

*La biblioteca digitale: statistiche d'uso*

*Studi ed esperienze al confronto*



Roma, Palazzo dei Congressi – 27 Ottobre 2004

---

**Strumenti di indagine conoscitiva sulle classi di utenza di una biblioteca digitale basati sull'analisi quantitativa delle informazioni di accesso**

*Ugo Contino – [ugo.contino@caspur.it](mailto:ugo.contino@caspur.it)*

# Sommario

---

- ❑ Il consorzio CIBER e l'Emeroteca Virtuale del CASPUR: una breve introduzione
- ❑ Le scelte contrattuali del consorzio CIBER per l'editore Elsevier
- ❑ Un'indagine quantitativa sugli accessi alle testate Elsevier: obiettivi, strategie di analisi e risultati
- ❑ Analisi degli accessi in funzione delle categorie e dell'anno di pubblicazione delle riviste Elsevier presenti sull'Emeroteca
- ❑ Conclusioni

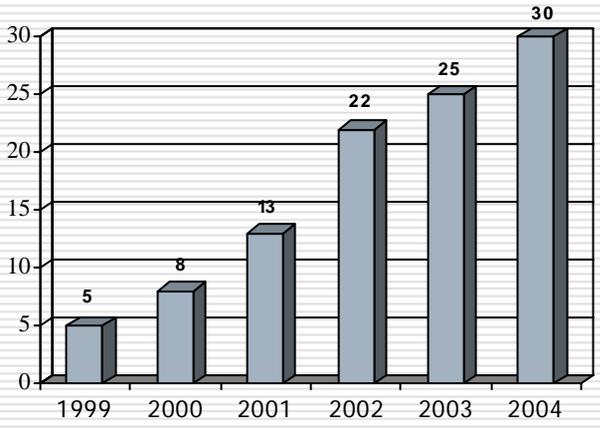
# Storia del CIBER

## Coordinamento **I**nteruniversitario **B**anche dati ed **E**ditoria in **R**ete

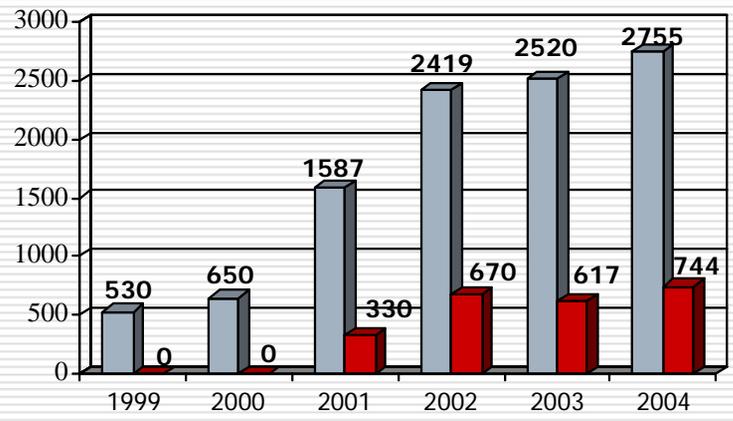
---

- Nato verso la fine degli anni '90
- Scopi:
  - **cooperare** per la costituzione e lo sviluppo di biblioteche digitali (informazioni scientifiche su matrice elettronica) nelle Istituzioni partecipanti e per promuovere la crescita professionale del loro personale tecnico-bibliotecario;
  - **facilitare** l'acquisizione, da terzi, di servizi bibliografici e documentari in rete;
  - **agevolare** il trasferimento di informazioni e di servizi tra le Istituzioni partecipanti;
  - **cooperare** alla scelta e la definizione delle risorse da sviluppare, acquistare o affittare;
  - **identificare** le tecnologie informatiche necessarie, al miglior rapporto costi/benefici;
  - **curare** lo sviluppo di prodotti per la consultazione e la archiviazione dell'informazione scientifica in formato elettronico;
  - **richiedere** alle autorità competenti, a nome delle Istituzioni partecipanti, il sostegno necessario per dare alle Biblioteche digitali le risorse di cui hanno bisogno per perseguire i loro fini.
- CASPUR identificato come *partner tecnologico* ed *amministrativo* del coordinamento per ciò che concerne il servizio di D.L.

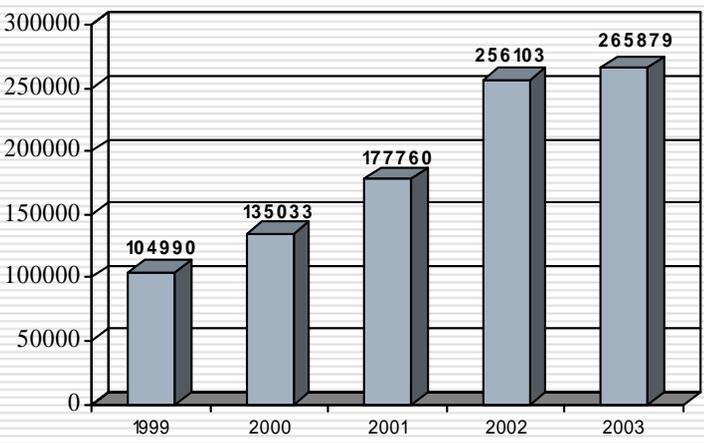
# 1999-2004: cosa è cambiato?



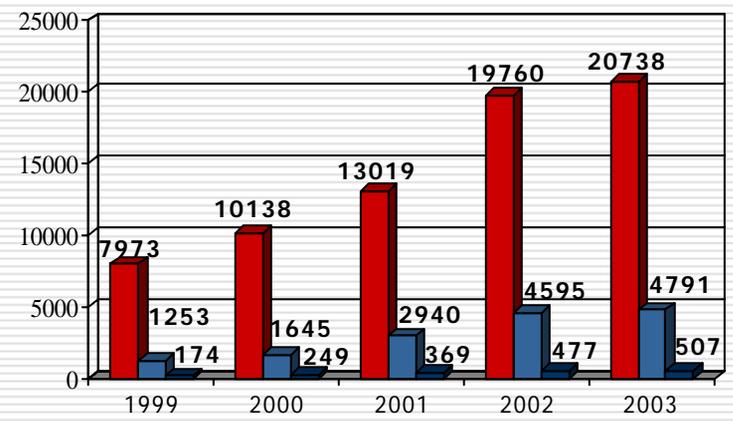
Numero di atenei/enti partecipanti al coordinamento CIBER (agg. Ottobre 2003)



Full Text Metadata (agg. ottobre 2004)



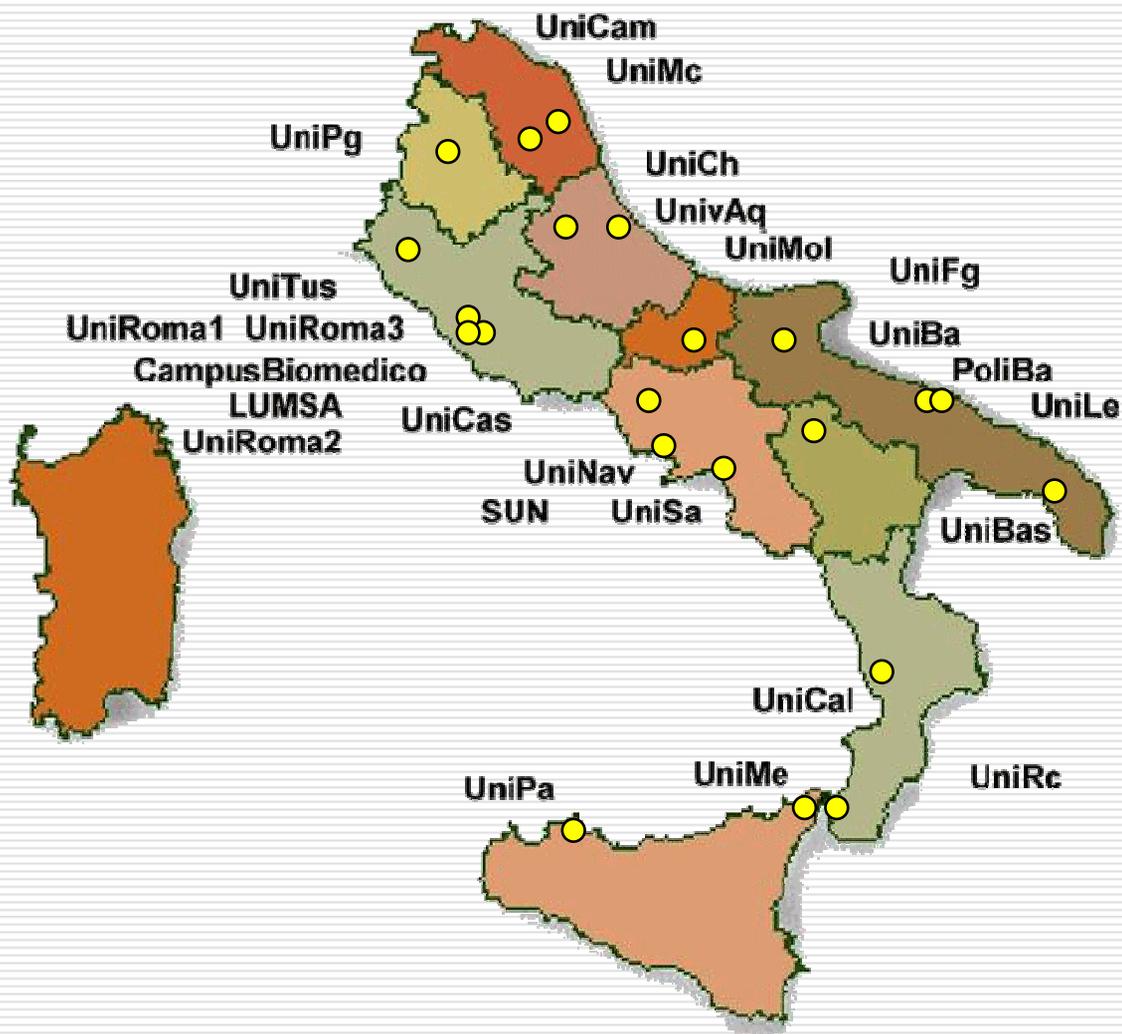
Studenti iscritti; Full Time Equivalent (fonte MIUR; dati '99)



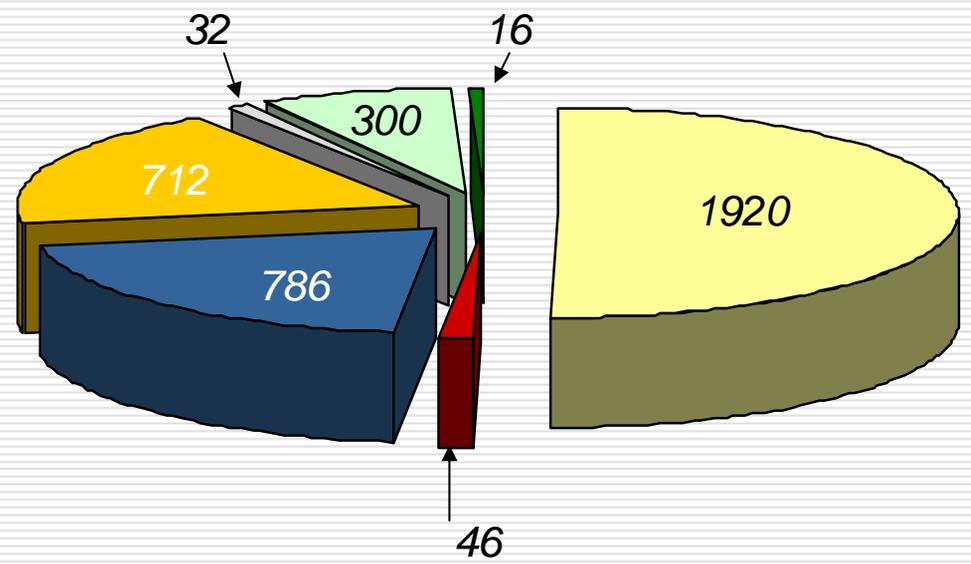
Professori Prof. a contratto lettori

(fonte MIUR; dati '02)

# Dove sono i nostri utenti?



# Scholarly e-journals: quale offerta?



- Elsevier (ft) dal 1995
- IOP(ft) dal 1991
- Kluwer (ft) dal 1995
- Blackwell dal 1996/1997
- ACS dal 1979
- Wiley (\*) dal 1996/1997
- Nature (\*) dal 1996/1997

(\*) *sul sito dell'editore*

Le riviste dell'editore Elsevier, oltre ad essere le più numerose sul sito dell'Emeroteca, sono anche le uniche per le quali l'accesso è possibile solo sul server del CASPUR

## Accesso alle riviste elettroniche di Elsevier

---

- Ultimo contratto stipulato nel 2003
- Durata triennale
- Modalità di accesso *OnSite* (SDOS)
- Due tipologie di accesso approvate in seno CIBER
  - *SDOS-UTL* : accesso alle riviste elettroniche possedute in cartaceo + quelle della Unique Title List (UTL) non comprese tra gli abbonamenti cartacei
  - *SDOS-Freedom* : accesso a tutte le riviste elettroniche di Elsevier
- Nel 2002 era in vigore un contratto di accesso alla collezione completa

## Accesso alle riviste elettroniche di Elsevier

---

- 10 (-1) enti/università CIBER "in UTL"
- 13 (+1) con contratto *Freedom*
- Concertate insieme all'editore fasce percentuali di pagamento a seconda del tipo di contratto
- Unique Title List composta da circa 740 riviste scelte in base a parametri di *impact factor* e di volume di accessi per l'anno 2002
- L'applicazione delle nuove politiche contrattuali ha imposto una ristrutturazione dei filtri di accesso alle riviste per gli enti che hanno scelto la UTL ("in UTL")
- Esiste una *finestra temporale* prima dell'entrata a regime del contratto dove valevano le vecchie condizioni (accesso consentito a tutti gli enti all'intero pacchetto)

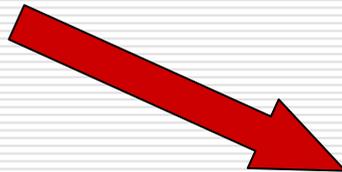
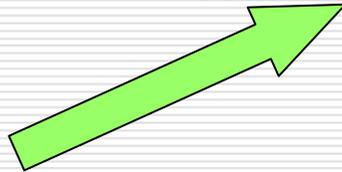
# Politica di accesso alle testate Elsevier (1)

Accesso ad una rivista NON in abbonamento  
ne in UTL

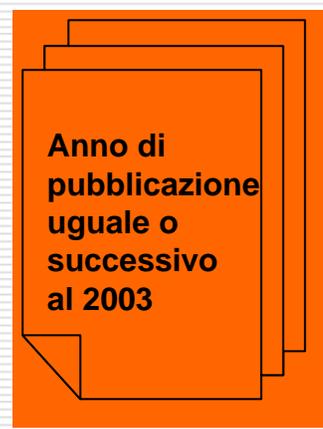


Utente da ente/università  
con contratto *UTL*  
("utente in *UTL*")

DEVE essere consentito  
l'accesso



DEVE essere proibito  
l'accesso



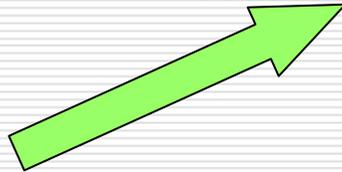
# Politica di accesso alle testate Elsevier (2)

Accesso ad una rivista NON in abbonamento  
ne in UTL

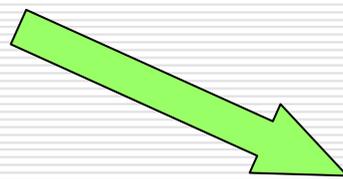


Utente da ente/università  
con contratto *Freedom*  
("utente in Freedom")

DEVE essere consentito  
l'accesso



DEVE essere consentito  
l'accesso



## Obiettivi dell'indagine

---

- Analizzare la distribuzione degli articoli scaricati in funzione delle due *classi* di utenza (*UTL* e *Freedom*)
- Stimare la distribuzione degli accessi sui diversi gruppi di testate (in abbonamento, appartenenti al gruppo *UTL*, appartenenti al gruppo *Freedom*) per le due classi di utenza
- Determinare la distribuzione degli accessi in funzione della *categorie delle riviste* e *dell'anno di pubblicazione* del fascicolo al quale l'articolo acquisito appartiene



## Analisi degli accessi: scelta del campione

---

- Sono state analizzate unicamente le testate Elsevier perché accessibili solo sull'Emeroteca Virtuale (completezza del campione)
- Sono stati considerati solo gli accessi agli articoli *full-text in formato PDF* (maggioranza del campione)
- Relativamente agli abbonamenti cartacei sono stati utilizzati i dati del posseduto Elsevier per l'anno 2003 dei vari enti/atenei analizzati

## Analisi degli accessi: scelta dei tempi

---

- Gli intervalli temporali scelti per l'analisi sono stati:

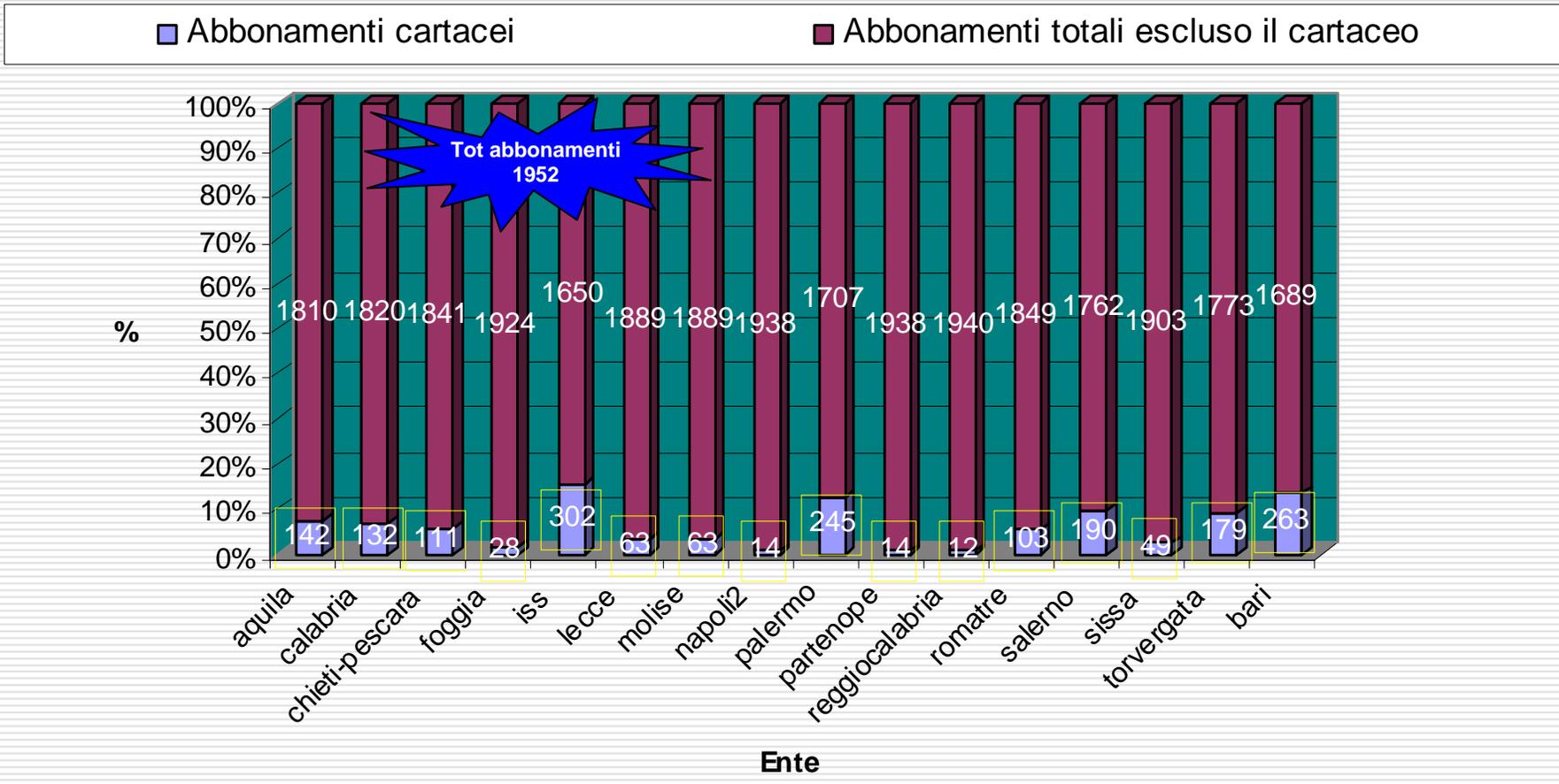
15 aprile -15 ottobre 2003 (T\_A)

1 gennaio – 31 giugno 2004 (T\_B)

- Durante T\_A non era ancora pienamente operativo il filtro sugli accessi, ed il set delle riviste Elsevier era in via di completamento

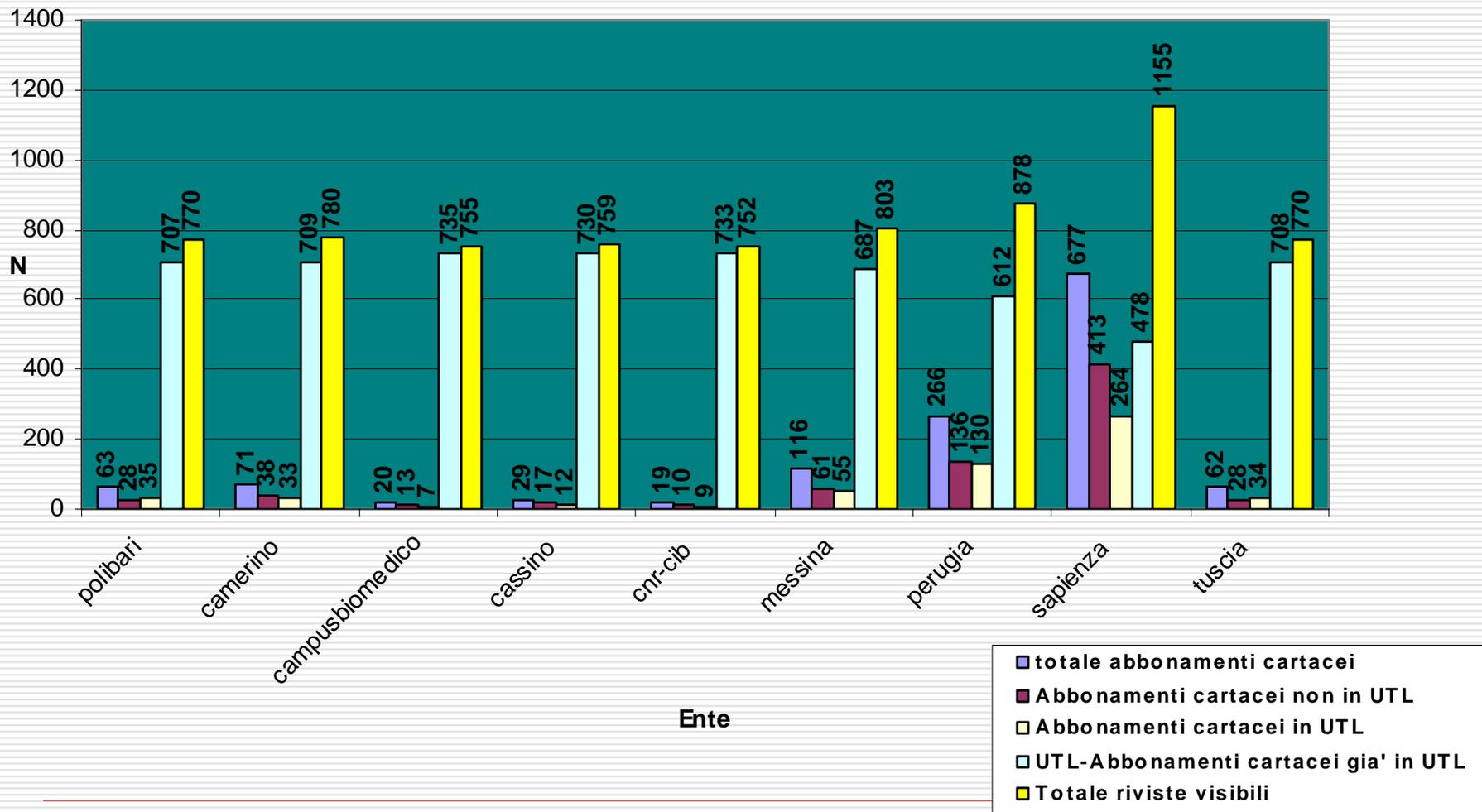
# Caratterizzazione degli abbonamenti (1)

Distribuzione % abbonamenti elettronici per enti con contratto Elsevier "Freedom"

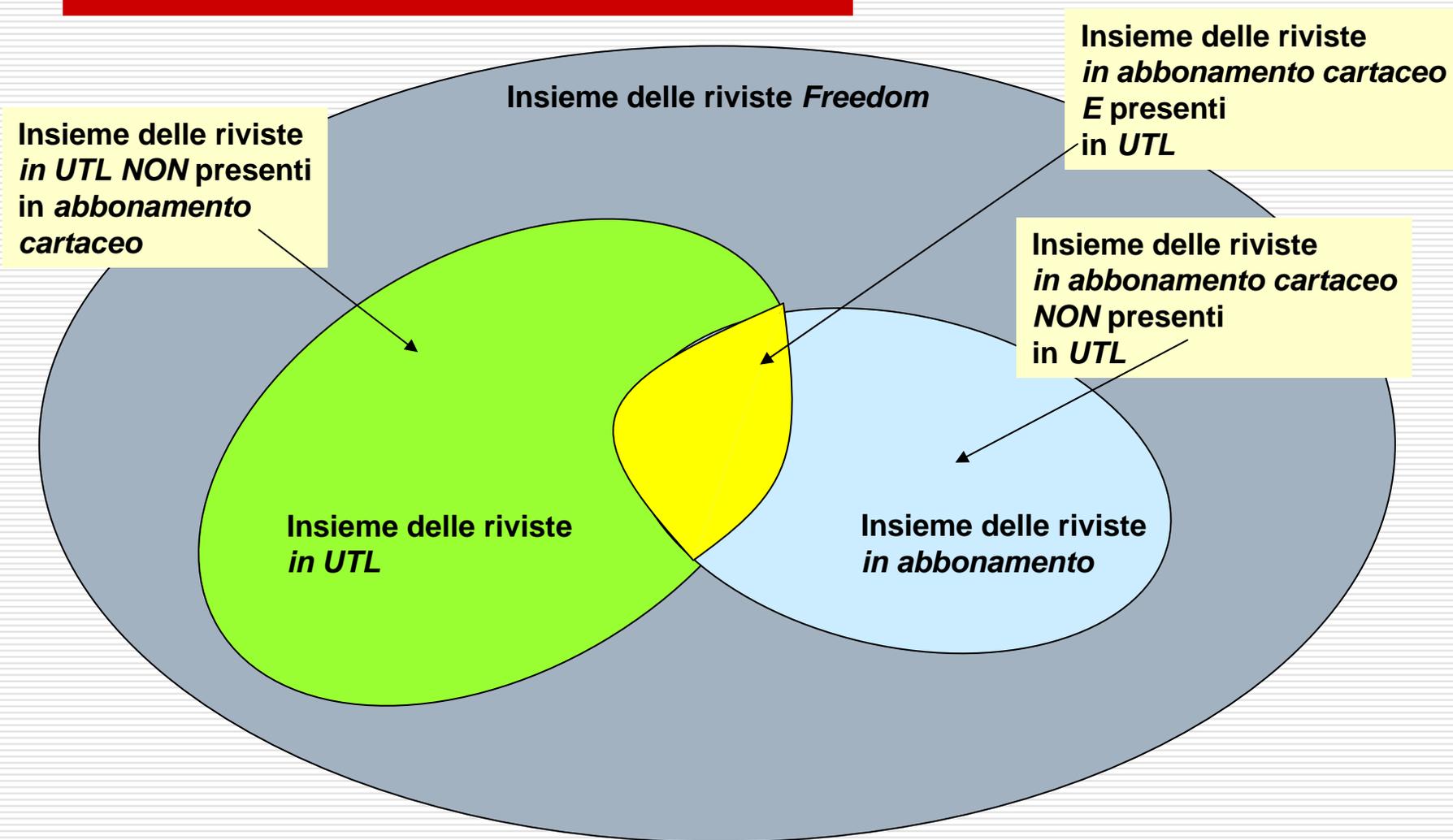


# Caratterizzazione degli abbonamenti (2)

Distribuzione abbonamenti per enti con contratto E Isevier-UTL

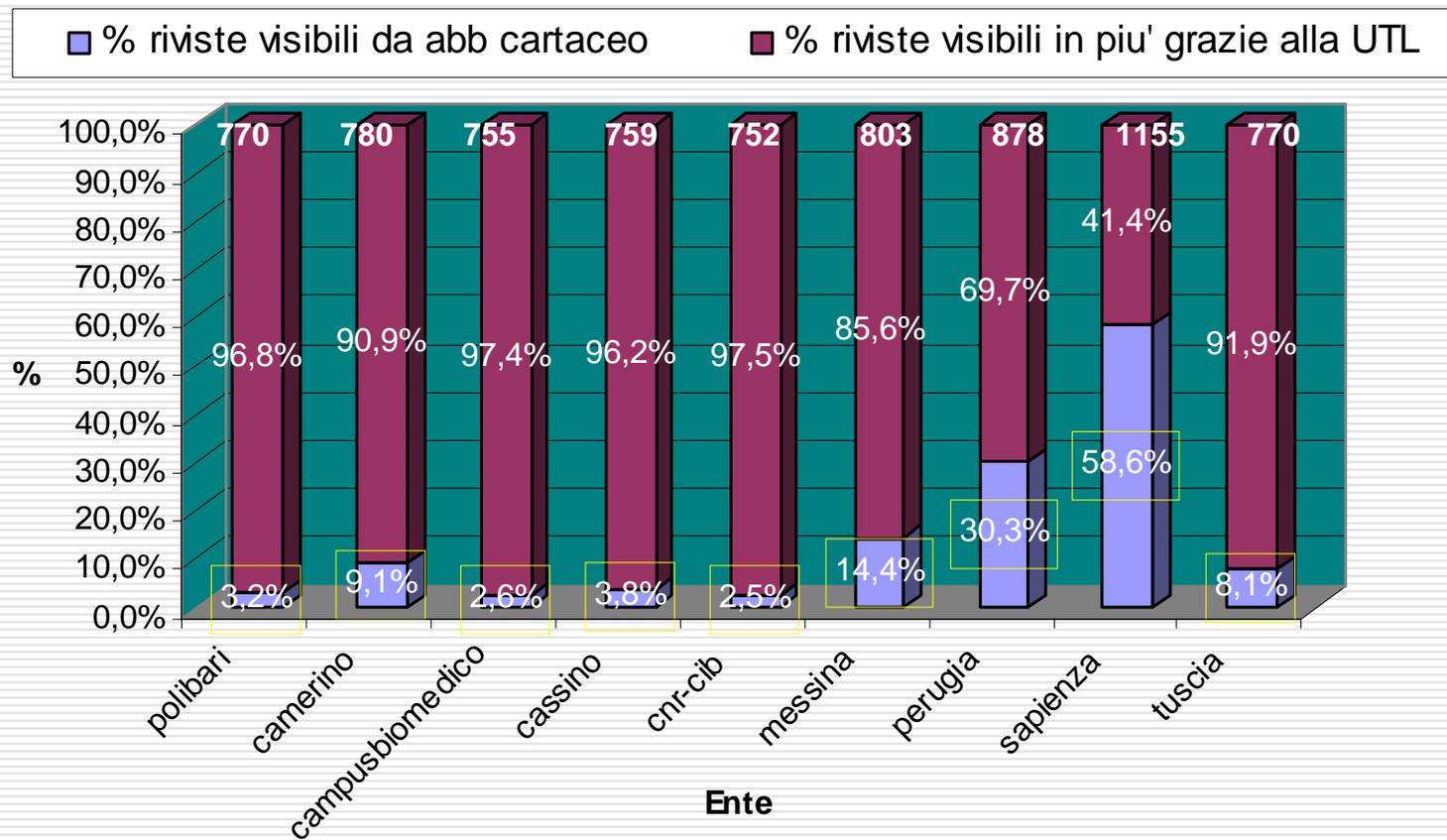


# Caratterizzazione degli abbonamenti (3)



# Accessibilità in UTL

## Riviste accessibili con contratto Elsevier-UTL



## Qualche cifra...

---

- ❑ Enti *UTL* analizzati nel periodo T\_A: 9
- ❑ Enti *Freedom* analizzati nel periodo T\_A: 14
- ❑ Enti *Freedom* analizzati nel periodo T\_B: 10
- ❑ Articoli totali analizzati per il periodo T\_A: 287.354 di cui 165.649 da enti *Freedom* e 121.705 da enti *UTL*
- ❑ Articoli totali analizzati per il periodo T\_B: 258.221

# Criteri di analisi

---

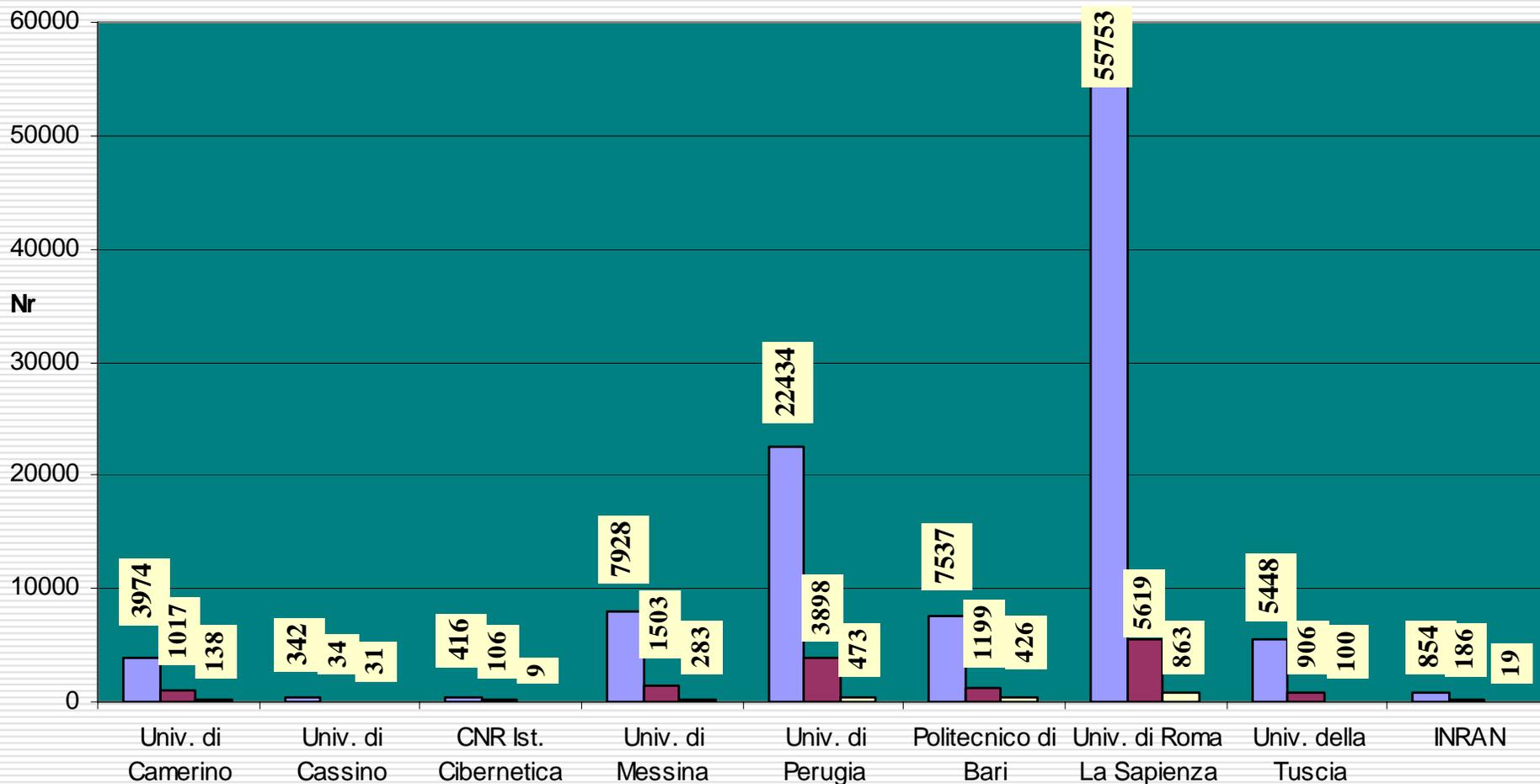
- Un articolo appartenente ad una rivista pubblicata prima del 2002 è conteggiato sempre come *in abbonamento* (parametro IA)
- Se l'articolo proviene da una rivista pubblicata nel 2003 se ne controlla la sua appartenenza ad una delle seguenti 3 categorie:
  - In abbonamento (parametro IA)
  - In UTL (parametro IUTL)
  - Fuori abbonamento (parametro FA)
- L'analisi è fatta per il primo periodo (T\_A) sulle 2 tipologie di enti; nel secondo (T\_B) si analizzano i soli enti con abbonamento *Freedom*

---

**Periodo di analisi T\_A: 15 aprile – 15 ottobre 2003**

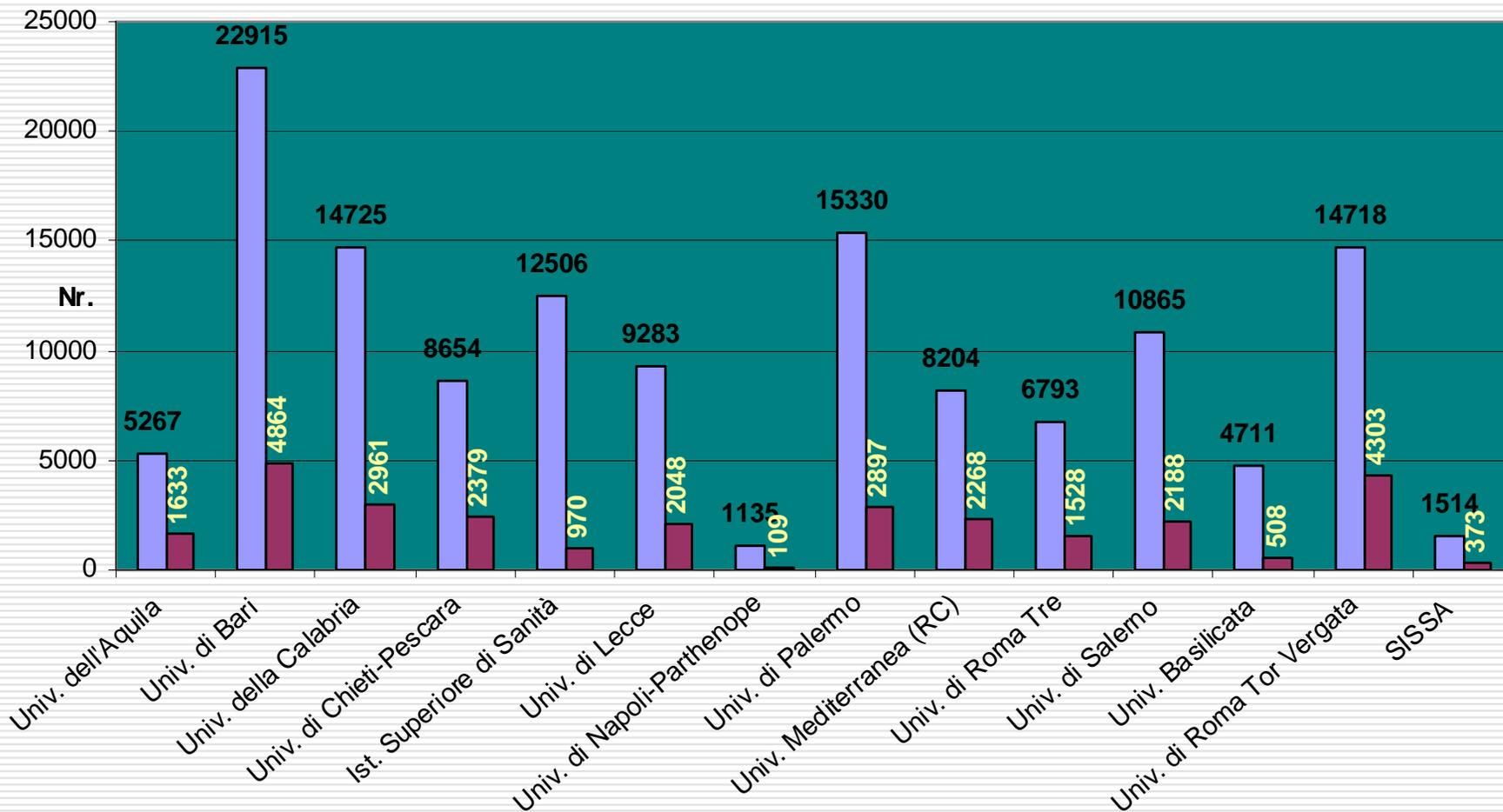
# Distribuzione degli articoli scaricati nel periodo: 15 aprile – 15 ottobre 2003 – *Enti in UTL*

- Nr. Articoli scaricati da riviste con anno di pubblicazione inferiore al 2003 o in abbonamento cartaceo (IA)
- Nr. Articoli scaricati da riviste in abbonamento UTL e con anno di pubblicazione uguale al 2003 (IUTL)
- Nr. Articoli scaricati da riviste non in abbonamento (cartaceo o UTL) con anno di pubblicazione uguale al 2003 (FA)



# Distribuzione degli articoli scaricati nel periodo: 15 aprile – 15 ottobre 2003 – *Enti in Freedom*

■ Nr. Totale articoli scaricati da riviste in abbonamento cartaceo o con anno di pubblicazione inferiore al 2003 (IA)  
 ■ Nr. Totale articoli scaricati da riviste non in abbonamento cartaceo e con anno di pubblicazione uguale al 2003 (FA)



## Calcolo % di accessi "fuori UTL"

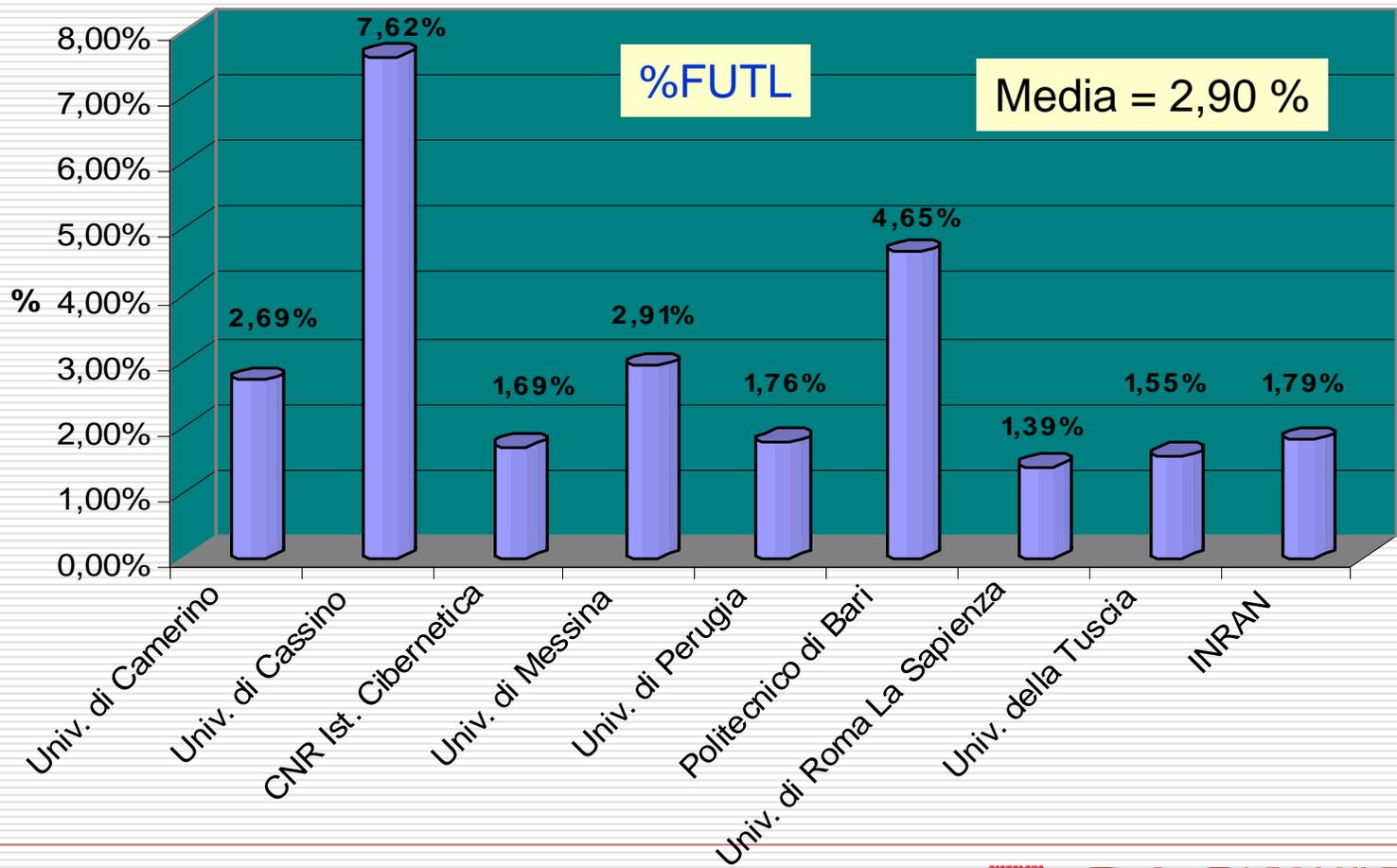
---

- Dato utile per stimare la bontà del *set* di riviste scelto per un determinato ente con contratto *in UTL*
- Per il calcolo si è utilizzato il seguente parametro:

$$\%FUTL = \frac{FA}{(IA + IUTL + FA)}$$

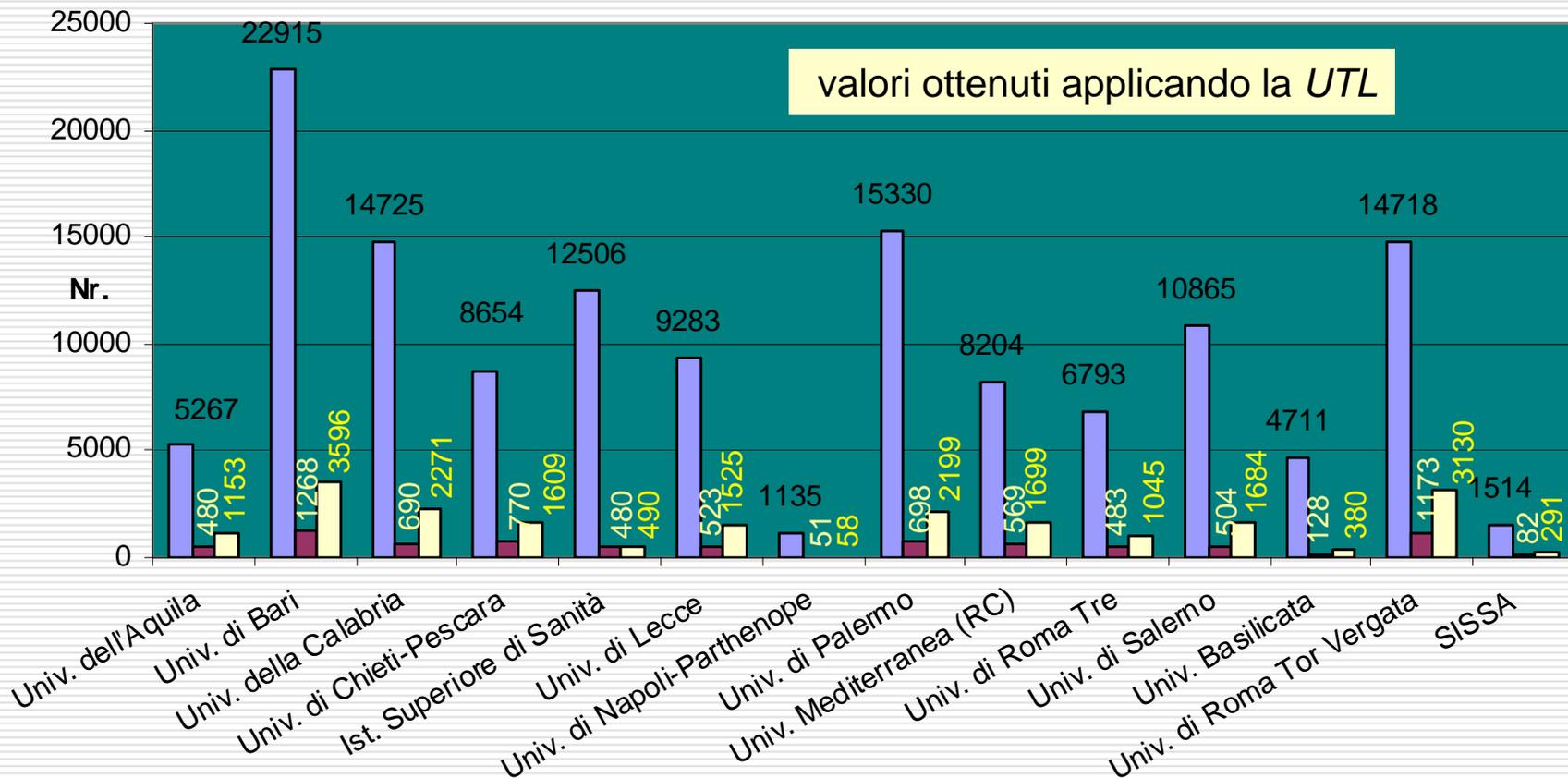
# Enti CIBER con abbonamento UTL

■ % articoli scaricati da riviste fuori abbonamento e con anno di pubblicazione uguale al 2003:  $FA/(IA+IUTL+FA)$



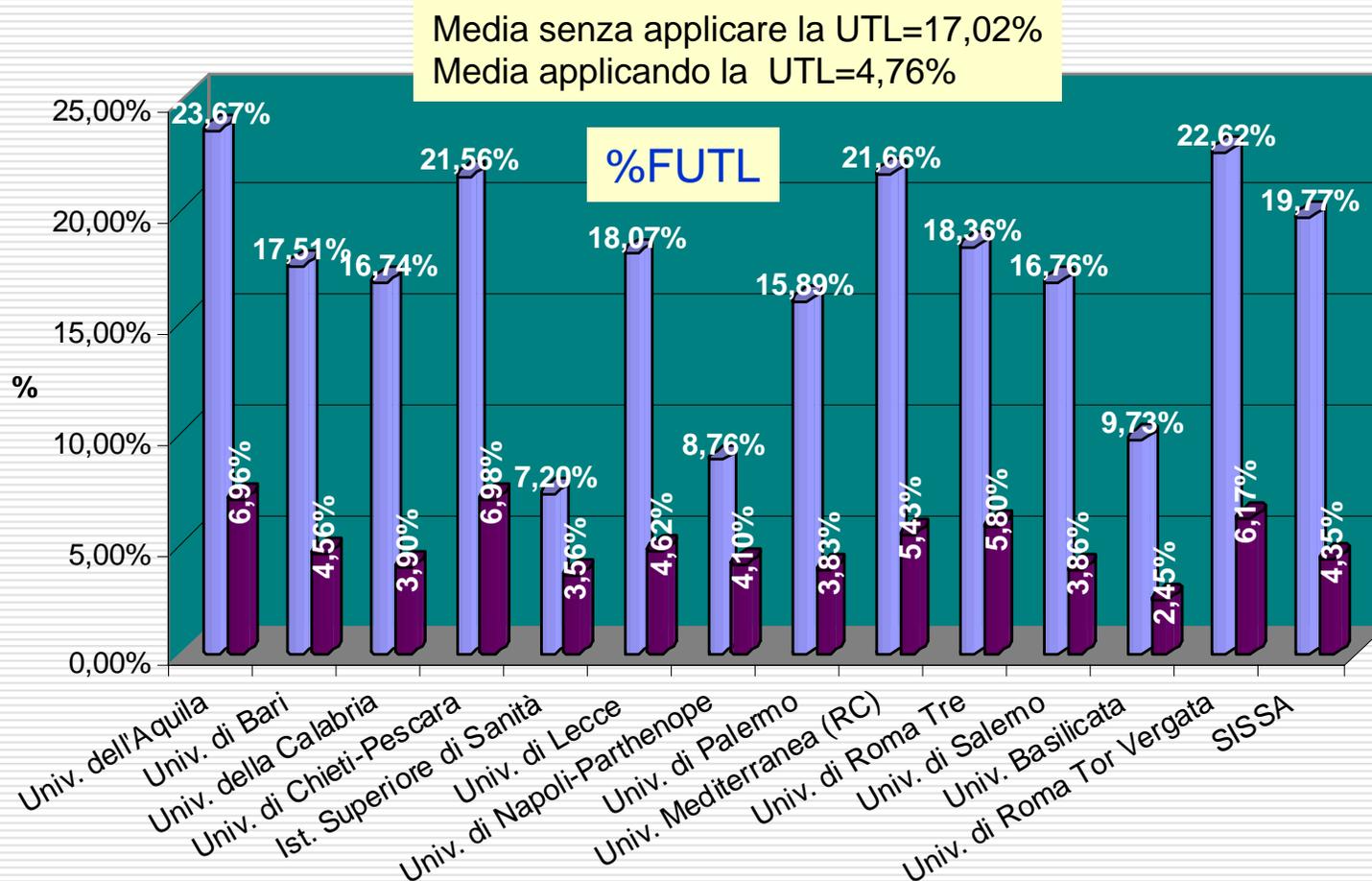
## Confronto con gli enti con abbonamento *Freedom* (1)

- Nr. Totale articoli scaricati da riviste in abbonamento cartaceo o con anno di pubblicazione inferiore al 2003 (IA)
- Nr. Totale articoli scaricati da riviste non in abbonamento cartaceo e con anno di pubblicazione uguale al 2003 (FA)
- Nr. Totale articoli scaricati da riviste appartenenti alla UTL e con anno di pubblicazione uguale al 2003 (IUTL)



## Confronto con gli enti con abbonamento *Freedom* (2)

- % articoli scaricati da riviste fuori abbonamento e con anno di pubblicazione uguale al 2003: FA/(IA+FA)
- % articoli scaricati da riviste fuori abbonamento e con anno di pubblicazione uguale al 2003 applicando il filtro UTL: FA/(IA+IUTL+FA)



---

**Periodo di analisi T\_B: 1 gennaio – 31 giugno 2004**

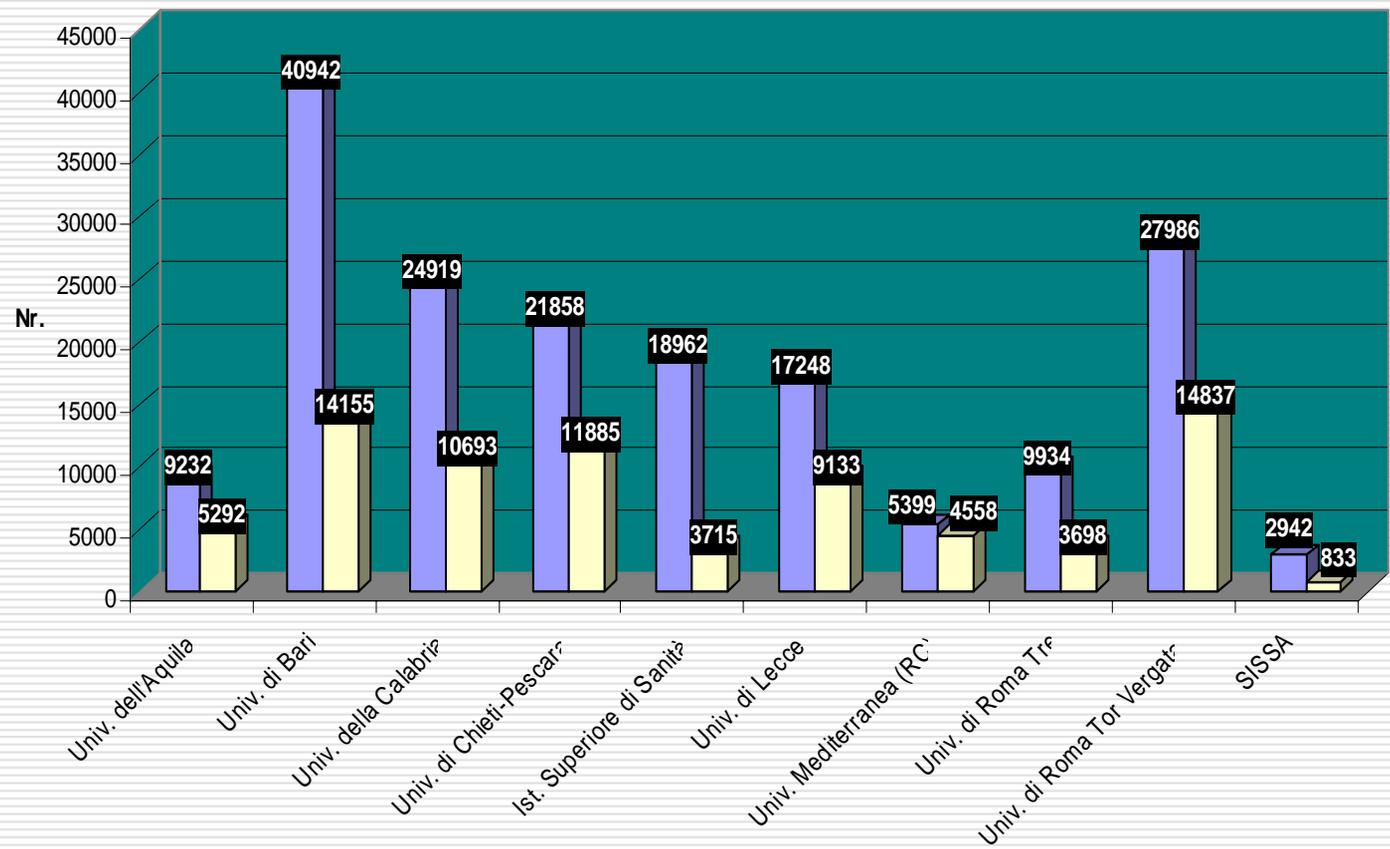
## Considerazioni generali

---

- A parità di *condizioni al contorno* (continuità del servizio, quantità di riviste offerte, tempi di risposta del servizio, etc.) non dovrebbero sussistere sostanziali differenze nei parametri di accesso calcolati in differenti periodi temporali
- Il comportamento del *campione utenza* dovrebbe quindi essere lo stesso per i due periodi di analisi (almeno per piccole differenze temporali)

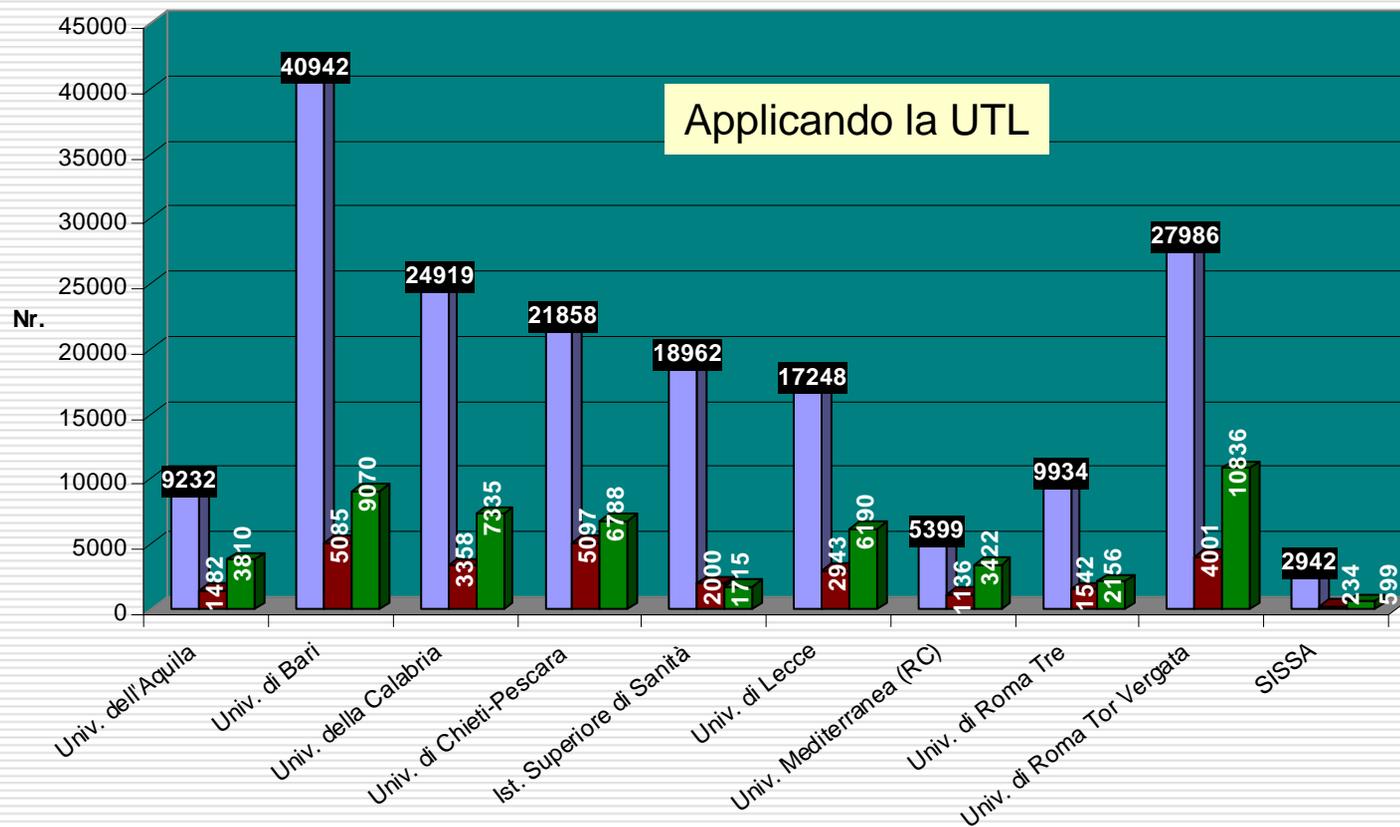
# Totale articoli scaricati in abbonamento (IA) e fuori abbonamento (FA) per enti con abbonamento *Freedom* (1)

- Nr. Totale articoli scaricati da riviste in abbonamento cartaceo o con anno di pubblicazione inferiore al 2003 (IA)
- Nr. Totale articoli scaricati da riviste non in abbonamento cartaceo e con anno di pubblicazione superiore o uguale al 2003 (FA)



## Totale articoli scaricati in abbonamento (IA) e fuori abbonamento (FA) per enti con abbonamento *Freedom* (2)

- Nr. Totale articoli scaricati da riviste in abbonamento cartaceo o con anno di pubblicazione inferiore al 2003 (IA)
- Nr. Totale articoli scaricati da riviste non in abbonamento cartaceo ne in UTL e con anno di pubblicazione superiore o uguale al 2003 (FA)
- Nr. Totale articoli scaricati da riviste non in abbonamento cartaceo ma in UTL e con anno di pubblicazione superiore o uguale al 2003 (IUTL)



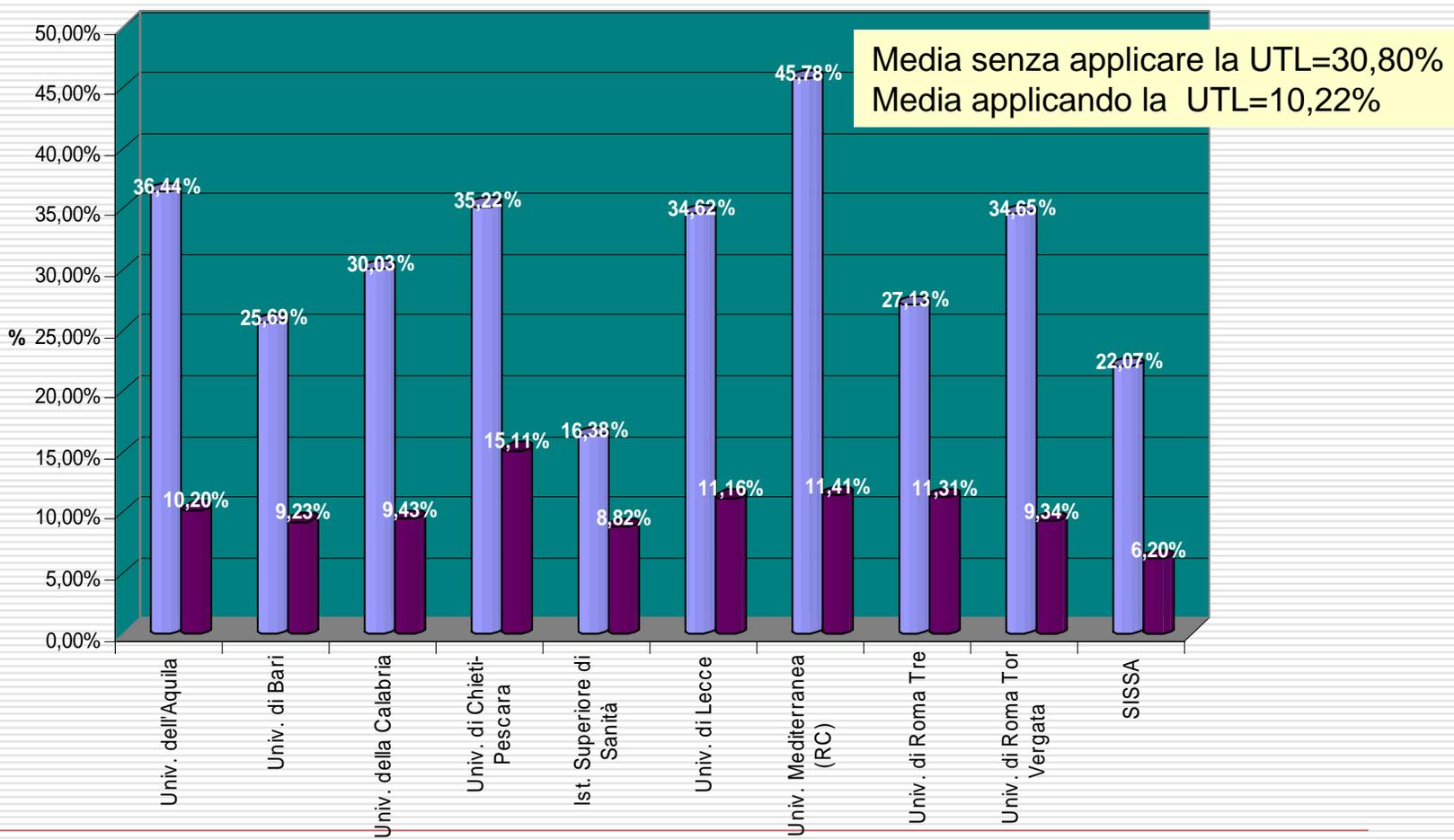
## Criteri di analisi per il periodo T\_B

---

- Un articolo appartiene ad una rivista pubblicata prima del 2002 è conteggiato sempre come *in abbonamento* (parametro IA)
- Se l'articolo proviene da una rivista pubblicata a partire dal 2003 se ne controlla la sua appartenenza ad una delle seguenti 3 categorie:
  - In abbonamento (parametro IA)
  - In UTL (parametro IUTL)
  - Fuori abbonamento (parametro FA)

# %FUTL (accesso a riviste fuori UTL)

■ % articoli scaricati da riviste fuori abbonamento e con anno di pubblicazione superiore al 2003: FA/(FA+IA)  
■ % articoli scaricati da riviste fuori abbonamento e con anno di pubblicazione superiore al 2003 applicando il filtro UTL: FA/(FA+IA+IUTL)



# Confronto tra i valori di %FUTL per i due periodi di analisi

## □ Valori medi:

Tipologia ente	Periodo	Valore
Con abbonamento UTL	15 apr-15 ott 2003	2,90 %
Con abbonamento Freedom	15 apr-15 ott 2003	4,76%
Con abbonamento Freedom	1 gen-31 lugl 2004	10,22%

Applicando la UTL

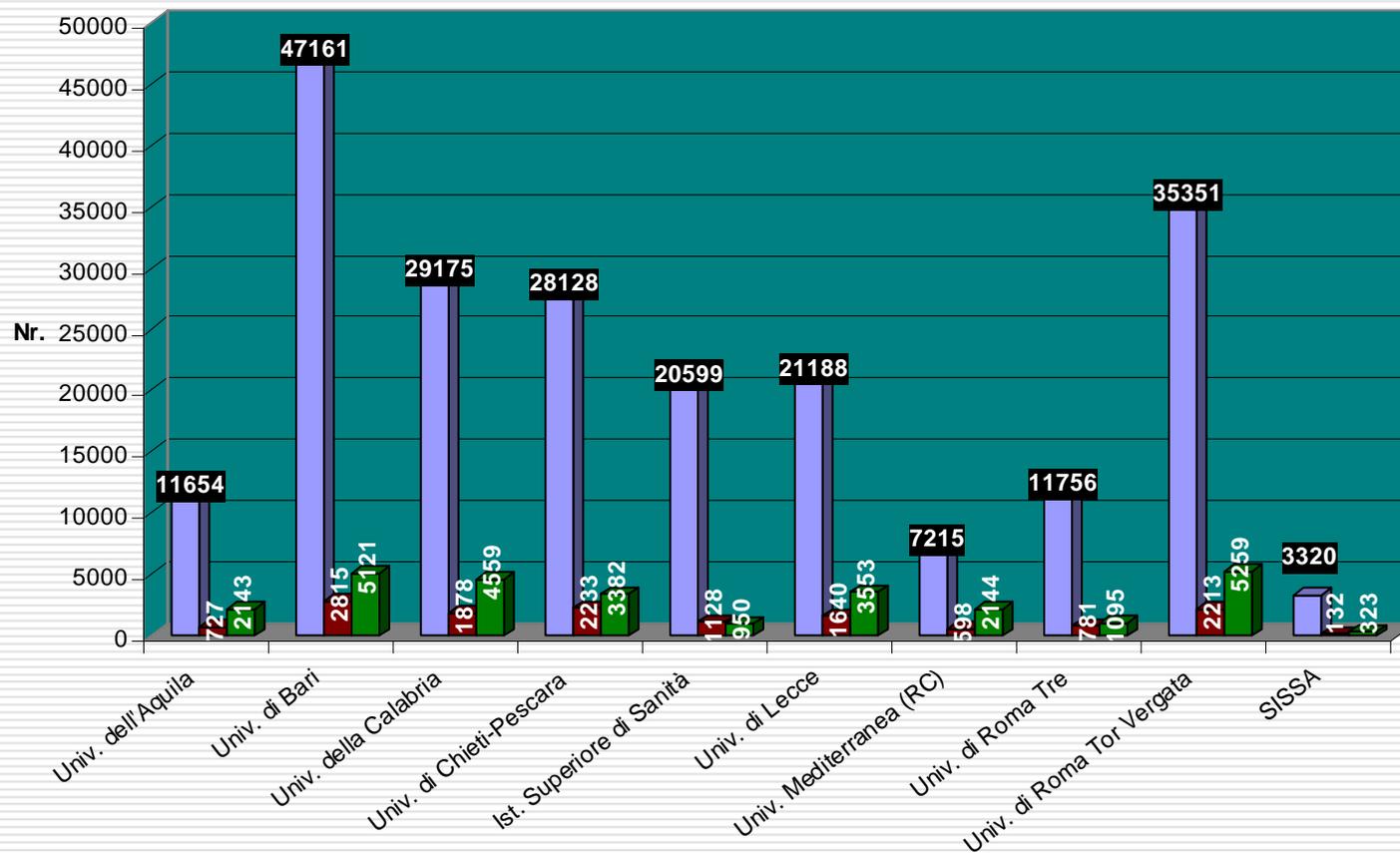
# Considerazioni generali

---

- Benchè i periodi di analisi abbiano durate simili, varia l'intervallo di tempo entro il quale si effettua il controllo "rivista in abbonamento" - "rivista fuori abbonamento" (solo un anno nel primo caso -2003-; quasi due nel secondo -2003,2004)
- Effettuata una nuova analisi considerando (nel calcolo del parametro %FUTL) i due insiemi "riviste pubblicate prima del 2004" e "riviste pubblicate nel 2004" (nuovo criterio di analisi)

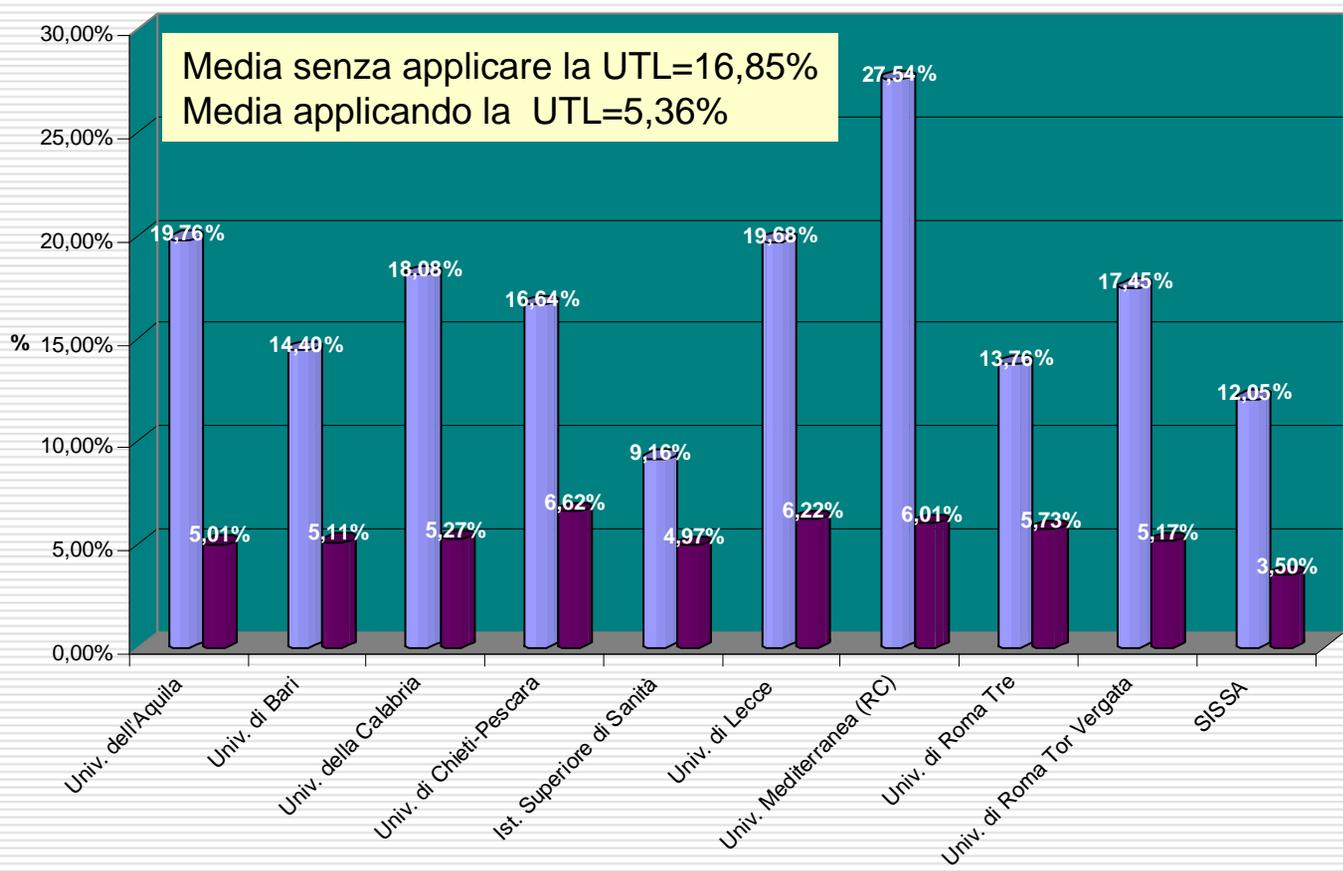
# Totale articoli scaricati in abbonamento (IA) e fuori abbonamento (FA) per enti con abbonamento *Freedom*

- Nr. Totale articoli scaricati da riviste in abbonamento cartaceo o con anno di pubblicazione inferiore al 2004 (IA)
- Nr. Totale articoli scaricati da riviste non in abbonamento cartaceo ne in UTL e con anno di pubblicazione uguale al 2004 (FA)
- Nr. Totale articoli scaricati da riviste non in abbonamento cartaceo ma in UTL e con anno di pubblicazione uguale al 2004 (IUTL)



# %FUTL (accesso a riviste fuori UTL)

■ % articoli scaricati da riviste fuori abbonamento e con anno di pubblicazione uguale al 2004 :  $FA/(FA+IA)$   
■ % articoli scaricati da riviste fuori abbonamento e con anno di pubblicazione uguale al 2004 applicando il filtro UTL:  $FA/(FA+IA+IUTL)$



Confronto tra i valori di %FUTL per i due periodi di analisi e considerando gli insiemi di riviste { <2003; =2003 } e { <2004; =2004 }

□ Valori medi:

Tipologia ente	Periodo	Valore
Con abbonamento UTL	15 apr-15 ott 2003	2,90 %
Con abbonamento Freedom	15 apr-15 ott 2003	4,76%
Con abbonamento Freedom	1 gen-31 lugl 2004	5,36%

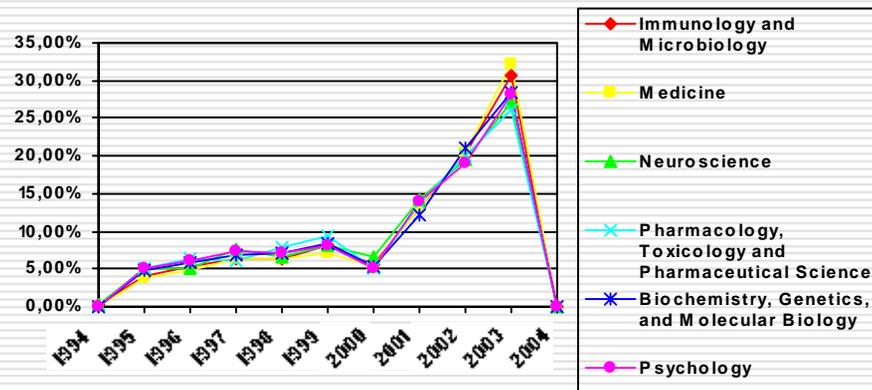
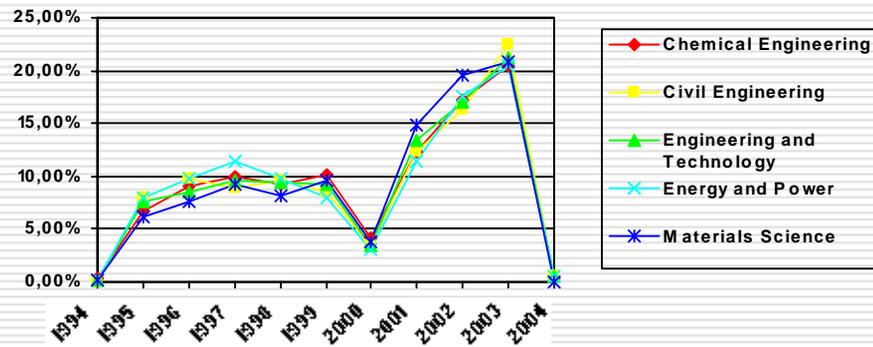
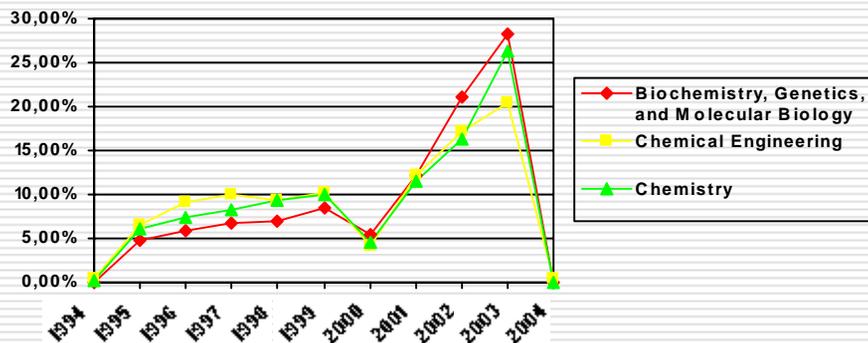
Applicando la UTL

# Analisi della distribuzione temporale del campione

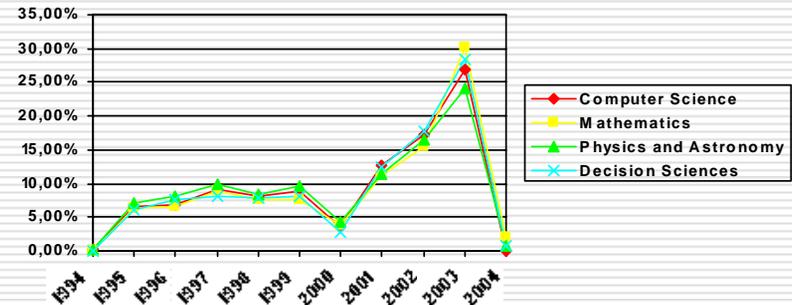
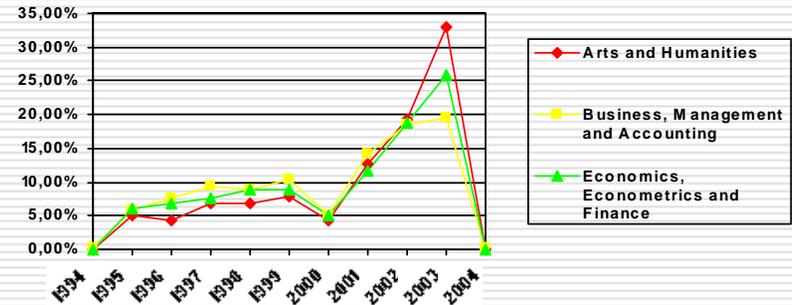
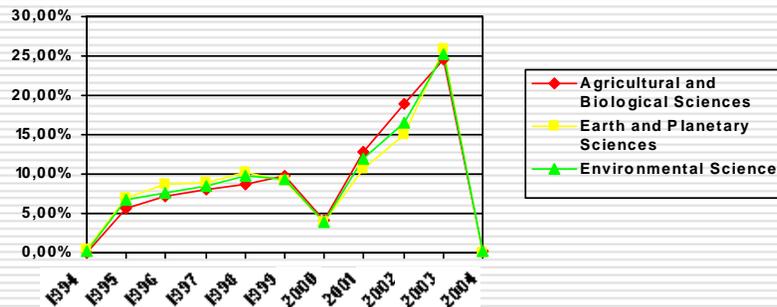
---

- Utile per capire come si distribuisce l'accesso alle riviste in funzione dell'anno di pubblicazione
- Verifica di un'eventuale correlazione con la categoria di appartenenza della rivista

# Distribuzione per categoria e anno di pubblicazione (periodo di analisi: 15 apr – 15 ott) – (1)

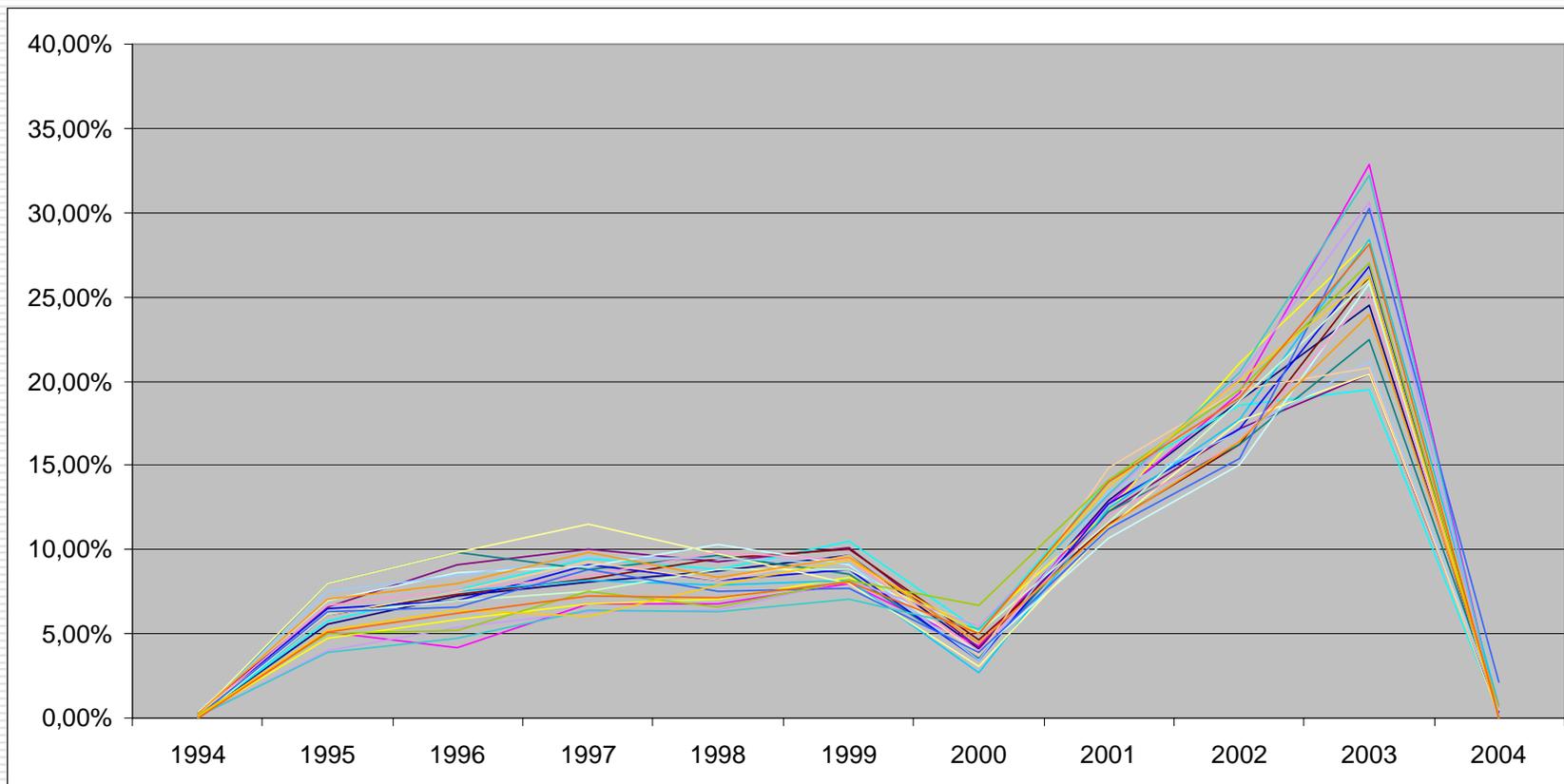


# Distribuzione per categoria e anno di pubblicazione (periodo di analisi: 15 apr – 15 ott) – (2)

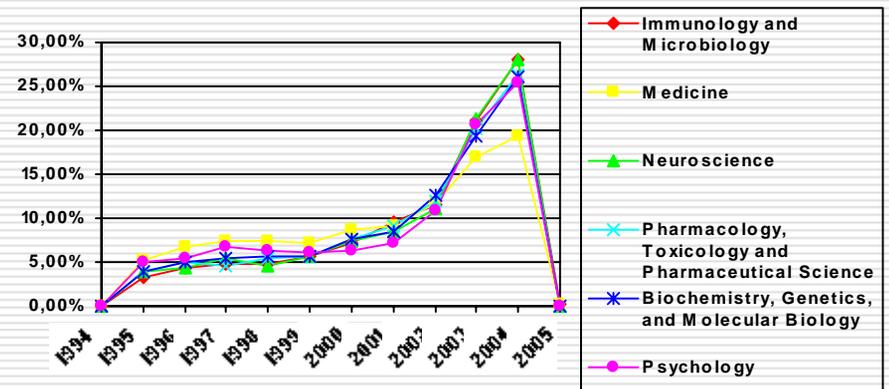
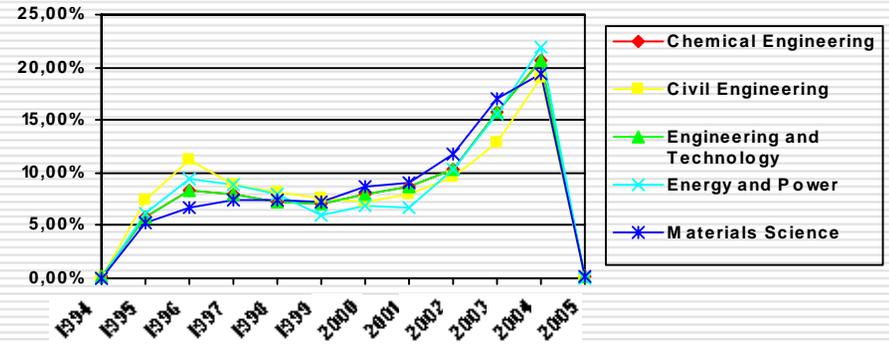
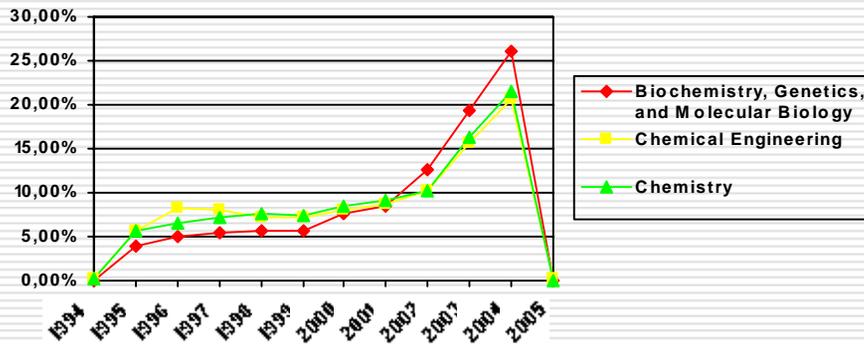


## Vista d'insieme (periodo di analisi: 15 apr – 15 ott)

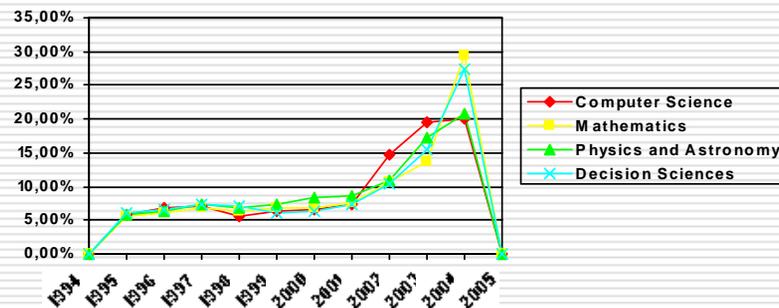
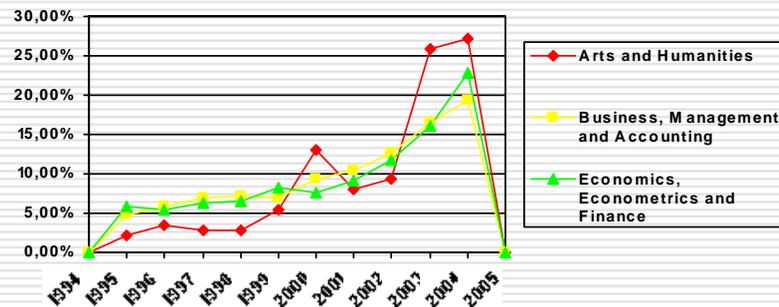
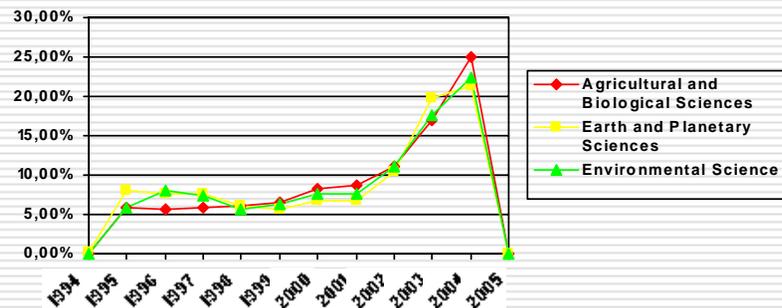
---



# Distribuzione per categoria e anno di pubblicazione (periodo di analisi: 1 gen – 31 lug) – (1)

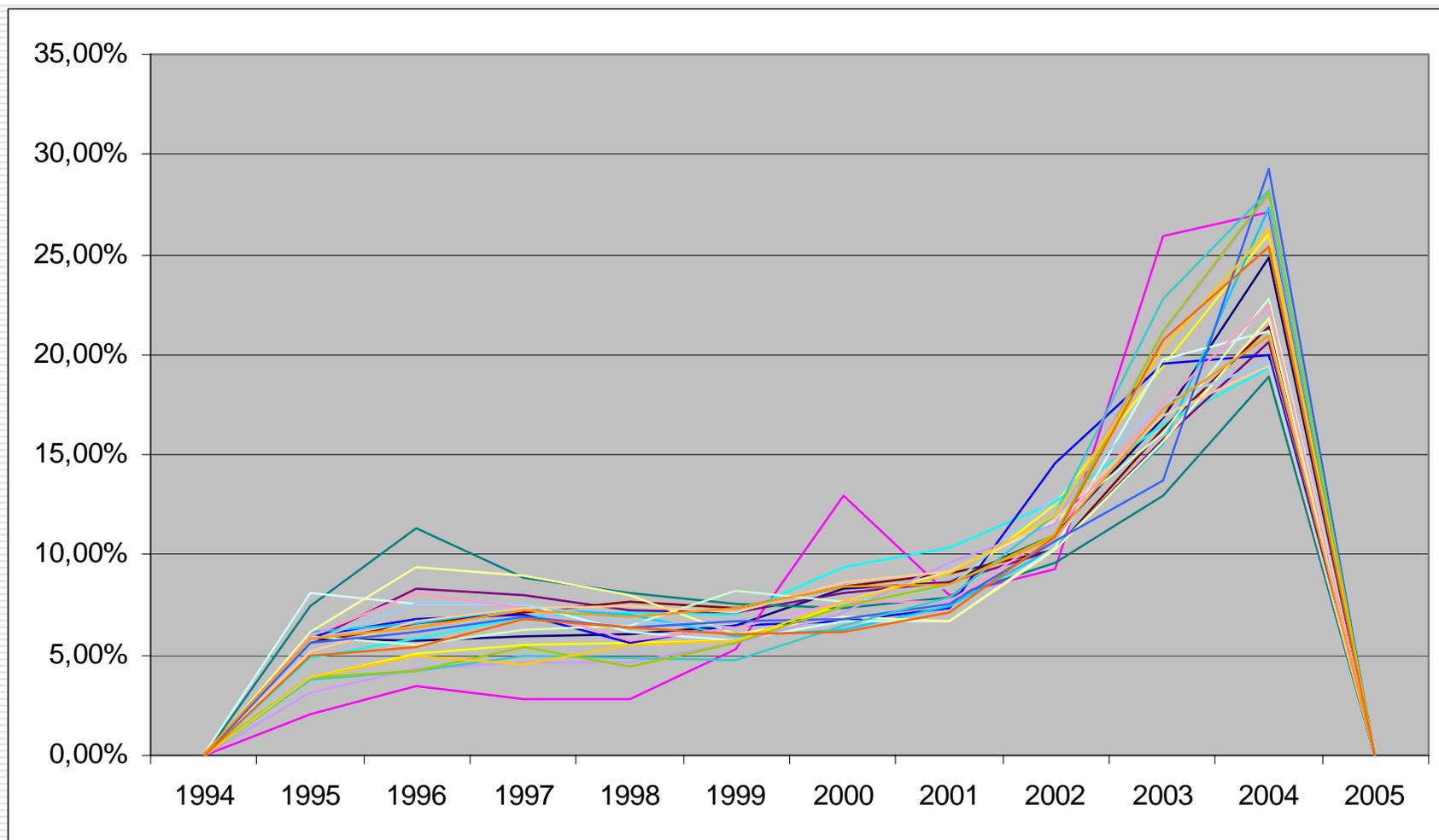


# Distribuzione per categoria e anno di pubblicazione (periodo di analisi: 1 gen – 31 lug) – (2)

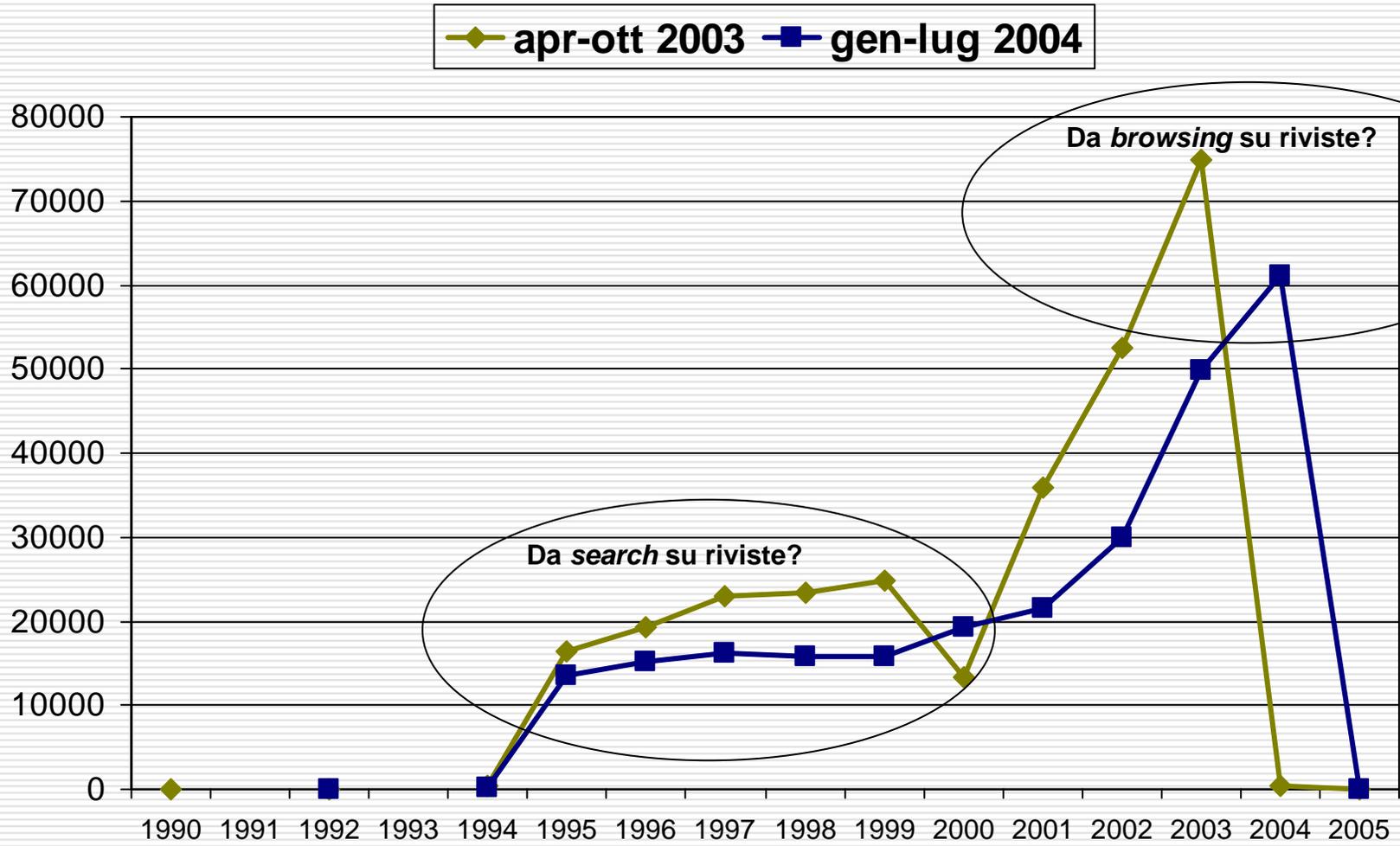


## Vista d'insieme (periodo di analisi: 1 gen – 31 lug)

---

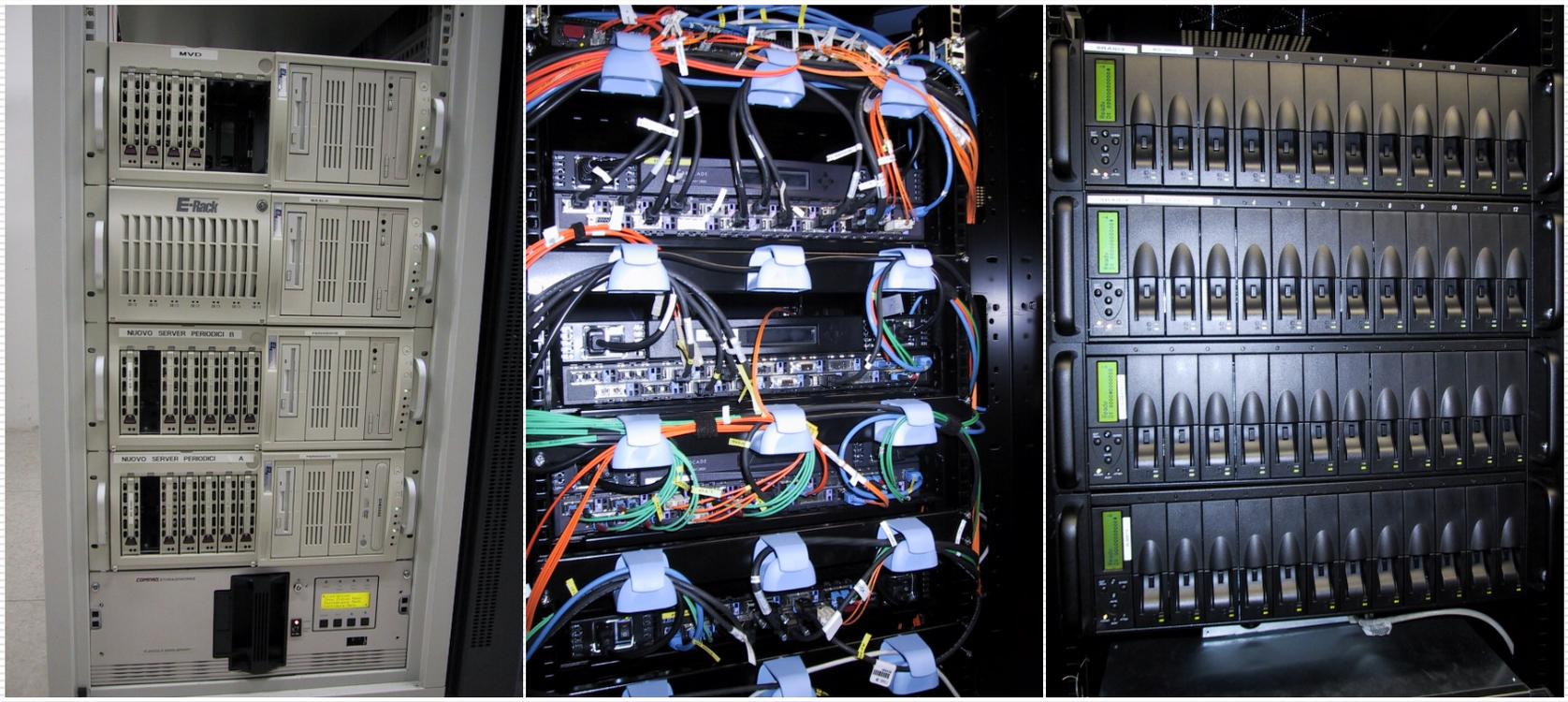


# Distribuzione per anno di pubblicazione



# Visti da vicino ...

---



# Conclusioni

---

- L'analisi delle **percentuali di accesso** ad articoli appartenenti a riviste *fuori abbonamento* e *fuori UTL* porta a risultati simili sia per enti *in UTL* che per enti *in Freedom* e **sembra scarsamente dipendente dal periodo di analisi** (di certo per intervalli di tempo vicini) a patto che vengano scelti criteri di indagine tra di loro coerenti
- La distribuzione degli **articoli scaricati in funzione dell'anno di pubblicazione e della categoria** di appartenenza della rivista, mostra **caratteri morfologici simili per tutte le categorie**, tanto più accentuati quanto più affini vengono scelte le categorie tra loro (effetto più evidente per T\_A che per T\_B)
- La **distribuzione cumulativa degli articoli scaricati** in funzione dell'anno di pubblicazione sembra **morfologicamente invariante** rispetto al **periodo di analisi** e potrebbe essere correlata con la **strategia di acquisizione dell'articolo** stesso (*search* o *browsing* sull'articolo)

---

*Grazie per la vostra attenzione!*

*[ugo.contino@caspur.it](mailto:ugo.contino@caspur.it)*

