rapporto tra la percentuale degli investimenti, ossia la quota versata rapportata alla somma totale consortile, e la percentuale degli FTE, ottenuta rapportando l’FTE di ciascun ente alla somma degli FTE degli enti sottoscrittori.

Quest’indice ci permette di analizzare la mole degli investimenti di un ente rispetto al suo FTE.

Indice investimenti/FTE	Numero di enti	Distribuzione
oltre 5	11	2L, 1M, 8S
tra 1 e 5	12	3L, 3M, 6S
tra 0,6 e 1	5	1L, 2M, 2S
tra 0 e 0,6	11	1L, 3M, 7S

Tab. 3 - Distribuzione degli enti rispetto al valore dell’indice investimenti / FTE

Gli enti in cui si riscontra un indice investimenti/FTE molto elevato, maggiore di 5, sono enti di ricerca o strutture ospedaliere come nel caso dell’indice accessi/FTE.

Questo è avvalorato dai dati che derivano dal calcolo della quantità di euro investiti per persona che è nettamente superiore rispetto a quello che, in media, è attribuibile agli enti accademici.

La media di euro investiti per persona per gli enti accademici è pari circa a 35 euro mentre per gli enti di ricerca sale a 485 euro.

Conclusioni

E’ sempre più di attualità la necessità di interpretare i dati degli accessi alle risorse elettroniche. Il ricorso a nuovi indici o parametri di valutazione delle risorse elettroniche è sempre più frequente.

In questo modo si potrà effettuare un reale monitoraggio dell’uso e dell’efficienza del servizio in ambito di accessi alle risorse elettroniche.

Questi indici, e le loro interpretazioni, possono essere di supporto nella fase di contrattazione tra i vari enti sottoscrittori e il consorzio mettendo in luce eventuali disequilibri fra enti, in quanto per alcuni sottoscrittori può essere rilevata una percentuale di accesso non proporzionata alla percentuale di investimenti.

Queste situazioni potranno essere corrette e, conseguentemente, ci potrà essere una nuova distribuzione delle quote tra i vari enti all’interno del contratto consortile.

Bibliografia

[1] Maffenini W., *Le indagini sull’utilizzo delle risorse elettroniche remote della Biblioteca dell’Università di Milano-Bicocca*, Proceedings Bibliocom 2003, Roma.