



# Più citazioni in Open Access? Tendenze, opinioni e dati

---

Convegno CNBA

Il peso della ricerca.

Valutare una materia umanistica: architettura per esempio

Bologna, 22 maggio 2009

Maria Laura Vignocchi

Università degli Studi di Bologna

Elena Giglia

Università degli Studi di Torino



## Di cosa parliamo oggi

---

- ✓ L'Open Access e la disseminazione dei risultati della ricerca
- ✓ Più citazioni in Open Access? studi e metodi a confronto
- ✓ Un indicatore tradizionale: dati sull'Impact Factor delle riviste Open Access
- ✓ Cosa resta da fare...





# Open Access

---

Dalla Dichiarazione di Berlino (2003):

*The author(s) and right holder(s) of such contributions grant(s) to all users a free, irrevocable, worldwide, right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship [...] as well as the right to make small numbers of printed copies for their personal use*

**Due canali** per garantire l'accesso aperto all'informazione scientifica:

- ✓ **Green road**  
**autoarchiviazione** di articoli pre-print o post-print in **Open Archives** (istituzionali o disciplinari)
- ✓ **Gold road**  
pubblicazione su **riviste Open Access**, che
  - garantiscono la peer-review
  - hanno diverso modello di sostenibilità economica (nessun abbonamento, ma Article Processing Charge)



# Più citazioni in Open Access ?

---

Ipotesi da verificare:

→ l'accesso aperto si traduce in un maggior numero di citazioni?

Panorama della letteratura:

- Hitchcock, S. The effect of Open Access and downloads ("hits") on citation impact: a bibliography of studies (Op-Cit project)
- De Robbio, A. Analisi citazionale e indicatori bibliometrici nel modello Open Access, "Bollettino AIB", 47 (2007) 3, p. 257-288
- Connotea, tag "citation" (E.Giglia) contributi suddivisi per metodo di ricerca (OA/TA same journal, OA/TA discipline, OA/TA multidiscipline, reviews, method...)



# Più citazioni in Open Access?

## Primi risultati

---

**Tutti** gli studi condotti finora **rilevano** – in misura più o meno accentuata - un **vantaggio citazionale** per gli articoli Open Access...

Oggi il dibattito si è spostato sulle possibili **cause** di questo vantaggio (e qui a volte entra in gioco l'ideologia...)

Vediamo allora alcune questioni di metodo...



# Più citazioni in Open Access?

## Questioni di metodo - 1

---

✓ è difficile **definire l'impatto**

- Science is a gift-based economy: value is defined as the degree to which one's ideas have contributed to knowledge and impacted the thinking of others

[Bollen, Van de Sompel et al. 2009](#)

✓ è ancora più difficile **misurarlo**

- tradizionale misura quantitativa: citazioni e Impact Factor
- nuove metriche per l'era digitale: analisi delle social networks, analisi dei log files, analisi dei downloads ([MESUR](#))

✓ è difficilissimo **confrontare OA/TA**

- OA non riguarda solo articoli: considerare i pre-prints
- enormi differenze disciplinari
- difficile creare set di dati comparabili



# Più citazioni in Open Access?

## Questioni di metodo - 2

---

- Seglen: citazioni fortemente **"skewed"**:
  - 15% degli articoli riceve 50% delle citazioni
  - 50% di questi più citati prende 90% delle citazioni

Seglen, P.O. The skewness of science, JASIST, 43 (9), 628-638
- Harnad:
  - il vero vantaggio dell'Open Access è l'**accesso**; vantaggio dell'accesso è **uso e impatto**, di cui le **citazioni** sono solo un indicatore
  - l'**accesso** è una **condizione necessaria** – per essere citabile un articolo deve essere accessibile - ma non sufficiente: è la **qualità** che dà valore intrinseco all'articolo
  - articoli di qualità hanno maggiori possibilità di essere citati (valore intrinseco, indipendente da OA); la possibilità cresce quanto più sono accessibili a tutti
  - vantaggio perché gli articoli migliori e più citati hanno più probabilità di essere messi OA dagli autori o perché gli articoli OA hanno più probabilità di essere citati?  
Entrambi..



# Più citazioni in Open Access?

## Questioni di metodo - 3

---

**OA Advantage** si compone di:

**EA** Early Advantage

archiviare i preprint aumenta le citazioni (sono disponibili prima)

**QA** Quality Advantage

~10% degli articoli riceve 90% delle citazioni

**AA** Accessibility Advance

accesso universale

**UA** Usage Advantage (o **DA** Download Adv.)

autoarchiviazione aumenta i download

**QB** Quality bias

articoli migliori hanno maggiori possibilità di essere autoarchiviati (self selection)

**CA** Competitive Advantage

vantaggio dell' accesso libero su TA

Harnad S., [The Open Access citation advantage: Quality Advantage or Quality Bias?](#), 2007

[i concetti di EA e QB sono già di [Kurtz](#), 2005]



# Più citazioni in Open Access?

## Questioni di metodo - 4

---

- siamo sempre nell'ambito delle CITAZIONI
- come misurare lo scarto fra citazioni OA e TA?
  - ✓ articoli OA/TA nella stessa rivista
  - ✓ articoli OA/TA (disciplinare)
  - ✓ articoli OA/TA (multidisciplinare)
  - ✓ pochi autori considerano anche i pre-print...
- vedremo qualche esempio, con riferimento al set di dati esaminato, per mostrare l'eterogeneità delle scelte e dei criteri



# Più citazioni in Open Access?

## Risultati - OA/TA nella stessa rivista

---

- Eysenbach su PNAS Dic. 2004-Apr.2006:  
Articoli OA hanno 2,1 possibilità in più di essere citati nei primi 10 mesi; 2,9 fra 10 e 16 mesi  
Eysenbach, G. (2006) [Citation Advantage of Open Access Articles](#)  
*PLoS Biology*, Volume 4, Issue 5, May 2006
- Kurtz su Astrophysical Journal 1997 OA e 1998 TA:  
Stesso trend di citazioni, nessun vantaggio OA. La vera differenza è EA (se esiste pre-print in arXiv) e QB (articoli migliori vengono scelti dagli autori)  
Kurtz-Henneken [Open Access does not increase citations for research articles from The Astrophysical Journal](#) arXiv, 6 Sept. 2007  
Stevan Harnad nota che dove non c'è problema di accesso (astrofisica è disciplina molto ben finanziata) non si dà vantaggio OA [[blog](#) 7 Sept. 2007]



# Più citazioni in Open Access?

## Risultati – OA/TA in una disciplina

---

- Zhang in Scienze della comunicazione: Journal of Computer-Mediated Communication (JCMC) e New Media & Society (NMS) art. pubblicati 2001 e 2002  
JCMC [OA]: 309 cit.; NMS [TA] 159 cit.  
Zhang, Y. (2006) [The Effect of Open Access on Citation Impact: A Comparison Study Based on Web Citation Analysis Libri](#), September 2006
- Davis 2008 in Biomedicina (11 riviste dal 2004 al 2008 solo considerando l'opzione Author's Choice)  
+ 21% citaz. per OA in media, ma con una riduzione del 7% all'anno dal 2004 al 2007  
Davis, P. [Author-choice open access publishing in the biological and medical literature: a citation analysis](#), arXiv, 2008

Secondo Davis è solo un fatto di QB

Secondo Harnad:

- difetti metodologici (PNAS usato solo parzialmente, calcolo non disaggregato per anni, non considerate le versioni preprint, non specificato quanti art. OA quanti TA...)
- significa al contrario che vantaggio OA cresce nel tempo
- vantaggio OA si compone proprio di EA+QA (QB)

[Blog](#) 25 Aug. 2008



# Più citazioni in Open Access?

## Risultati – OA/TA multidisciplinare

---

- Norris et al. 2009 su 4 discipline  
(ecol, mat appl, sociol, econ)

Qualsiasi ne sia la causa, vantaggio OA  
esiste: 65 % in media (sociol: 108%)

Norris, M.-Oppenheim, C.- Rowland, F.

[The citation advantage of open-access articles,](#)

*JASIST*, Vol. 59, No. 12, 2008

- Hajjem et al. 2005

10 discipline, vantaggi vanno dal 25 al 250%

Hajjem, C., Harnad, S. and Gingras, Y. (2005)

[Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access  
and How it Increases Research Citation Impact,](#)

IEEE Data Engineering Bulletin, Vol. 28 No. 4, December 2005



# Più citazioni in Open Access?

## Risultati - Due articoli controversi - 1

---

- Craig et al. 2007 revisione critica della letteratura i cui risultati sembrano porre seri dubbi... **ma**
  - gli autori sono affiliati alle maggiori case editrici commerciali
  - si basano soprattutto sui pochi articoli che mostrano dubbi o dati non eclatanti
  - in ogni caso, NON dimostrano che il vantaggio non c'è, invitano solo a studiarne le cause e le correlazioni causa/effetto...

Craig, I. D., Plume, A. M., McVeigh, M. E., Pringle, J. Amin, M.  
[Do Open Access Articles Have Greater Citation Impact?](#)  
[A critical review of the literature](#)  
Publishing Research Consortium, *Journal of Informetrics*,  
1 (3): 239-248, July 2007

Harnad: OA è causa non effetto collaterale...

[blog](#), 26 May 2007



# Più citazioni in Open Access?

## Risultati – Due articoli controversi - 2

---

- Davis et al. 2008: trial randomizzato controllato
  - 11 riviste American Physiological Society, con articoli assegnati a caso al gruppo OA o TA (controllo)
  - previsto studio su 4 anni, pubblicato dopo solo 1
  - OA: 89% downloads in più, ma non molte citazioni in più a 12 mesi dall'uscita
  - non esiste vantaggio citazionale OA, o è dovuto ad altre cause (QB)

Davis, P.M., Lewenstein, B.V., Simon, D.H., Booth, J.G., Connolly, M.J.L.

[Open access publishing, article downloads, and citations: randomised controlled trial](#)

*BMJ*, 2008;337:a568, 31 July 2008

- dopo un anno nessun effetto, è vero (il picco è nel 3° anno [H. Moed])
- il controllo andava fatto con articoli OA scelti dall'autore e autoarchiviati (per verificare QB)

Harnad, S. [blog](#), 31 Jul 2008



# Più citazioni in Open Access?

## Risultati – Un nuovo approccio

---

- Gargouri - Harnad 2009
    - per evitare QB, dataset comprende anche depositi istituzionali con "mandato"
    - non considera riviste OA (se no non ci sarebbe articolo di controllo)
    - compara una serie notevole di variabili, escluse le quali si nota un vantaggio OA puro
- articoli OA hanno maggiori citazioni, tanto più a livello alto (dimostra QA)
- articoli OA in depositi con mandato hanno maggiori citazioni a livello medio alto

Gargouri, Y – Harnad S. [Logistic regression of potential explanatory variables on citation counts](#), preprint, Apr. 2009



# Più citazioni in Open Access? Opinioni...

---

Studi fortemente connotati dal punto di vista:

- ✓ **ideologico**  
i numeri vengono letti a supporto di una tesi (in particolare per chi è avverso all'OA)
- ✓ **metodologico**  
i disegni di studio non sono neutri: cosa studio e come lo studio
- ✓ **disciplinare**  
profonde differenze fra discipline:
  - ✓ abitudine al deposito dei pre-print
  - ✓ accesso illimitato alle risorse (es. astrofisica)
  - ✓ "mandati" a depositare
  - ✓ attitudine alle citazioni



# I dati: IF e riviste Open Access

---

Proviamo a testare un indicatore tradizionale:  
l'Impact Factor

- riconosciuto
- utilizzato da anni
- confrontabile
- ... criticato (giustamente) da più parti

→ non entreremo nel merito delle polemiche sull'IF...

I dati che vedrete sono **dati preliminari** di uno studio originale, non ancora sottoposto a referaggio, da usare solo dietro esplicito consenso da richiedere agli autori, Elena Giglia e Sandra Migliore dell'Università di Torino ([elena.giglia@unito.it](mailto:elena.giglia@unito.it), [sandra.migliore@unito.it](mailto:sandra.migliore@unito.it))



# Di cosa parliamo...

---

## ➤ **Journal Citation Reports (JCR)**

- prodotto da ISI Thomson Reuters
- due edizioni annuali, Science [1975-] e Social Science [1977-]
- contiene i dati sull'IF e IMM IND (e altro...)

## ➤ **Impact Factor**

- numero medio di **citazioni** ricevute nell'**anno** di riferimento dagli articoli pubblicati nei **due anni precedenti**
- es. IF 2007 per BMC Biology:

$$\frac{[168 \text{ (cit 2006)} + 176 \text{ (cit 2005)}] = 344}{[41 \text{ (art 2006)} + 27 \text{ (art 2005)}] = 67} = 5,059$$

## ➤ **Immediacy Index**

- numero medio di **citazioni** ricevute nell'anno di riferimento dagli articoli pubblicati **nell'anno stesso**
- N.B. favorisce riviste di grandi dimensioni e con periodicità alta

## ➤ **Article Influence Score/Eigenfactor**

- creato da Carl T. Bergstrom (Univ. Washington)
- si calcola su Eigenfactor diviso la frazione del numero di articoli pubblicato da una rivista
- Eigenfactor = numero medio di **citazioni** ricevute nei **5 anni precedenti** l'anno di riferimento, **pesato** per disciplina e importanza della rivista citante
- **NON** ancora significativo per le riviste OA perché troppo "giovani"



# Ipotesi e metodo - 1

---

- L'**ipotesi di lavoro** è che le riviste Open Access, essendo liberamente accessibili, ottengano in breve buoni indici di impatto
- I **dati** analizzati:
  - [Journal Citation Reports \(JCR\)](#) Science edition 2007
  - [Journal Citation Reports \(JCR\)](#) Social Science edition 2007  
(data l'esiguità dei titoli OA presenti, non si è approfondito lo studio)
  - [Directory of Open Access Journals \(DOAJ\)](#), scelta come repertorio più accreditato per le riviste OA (i precedenti studi avevano considerato anche SCiELO – ormai listato in DOAJ - e JSTAGE, piattaforma giapponese con testi free non sempre OA)



## Ipotesi e metodo – 2

---

### ➤ Il **metodo**:

- per verificare la **copertura** : confronto fra Journal of Citation Reports (JCR) e Directory of Open Access Journals (DOAJ) al **31 dicembre dell'anno corrispondente**
- aggregazione **geografica** (paese di pubblicazione) e **temporale** (anno di nascita, difficile risalire alla conversione OA, lista in Open Access Directory, [http://oad.simmons.edu/oadwiki/Main\\_Page](http://oad.simmons.edu/oadwiki/Main_Page) )
- verifica del **ranking** sia per l'**IF** sia per l'**Immediacy Index** all'interno della singola categoria (se un titolo appartiene a più categorie, è stato sdoppiato)
- **comparazione** grazie alla riduzione in **percentili** e all'aggregazione in **macro aree** disciplinari:
  - ✓ A – CHEMISTRY
  - ✓ B - MATHEMATICS, PHYSICS, ENGINEERING
  - ✓ C – LIFE SCIENCES
  - ✓ D – MEDICINE
  - ✓ E – HUMAN SCIENCES
- inizialmente si è cercato di aggregare i dati per un confronto con Mc Veigh 2004: il confronto è reso arduo dalla diversa copertura (non solo DOAJ) e dai criteri di lavoro non sempre esplicitati



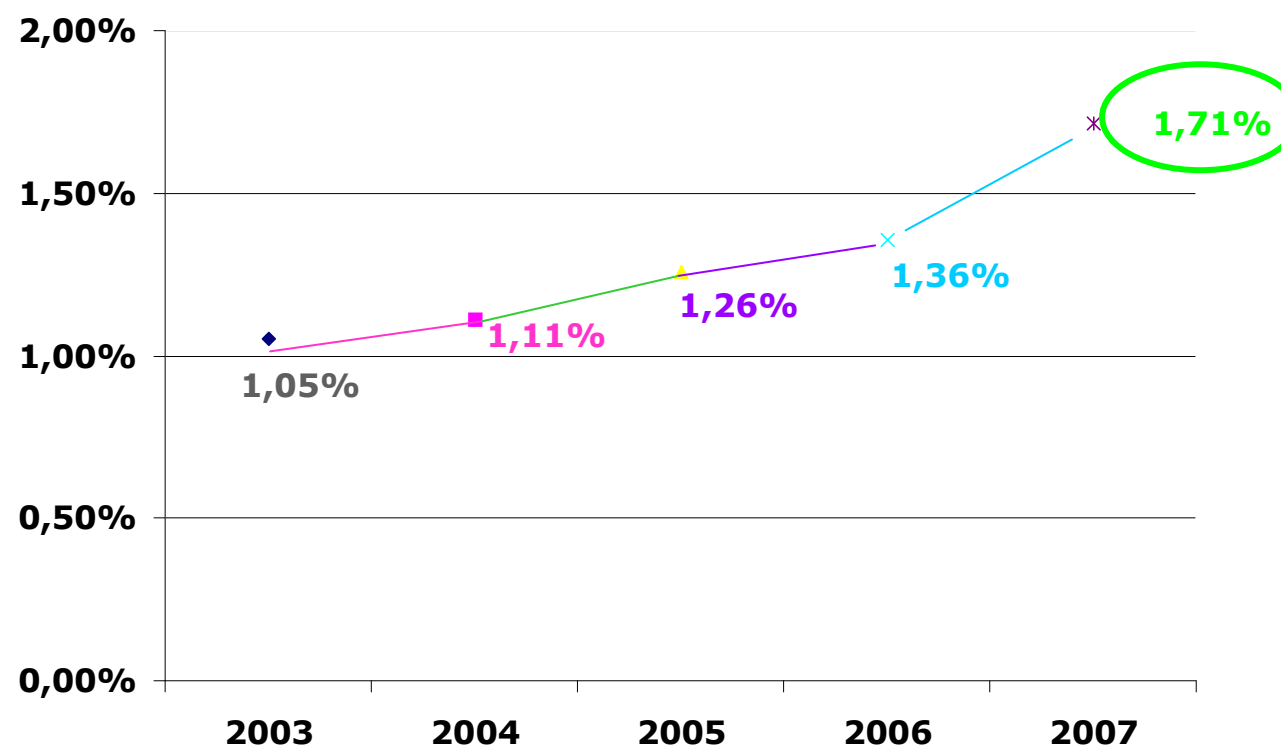
# Titoli Open Access in JCR SOCIAL SCIENCES

---

Criteria di confronto: Journal of Citation Reports 2007 [18 giugno 2008]  
con lista titoli presenti in Directory of Open Access Journals  
al 31 dicembre 2007 [=riviste su cui si è calcolato IF nel 2007]

<b>Anno</b>	<b>Titoli in JCR</b>	<b>Titoli in DOAJ</b>	<b>OA con IF</b>	<b>OA con IF</b>
2003	1714	602	18	1,05%
2004	1712	1194	19	1,11%
2005	1747	1811	22	1,26%
2006	1768	2357	24	1,36%
2007	1866	2954	32	1,71%

# Titoli Open Access in JCR SOCIAL SCIENCES



...e con questo chiudiamo il discorso su JCR Social Science,  
perché non ancora significativo...



# Titoli Open Access in JCR SCIENCE

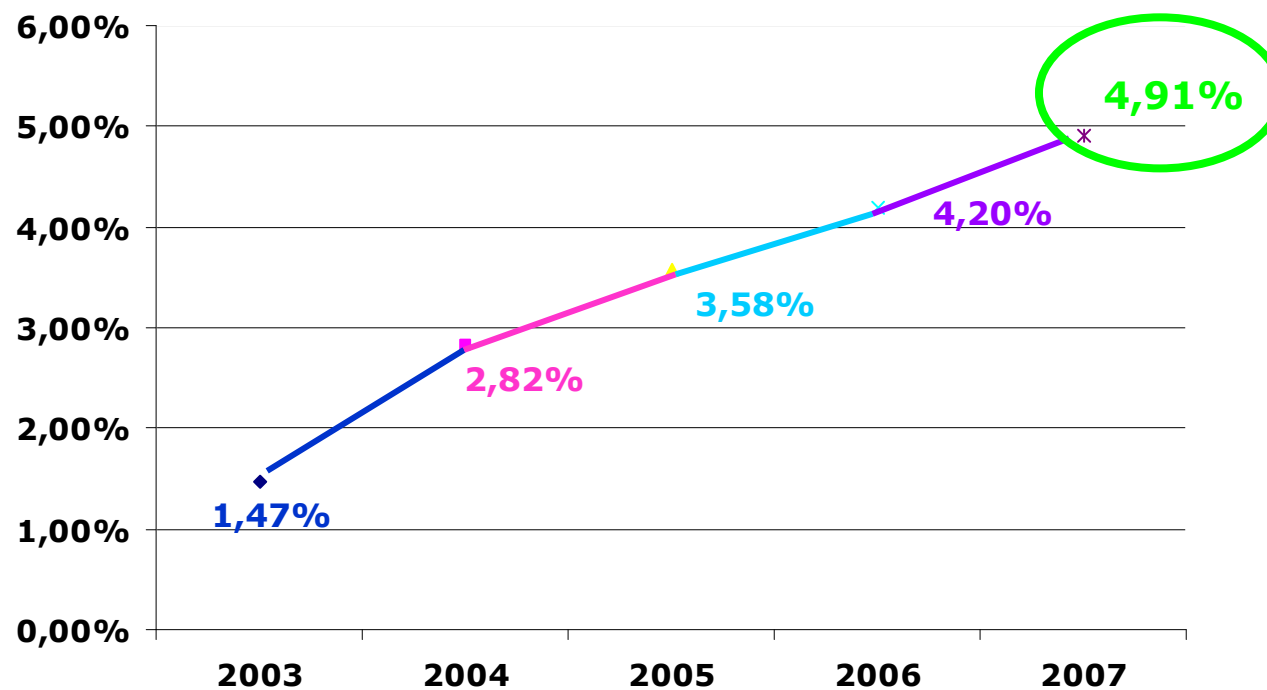
---

Criteria di confronto: Journal of Citation Reports 2007 [18 giugno 2008]  
con lista titoli presenti in Directory of Open Access Journals  
al 31 dicembre 2007 [=riviste su cui si è calcolato IF nel 2007]

<b>Anno</b>	<b>Titoli in JCR</b>	<b>Titoli in DOAJ</b>	<b>OA con IF</b>	<b>OA con IF</b>
2003	5907	602	87	1,47%
2004	5968	1194	168	2,82%
2005	6088	1811	218	3,58%
2006	6164	2357	259	4,20%
2007	6417	2954	315	4,91%

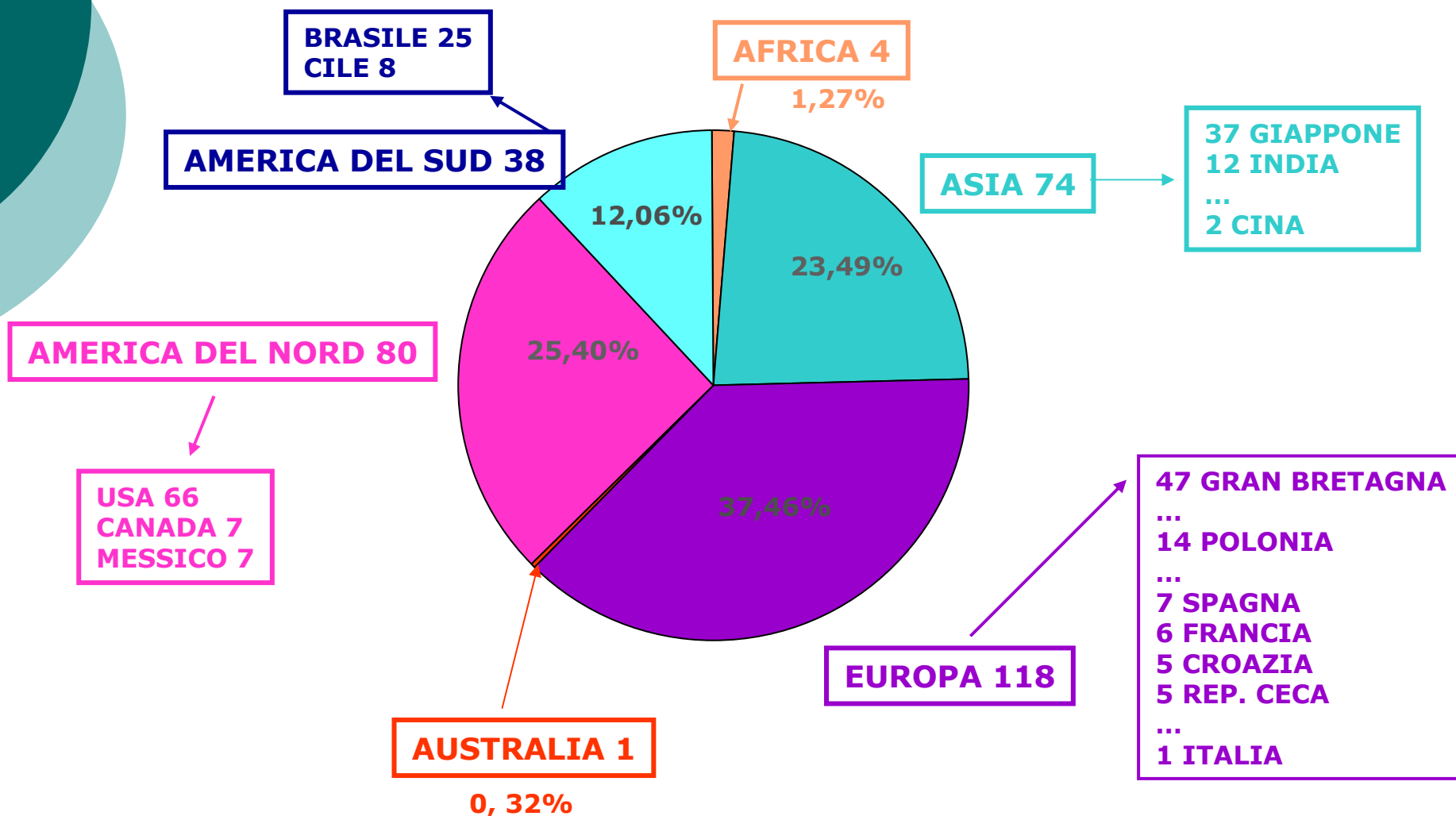
# Titoli Open access in JCR SCIENCE

---



# JCR SCIENCE 2007

## Titoli per area geografica





# Anno di inizio titoli

---

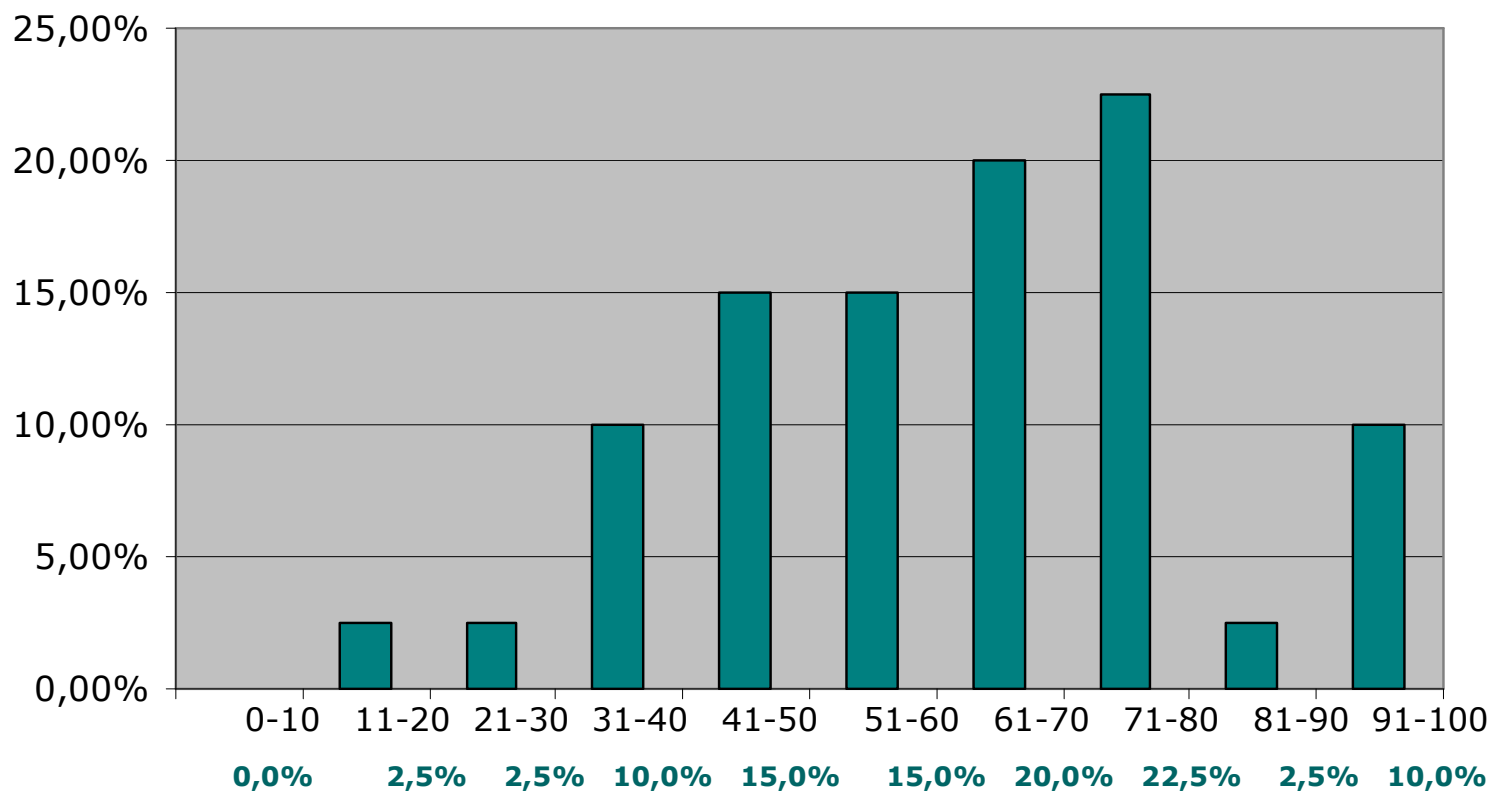
Valori mediani, dal primo all'ultimo percentile

<b>PERC.</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
<b>1</b>		<b>1997</b>	<b>2003</b>	<b>1996</b>
<b>2</b>	<b>2006</b>	<b>1996</b>	<b>1998</b>	<b>1999</b>
<b>3</b>	<b>2001</b>	<b>1995</b>	<b>2000</b>	<b>1998</b>
<b>4</b>	<b>1997</b>	<b>1997</b>	<b>2000</b>	<b>2002</b>
<b>5</b>	<b>1998</b>	<b>1997</b>	<b>1998</b>	<b>2001</b>
<b>6</b>	<b>1997</b>	<b>1997</b>	<b>2000</b>	<b>2001</b>
<b>7</b>	<b>1997</b>	<b>1997</b>	<b>2000</b>	<b>2000</b>
<b>8</b>	<b>2000</b>	<b>1999</b>	<b>2000</b>	<b>2000</b>
<b>9</b>	<b>1997</b>	<b>1998</b>	<b>2000</b>	<b>2000</b>
<b>10</b>	<b>2001</b>	<b>2001</b>	<b>2000</b>	<b>2000</b>

# Percentili di ranking – Dati in percentuale

## IMPACT FACTOR

### A - CHEMISTRY 40 TITOLI

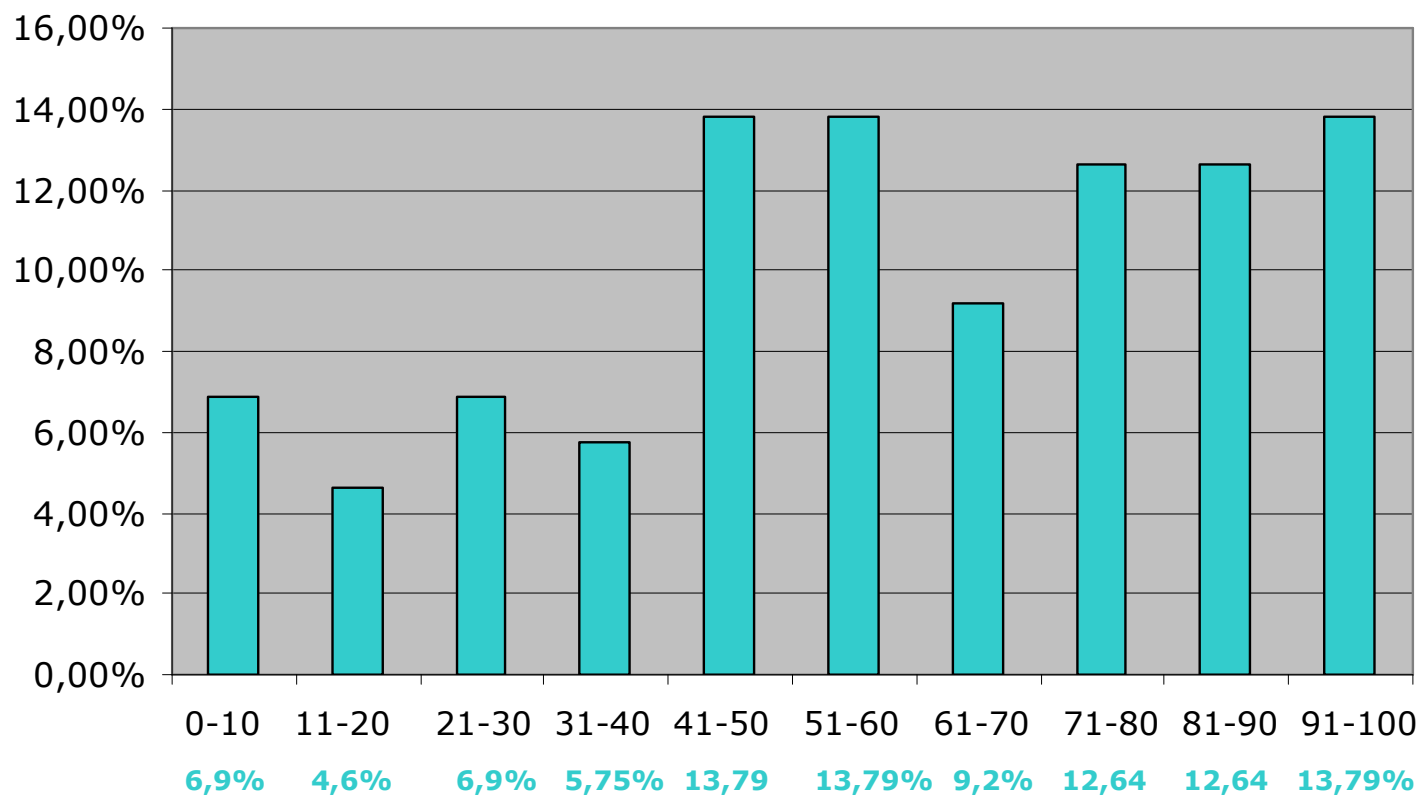


Si tratta di dati preliminari, non ancora sottoposti a referaggio, da usare solo dietro esplicito consenso da richiedere agli autori ([elena.giglia@unito.it](mailto:elena.giglia@unito.it), [sandra.migliore@unito.it](mailto:sandra.migliore@unito.it))

# Percentili di ranking – Dati in percentuale

## IMPACT FACTOR

### B - MATHEMATICS PHYSICS ENGINEERING 87 TITOLI

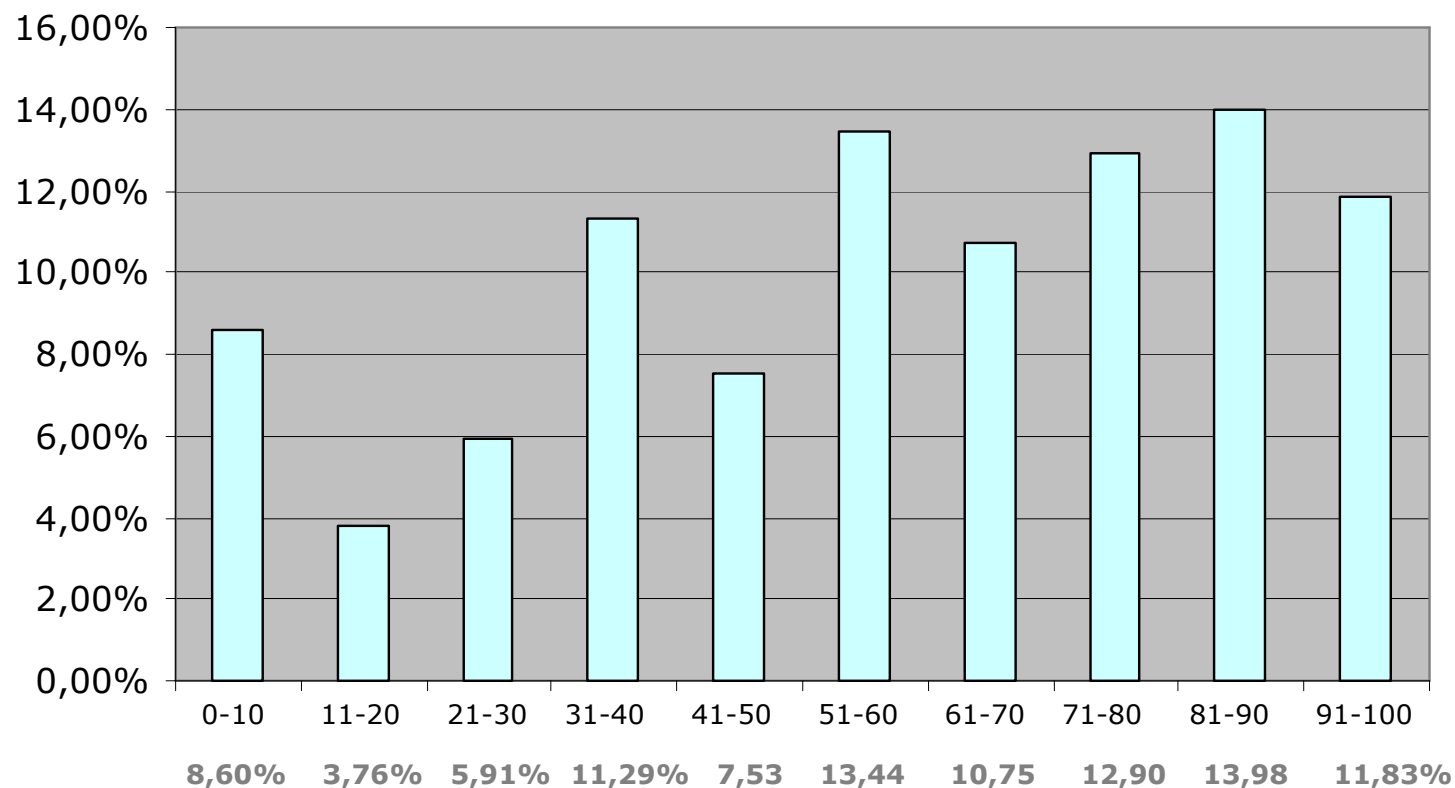


Si tratta di dati preliminari, non ancora sottoposti a referaggio, da usare solo dietro esplicito consenso da richiedere agli autori ([elena.giglia@unito.it](mailto:elena.giglia@unito.it), [sandra.migliore@unito.it](mailto:sandra.migliore@unito.it))

# Percentili di ranking – Dati in percentuale

## IMPACT FACTOR

### C - LIFE SCIENCES 186 TITOLI

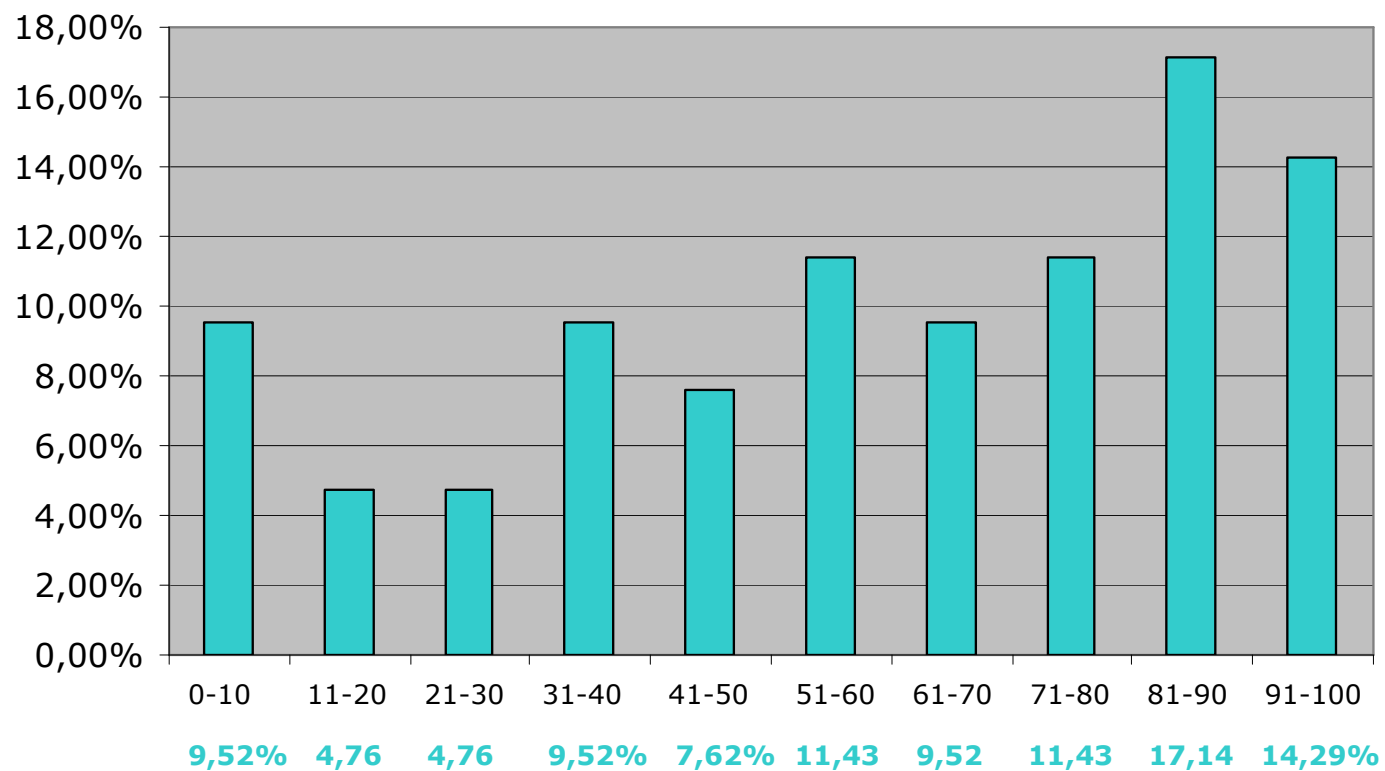


Si tratta di dati preliminari, non ancora sottoposti a referaggio, da usare solo dietro esplicito consenso da richiedere agli autori ([elena.giglia@unito.it](mailto:elena.giglia@unito.it), [sandra.migliore@unito.it](mailto:sandra.migliore@unito.it))

# Percentili di ranking – Dati in percentuale

## IMPACT FACTOR

### D - MEDICINE 105 TITOLI

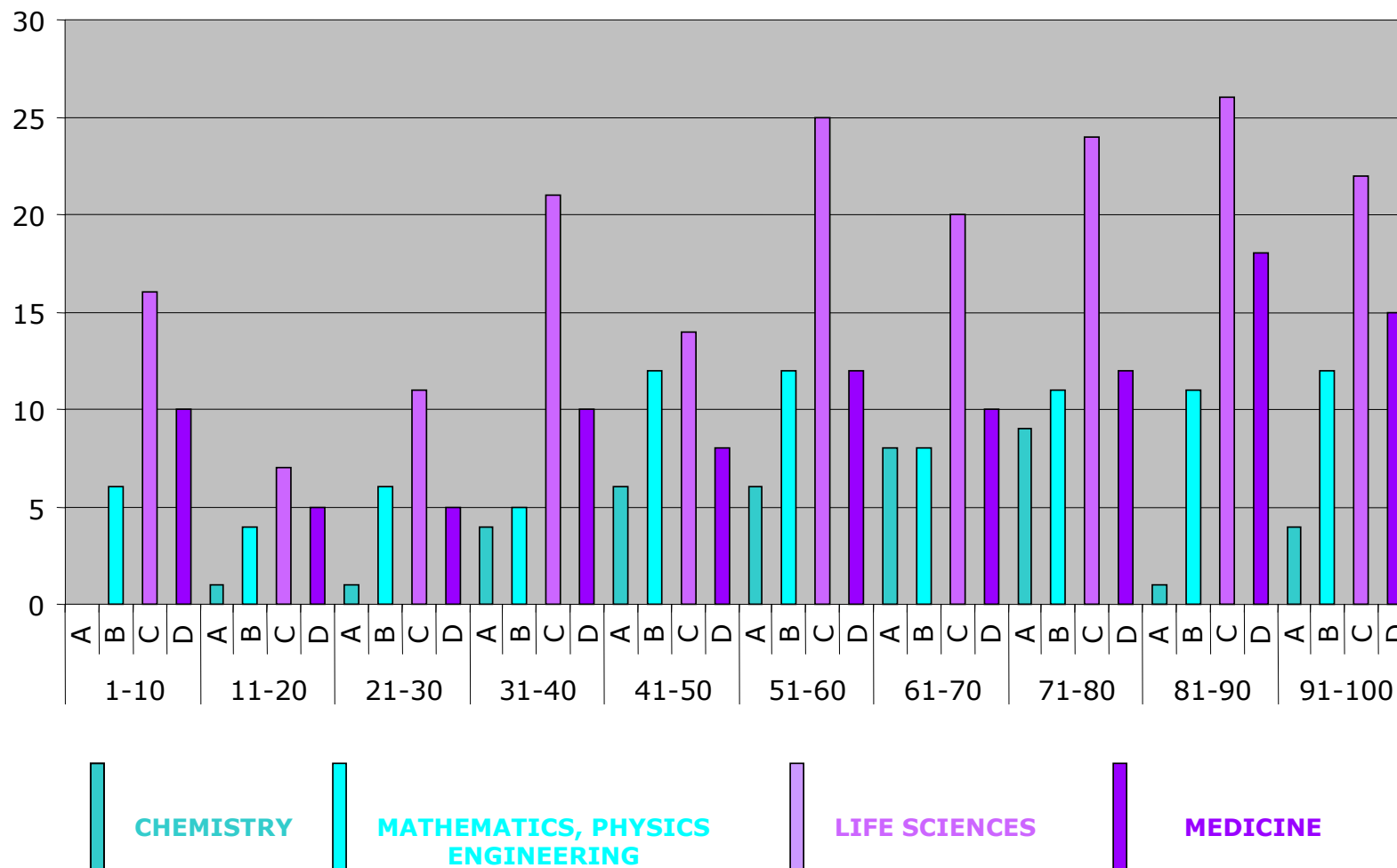


Si tratta di dati preliminari, non ancora sottoposti a referaggio, da usare solo dietro esplicito consenso da richiedere agli autori ([elena.giglia@unito.it](mailto:elena.giglia@unito.it), [sandra.migliore@unito.it](mailto:sandra.migliore@unito.it))

# Percentili di ranking – dati aggregati

## IMPACT FACTOR

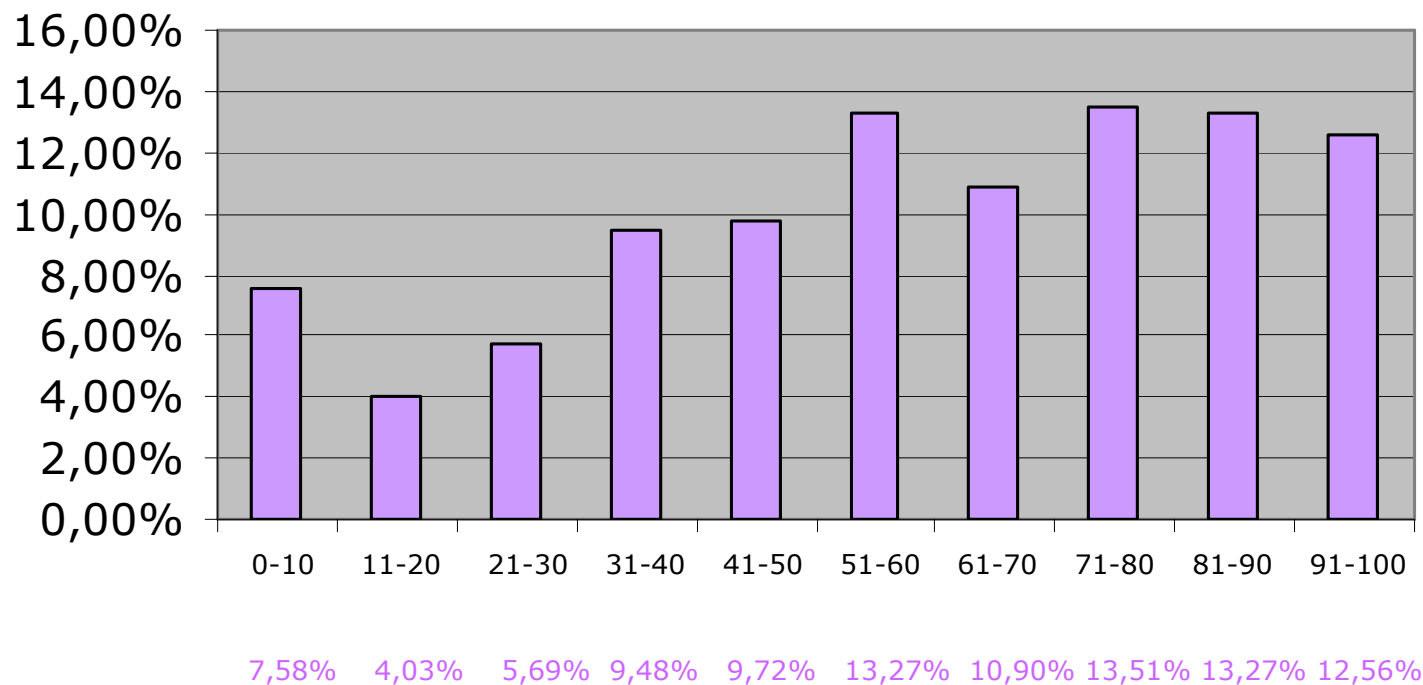
Si tratta di dati preliminari, non ancora sottoposti a referaggio, da usare solo dietro esplicito consenso da richiedere agli autori (elena.giglia@unito.it, sandra.migliore@unito.it)



# Percentili di ranking - Totali dati in percentuale

## IMPACT FACTOR

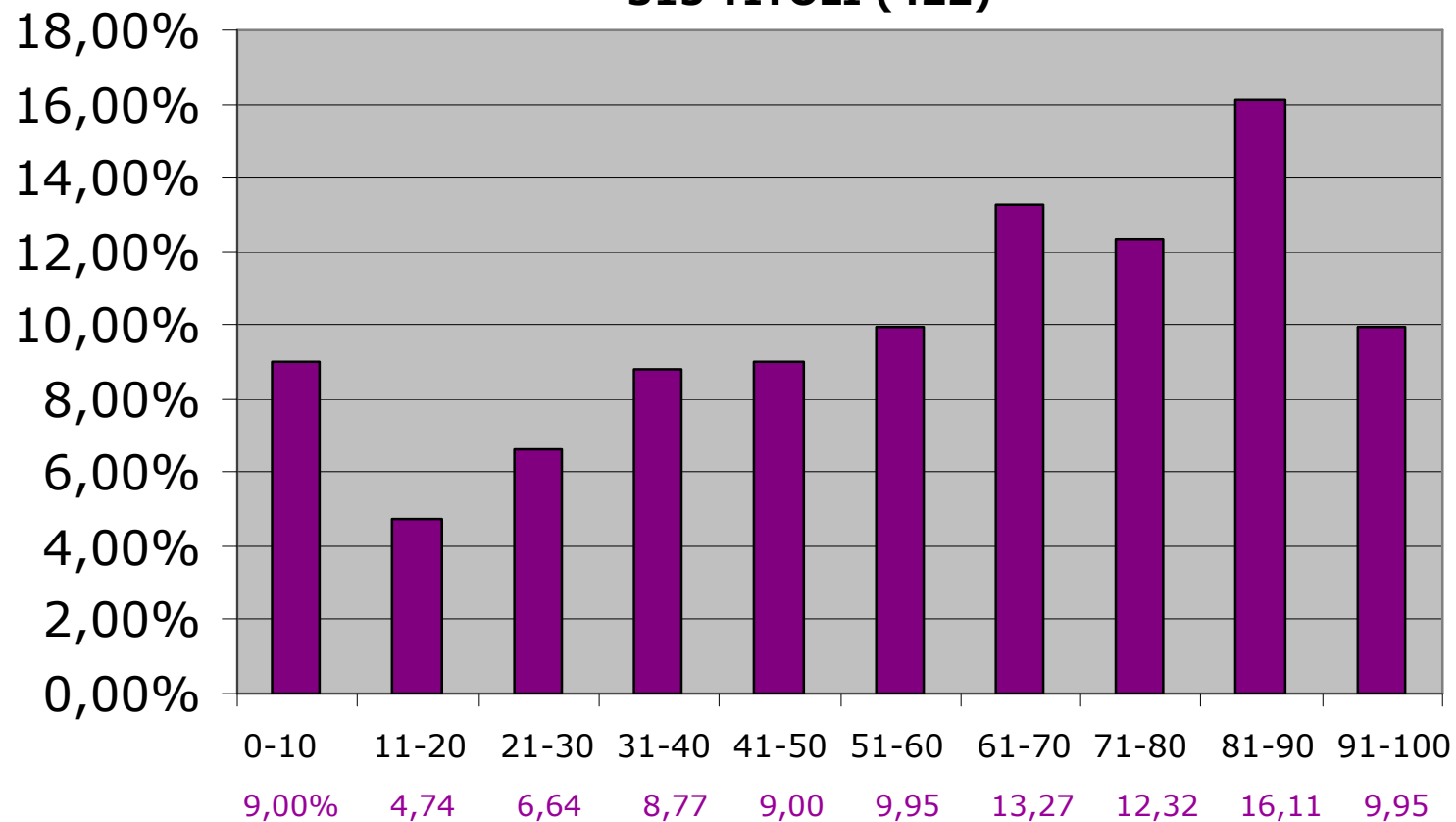
**JCR SCIENCE 2007**  
**315 TITOLI (422)**



Si tratta di dati preliminari, non ancora sottoposti a referaggio, da usare solo dietro esplicito consenso da richiedere agli autori ([elena.giglia@unito.it](mailto:elena.giglia@unito.it), [sandra.migliore@unito.it](mailto:sandra.migliore@unito.it))

# Percentili di ranking - Totali dati in percentuale IMMEDIACY INDEX

**JCR SCIENCE 2007  
315 TITOLI (422)**

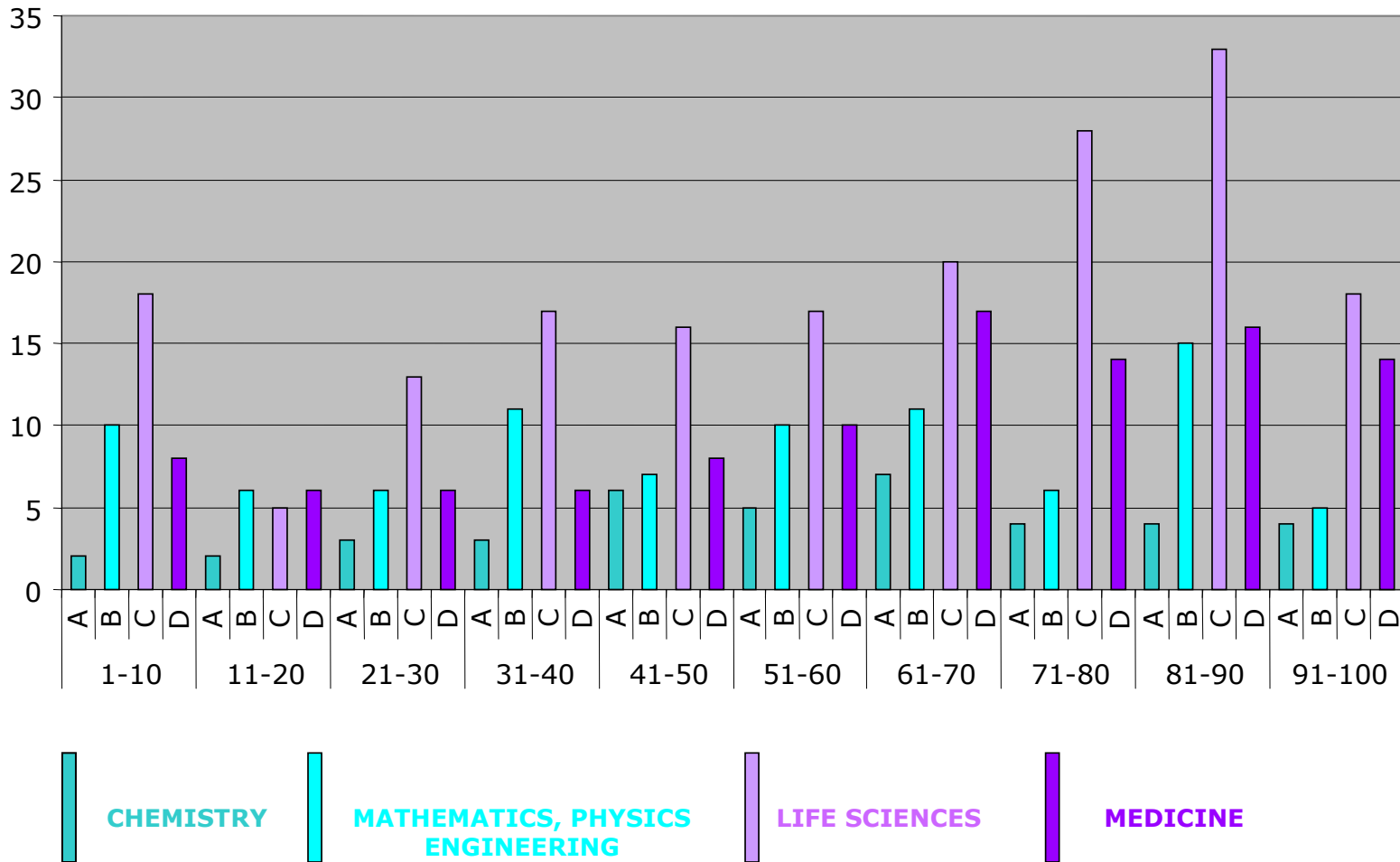


Si tratta di dati preliminari, non ancora sottoposti a referaggio, da usare solo dietro esplicito consenso da richiedere agli autori ([elena.giglia@unito.it](mailto:elena.giglia@unito.it), [sandra.migliore@unito.it](mailto:sandra.migliore@unito.it))

# Percentili di ranking – dati aggregati

## IMMEDIACY INDEX

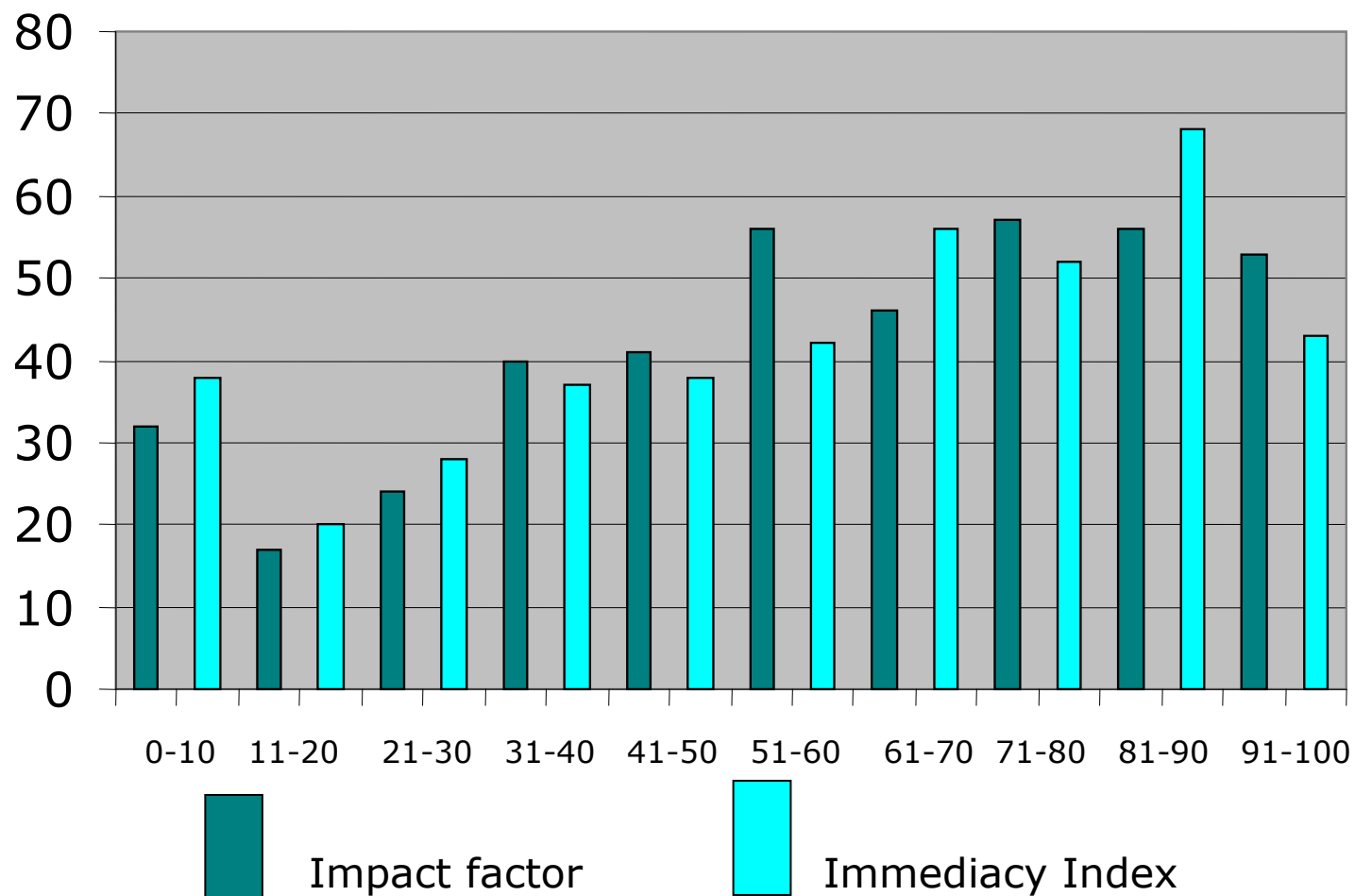
Si tratta di dati preliminari, non ancora sottoposti a referaggio, da usare solo dietro esplicito consenso da richiedere agli autori ([elena.giglia@unito.it](mailto:elena.giglia@unito.it), [sandra.migliore@unito.it](mailto:sandra.migliore@unito.it))



# Impact factor/Immediacy index

## Dati aggregati per le 4 aree

Si tratta di dati preliminari, non ancora sottoposti a referaggio, da usare solo dietro esplicito consenso da richiedere agli autori (elena.giglia@unito.it, sandra.migliore@unito.it)

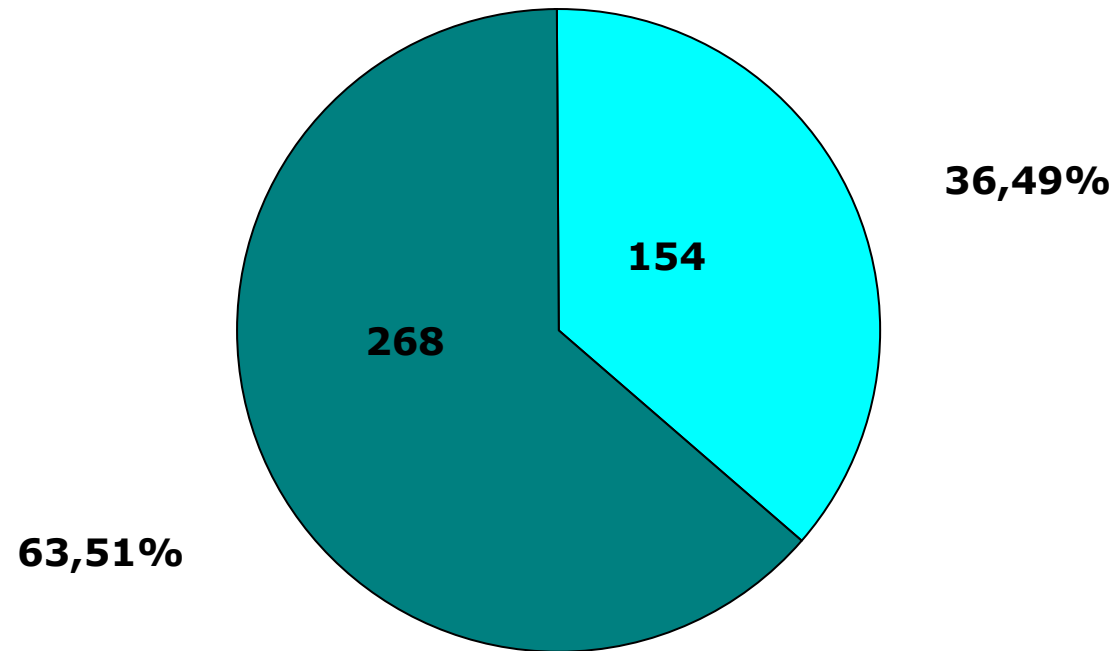


# Percentili di ranking – dati aggregati

## IMPACT FACTOR

---

Si tratta di dati preliminari, non ancora sottoposti a referaggio, da usare solo dietro esplicito consenso da richiedere agli autori ([elena.giglia@unito.it](mailto:elena.giglia@unito.it), [sandra.migliore@unito.it](mailto:sandra.migliore@unito.it))



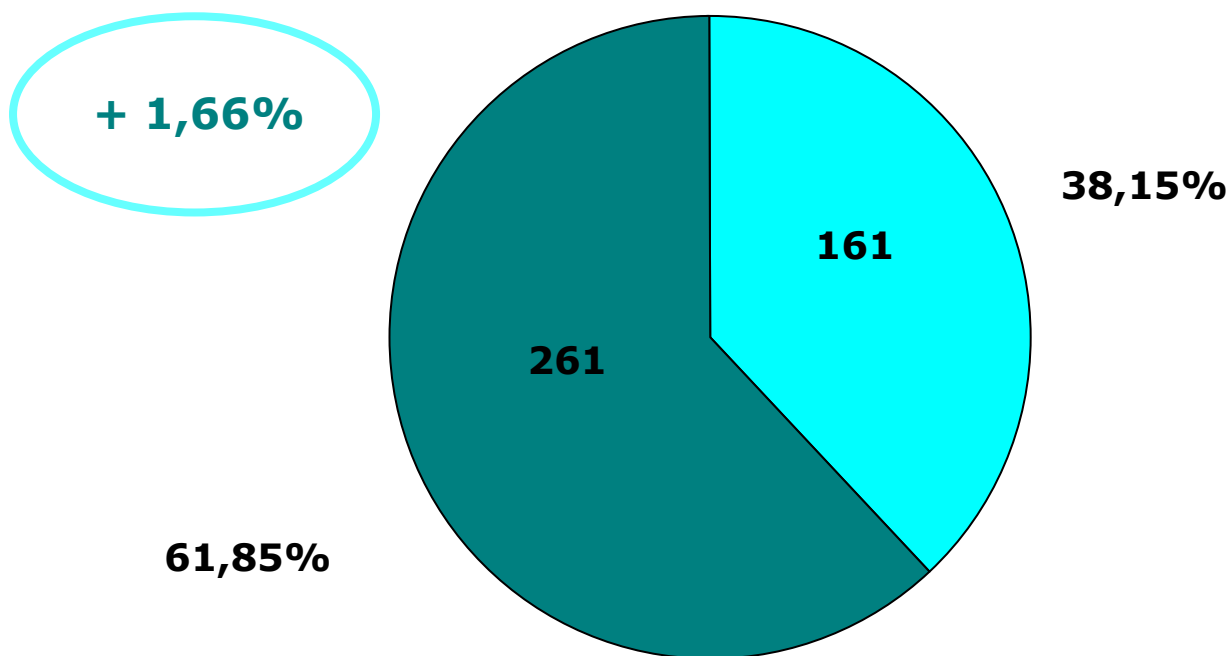
Primi 5 percentili: 154 titoli pari a 36,49%  
Ultimi 5 percentili: 268 titoli pari a 63,51%

# Percentili di ranking – dati aggregati

## IMMEDIACY INDEX

---

Si tratta di dati preliminari, non ancora sottoposti a referaggio, da usare solo dietro esplicito consenso da richiedere agli autori ([elena.giglia@unito.it](mailto:elena.giglia@unito.it), [sandra.migliore@unito.it](mailto:sandra.migliore@unito.it))



Primi 5 percentili: 161 titoli pari a 38,15%  
Ultimi 5 percentili: 261 titoli pari a 61,85%



## Qualche riflessione...

---

- ✓ in ogni categoria ci sono uno o più titoli OA nei primi due percentili
- ✓ le tabelle di Immediacy Index – pur con tutti i limiti evidenziati da ISI - rivelano in media dati più alti rispetto all'IF: è vero che le riviste OA entrano in circolo prima
- ✓ riviste anche molto recenti mostrano già buoni indici di IF, o entrano direttamente il primo anno nei primi percentili (su 56 media ingresso: 7°; 3 nel 1°; 5 nel 2°)
- ✓ mediamente, le riviste OA hanno un andamento nel trend di IF meno lineare delle riviste tradizionali, con incrementi più sostanziali di anno in anno




# Cosa resta da fare...

---

Nuove prospettive di ricerca:

- ✓ **sul vantaggio citazionale:**
  - OA è giovane...condurre di nuovo gli stessi studi fra qualche anno, quando si vedrà l'effetto dei mandati a depositare
  - attendere esiti del progetto [PEER](#) - Publishing and the Ecology of European Research, un'analisi degli effetti di un massiccio selfarchiving concordato con editori (7°FP EU)
- ✓ **sulle citazioni:**
  - testare SciMago
  - testare Eigenfactor
  - testare IF sui 5 anni...



Grazie per l'attenzione!

---

Domande?

Elena Giglia  
Università degli Studi di Torino  
[elena.giglia@unito.it](mailto:elena.giglia@unito.it)

Maria Laura Vignocchi  
Università degli Studi di Bologna  
[vignocchi@cib.unibo.it](mailto:vignocchi@cib.unibo.it)