



# Diffusione dell'Open Source nei Siti Web dei Comuni Italiani

*Giuseppe Destefanis, Roberto Tonelli, **Giulio Concas**, Michele Marchesi*



**GULCh**



2012  
**LINUX  
DAY**  
I T A L I A



## Introduzione

- Negli ultimi anni il numero e l'importanza delle applicazioni del software libero sono cresciuti oltre le più ottimistiche previsioni e hanno dato vita a numerosi modelli di business che hanno coinvolto le aree applicative più importanti.
- Anche molte istituzioni governative, nazionali e locali, hanno deciso di promuovere politiche per l'adozione di software Open Source. La scelta del software Open Source può rappresentare una soluzione per i numerosi problemi che si presentano nell'attuazione di una politica dell'ICT.
- La principale ragione è di ordine economico: il software libero consente notevoli risparmi sia nel capitolo degli investimenti, sia in quello dei costi di esercizio



## Introduzione

- Diversi studi sono stati condotti sulla diffusione dei tradizionali software Open Source, ma poco o niente è stato fatto per analizzare la diffusione dell'Open Source nel web.
- Questa analisi svolta sui siti dei comuni italiani ci permette di avere informazioni su un campione abbastanza grande di siti web e ci consente di capire qual è, allo stato attuale, l'orientamento tecnologico delle pubbliche amministrazioni italiane.
- La nostra analisi è stata ripetuta due volte: la prima volta nel **Febbraio 2011** e la seconda nel **Febbraio 2012**, con lo scopo di comprendere il trend di diffusione dei CMS Open Source nel web.



# Introduzione

Lo studio dimostra che la diffusione dei CMS Open Source **sta aumentando**; un altro dato importante è dato dal fatto che le Pubbliche Amministrazioni stanno iniziando a percepire le opportunità (in termini di qualità, sicurezza, prestazioni e risparmio) offerte da tali tecnologie.



# Metodologia

Per poter analizzare la diffusione dei CMS Open Source nei siti web dei comuni italiani, abbiamo sviluppato un'applicazione in Java che lavora in due passi

1. Nella prima fase vengono collezionati e memorizzati in un database tutti gli url dei siti web dei comuni. Gli indirizzi sono stati raccolti interrogando il sito web dell'ANCI (il portale dell'Associazione Nazionale Comuni Italiani, contenente informazioni puntuali e abbastanza aggiornate sui comuni).

2. Nella seconda fase lo strumento interroga il database e inizia il processo di riconoscimento della tecnologia utilizzata per lo sviluppo del sito.



# Metodologia

I CMS Open Source sono tanti e differenti; per il nostro studio abbiamo deciso di considerare i più comuni e più conosciuti:

- 1) Joomla!
- 2) Drupal;
- 3) Wordpress;
- 4) Open CMS;
- 5) Mambo;
- 6) XOOPS;
- 7) Plone



## Metodologia

- Per riconoscere in maniera automatica il tipo di CMS utilizzato, abbiamo provveduto a creare delle installazioni base dei CMS considerati nello studio (senza inserire contenuti personalizzati), in modo tale da avere punti di riferimento per poter confrontare i siti web.
- Il software, nella seconda fase, analizza le home page dei siti web dei comuni italiani (una alla volta) depurandole (tramite l'utilizzo di espressioni regolari) dai contenuti testuali, e mantenendo soltanto codici HTML, PHP o Javascript.
- Il codice ottenuto viene confrontato con ciascuna installazione base dei diversi CMS, e viene calcolata la distanza di Levhestain e di Hamming tra il codice del sito web comunale e il codice dell'installazione base dei CMS



# Metodologia

La distanza di Levhestain tra due stringhe è il numero minimo di cambiamenti elementari per convertire una stringa  $x$  in una stringa  $y$ . I cambiamenti elementari possono essere di tre tipi:

- 1. cancellazione di un carattere;***
- 2. rimpiazzamento di un carattere con un altro;***
- 3. inserimento di un carattere.***





# Metodologia

La distanza di Hamming tra due stringhe di lunghezza identica è il numero di posizioni nelle quali i simboli corrispondenti sono diversi.

La distanza di Hamming misura il numero di sostituzioni necessarie per convertire una stringa nell'altra. Il peso di Hamming di una stringa è la sua distanza di Hamming dalla stringa nulla (stringa costituita solo da zeri) della stessa lunghezza.



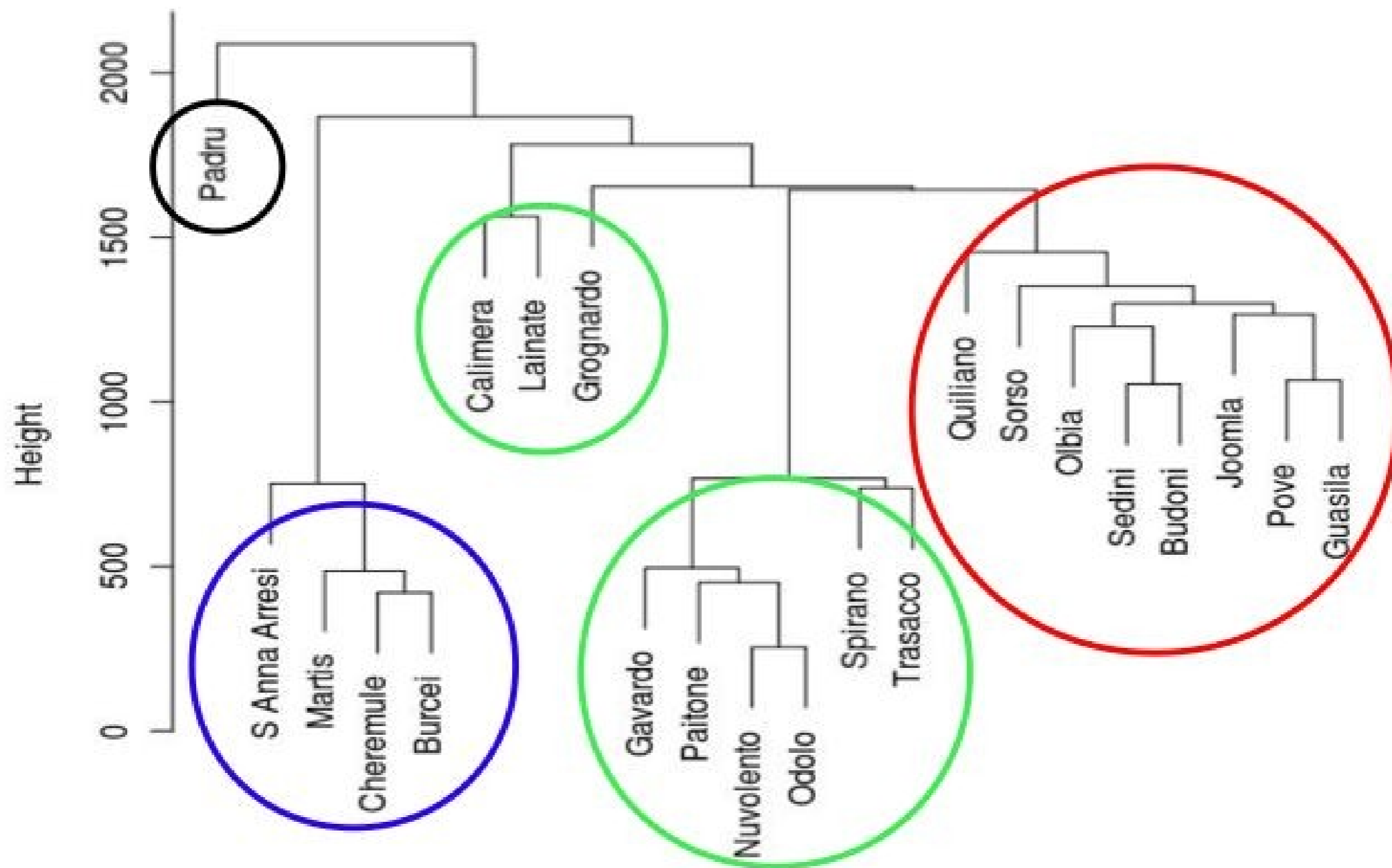
## Metodologia

- Per migliorare la qualità del riconoscimento del tipo di CMS utilizzato, viene fatta una ricerca di parole chiave (come Joomla!, Drupal, ecc.) nel codice sorgente della pagina web, e in caso di presenza di queste parole viene diminuito il valore di distanza con le installazioni base dei sistemi.
- Come terzo passaggio, viene effettuato un clustering gerarchico tra tutti i siti e le installazioni base dei CMS in modo tale da raggrupparli in insiemi simili per evidenziare le caratteristiche comuni.



# Metodologia

Cluster Dendrogram





# Risultati

Table 1. Risultati

	% 2011	% 2012	% crescita	% 2011	% 2012
Drupal	1,36	2,15	0,8	8,6	10,1
Joomla!	8,77	12,32	3,5	55,4	57,9
Mambo	0,47	0,25	-0,2	3,0	1,2
OCMS	2,58	3,10	0,5	16,3	14,6
Plone	1,03	1,23	0,2	6,5	5,8
Wordpress	1,36	1,91	0,5	8,6	9,0
Xoops	0,20	0,15	0,05	1,2	0,7
Minors	0,06	0,14	0,1	0,4	0,7
totale	15,8	21,2	5,4	100	100
Altri	84,17	78,8			



## Risultati

- I dati mostrano una crescita dei CMS Open Source nelle pubbliche amministrazioni comunali, con una penetrazione dei CMS open source pari al 21.2% nel 2012, in aumento del 5% se si confronta il valore del 2011 (15.8% è il valore registrato nel 2011).
- La colonna "% crescita" indica la crescita percentuale di ogni CMS Open Source sul totale dei siti web analizzati. Le ultime due colonne invece mostrano il peso relativo di ogni CMS analizzato nella categoria dei CMS Open Source. La categoria "Minors" contiene la somma dei siti che utilizzano PhpBB ed altri (con una bassa percentuale di utilizzo).



## Risultati

- Joomla! risulta essere il CMS che contribuisce maggiormente a guidare la crescita dei CMS Open Source, seguito subito da Drupal, Wordpress e Open CMS. Il dato negativo mostrato da Mambo è dovuto al fatto che si tratta di una distribuzione obsoleta (sostituita da Joomla!) e per la quale è consigliata la migrazione verso Joomla!
- Nell'insieme dei CMS Open Source analizzati c'è una forte crescita di Joomla! che era già predominante nel 2011, e una netta crescita di Drupal (tasso di crescita del 15%).
- Al contrario, Open CMS, il quale ancora cresce e ha contribuito alla crescita di utilizzo dei CMS Opens Source nel 2012, ha diminuito la sua rappresentatività all'interno dell'insieme dei CMS Open Source analizzati.



## Risultati

Il motivo della predominanza di Joomla! va attribuito al fatto che si tratta di un sistema molto conosciuto e completo, per il quale ci sono molti plug-in utilizzabili per qualsiasi tipo di funzionalità richiesta per il sito web.

Si tratta di un CMS estremamente semplice da usare da parte di utenti poco esperti, è completamente compatibile con tutti i tipi di browser web.

Essendo uno dei CMS più utilizzati è soggetto ad esternalità di rete: ci sono molti tutorial, manuali, forum, e molti casi d'uso ben spiegati.

Le motivazioni che spiegano la crescita di Drupal sono simili a quelle di Joomla!, e nell'ultimo periodo Drupal ha goduto di diversa pubblicità da parte di importanti enti governativi: basti pensare alla pubblicità fatta dalla Casa Bianca Americana.

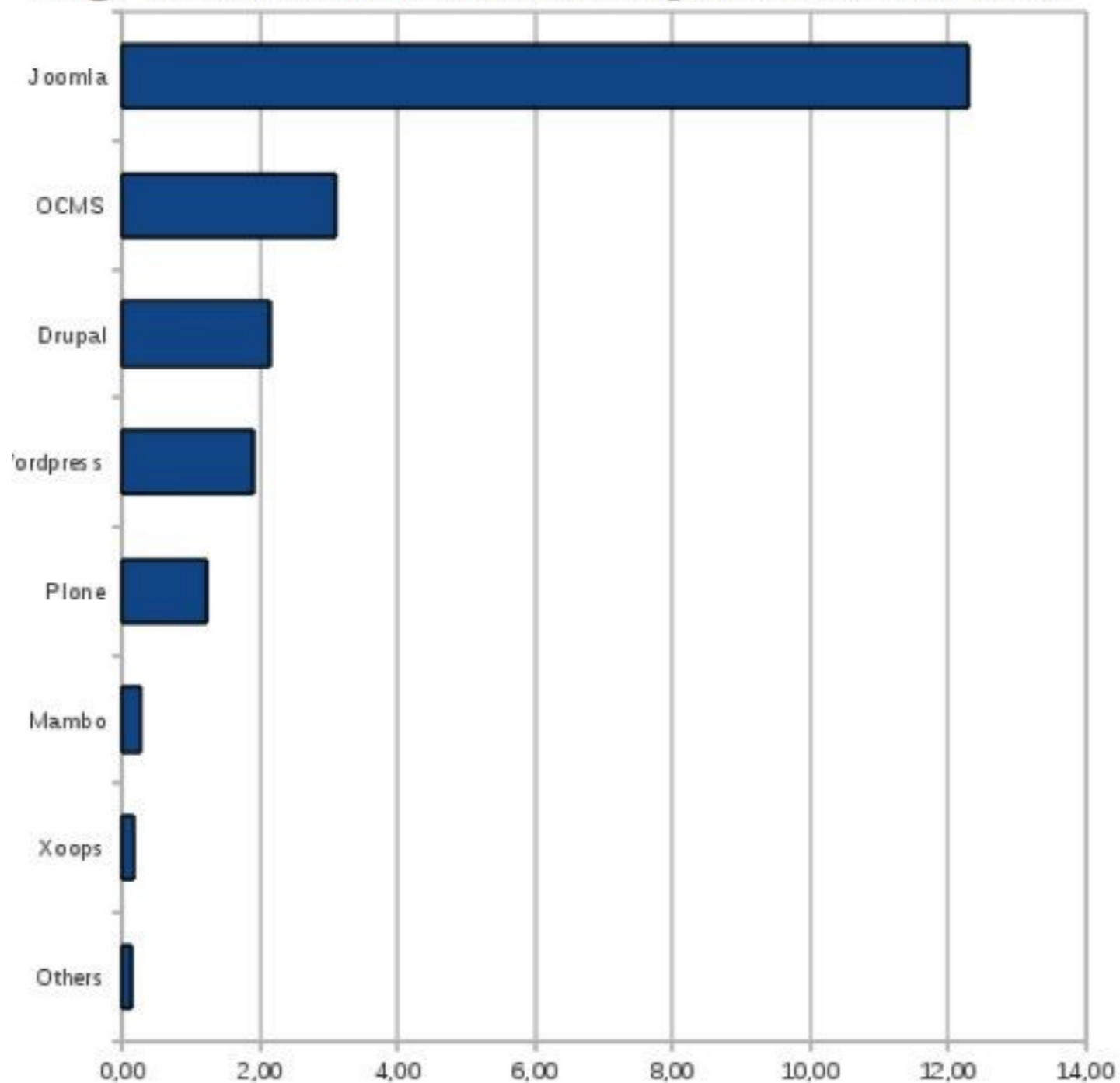


- La rapida crescita di Wordpress va probabilmente attribuita alla semplicità delle procedure di installazione del sistema. Si tratta di un'operazione che non richiede competenze tecniche (a differenza di Joomla! e Drupal, ad esempio) e sta crescendo la diffusione di template grafici e di plug-in studiati appositamente per Wordpress.
- La decrescita di Open CMS va probabilmente attribuita al tipo di linguaggio di sviluppo utilizzato: è scritto in Java e ciò comporta una richiesta di forti conoscenze tecniche da parte degli utenti finali.



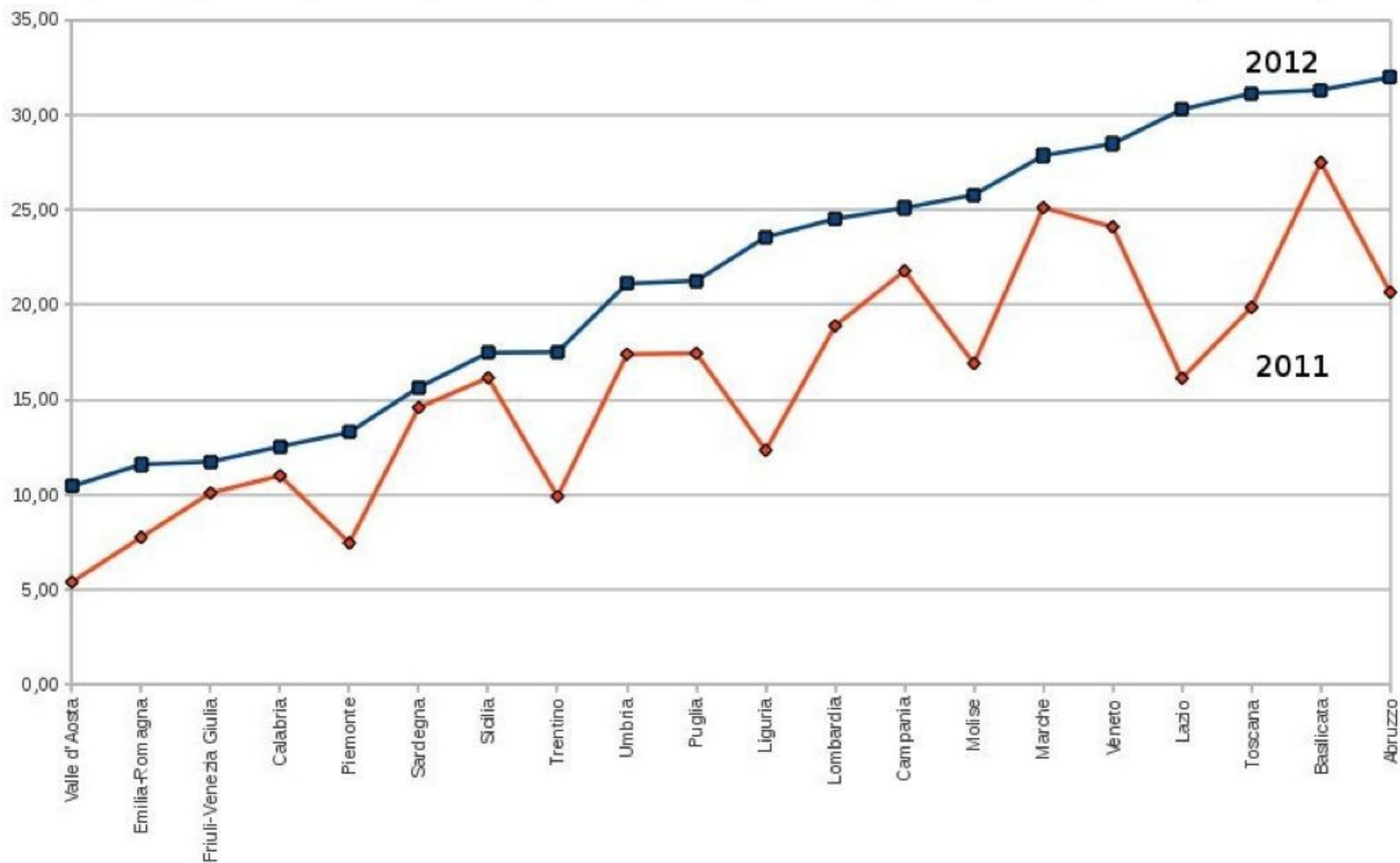


**Fig. 2.** Percentuale dei CMS Open Source nel 2012





**Fig. 3.** Percentuale di diffusione dei CMS Open Source per regione





## Conclusioni

Gli obiettivi di questa ricerca sono stati quelli di analizzare la diffusione dei CMS Open Source nei comuni italiani. Abbiamo sviluppato un software in Java che lavora in due fasi:

1. Durante la prima fase vengono memorizzati in un database gli URL dei siti web di tutte le amministrazioni comunali italiane (8092 comuni), recuperati dal sito web dell'ANCI.

2. Durante la seconda fase viene interrogato il database creato nella prima fase e si esegue l'elaborazione del contenuto del sito web. Infine, si applica un clustering gerarchico tra tutti i siti e le installazioni di base del CMS, al fine di raggrupparli in insiemi simili e in modo da evidenziarne le caratteristiche comuni.



## Conclusioni

Abbiamo condotto due analisi, una nel mese di Febbraio 2011 e una nel mese di Febbraio 2012 e i risultati mostrano che il numero di comuni italiani che utilizzano CMS Open Source sono in aumento: il 21,2% delle amministrazioni comunali utilizza CMS Open Source, come osservato nel mese di febbraio 2012.

I dati mostrano che dopo un anno la diffusione dei CMS Open Source è aumentata del 5%, e si prevede una crescita anche per l'immediato futuro



# Siti rilevanti in Joomla

## *Enti italiani*

- [www.regione.calabria.it](http://www.regione.calabria.it) (Sito istituzionale della Regione Calabria)
- [www.regione.liguria.it](http://www.regione.liguria.it) (Sito istituzionale della Regione Liguria)
- [www.provincia.salerno.it](http://www.provincia.salerno.it) (Sito istituzionale della Provincia di Salerno)
- [www.provincia.brindisi.it](http://www.provincia.brindisi.it) (Sito istituzionale della Provincia di Brindisi)
- [www.provincia.asti.it](http://www.provincia.asti.it) (Sito istituzionale della Provincia di Asti)
- [www.pariopportunita.gov.it](http://www.pariopportunita.gov.it) (Dipartimento per le Pari Opportunità)
- [www.sviluppoeconomico.gov.it](http://www.sviluppoeconomico.gov.it) (Sito del Ministero dello Sviluppo Economico)
- [www.turistia4zampe.it](http://www.turistia4zampe.it) (Campagna Ministero del Turismo)
- [www.anci.puglia.it](http://www.anci.puglia.it) (L'associazione dei comuni pugliesi)
- [www.consumatori.it](http://www.consumatori.it) (Unione Nazionale Consumatori)
- [www.enit.it](http://www.enit.it) (Agenzia Nazionale del Turismo)
- [www.coni.it](http://www.coni.it) (Comitato Olimpico Nazionale Italiano)



## Calcio, Serie A (20% di siti fatti in Joomla)

- - [www.sampdoria.it](http://www.sampdoria.it) (Sampdoria Calcio)
- - [www.pescaracalcio.com](http://www.pescaracalcio.com) (Pescara Calcio)
- - [www.sslazio.it](http://www.sslazio.it) (Lazio calcio)
- - [www.torinofc.it](http://www.torinofc.it) (Torino Fc)



## Altre organizzazioni e aziende

- [www.salone libro.it](http://www.salone libro.it) (Salone Internazionale del libro)
- [www.giffonifilmfestival.it](http://www.giffonifilmfestival.it) (Giffoni Film Festival)
- [www.cittadinanzattiva.it](http://www.cittadinanzattiva.it) (Cittadinanzattiva onlus)
- [www.federturismo.it](http://www.federturismo.it) (Federturismo Confindustria)
- [www.elisatoffoli.com](http://www.elisatoffoli.com) (Cantante)
- [www.totalerg.it](http://www.totalerg.it) (TotalErg)
- [www.leroymerlin.it](http://www.leroymerlin.it) (Leroy Merlin)
- [www.foppapedretti.it](http://www.foppapedretti.it) (Foppapedretti S.p.a)



A questo link:

- <http://www.joomla.it/notizie/4349-i-comuni-utilizzano-joomla.html>

è possibile trovare l'elenco dei domini dei comuni italiani che utilizzano Joomla.





# Grazie per l'attenzione!

Per info: [concas@diee.unica.it](mailto:concas@diee.unica.it)