

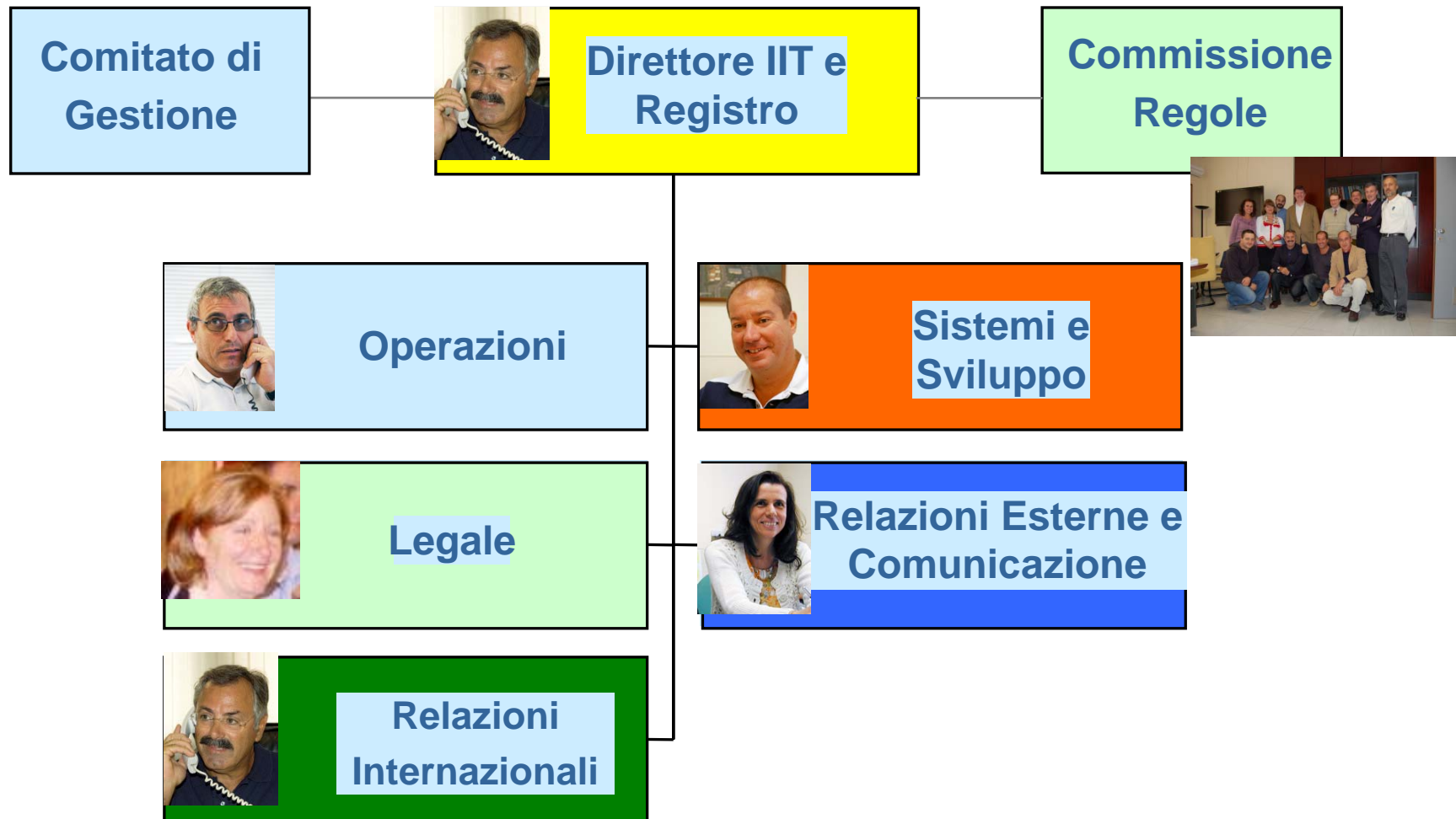
Il monitoraggio dei siti Web della Pubblica Amministrazione: l'esperienza del Registro .it

Maurizio Martinelli

*Responsabile Servizi Internet e Sviluppo Tecnologico
Istituto di Informatica e Telematica CNR – Pisa*

- È stato delegato al CNR il 23 dicembre 1987
- Circa 2.200.000 nomi a dominio registrati
- Nuovo sistema di registrazione sincrono dal 28 settembre 2009
- Il “vecchio” sistema di registrazione asincrono è stato definitivamente dismesso il 1 maggio 2011
- Circa 1.300 Registrar sincroni
- Aperto a tutti coloro che risiedono in un paese EU

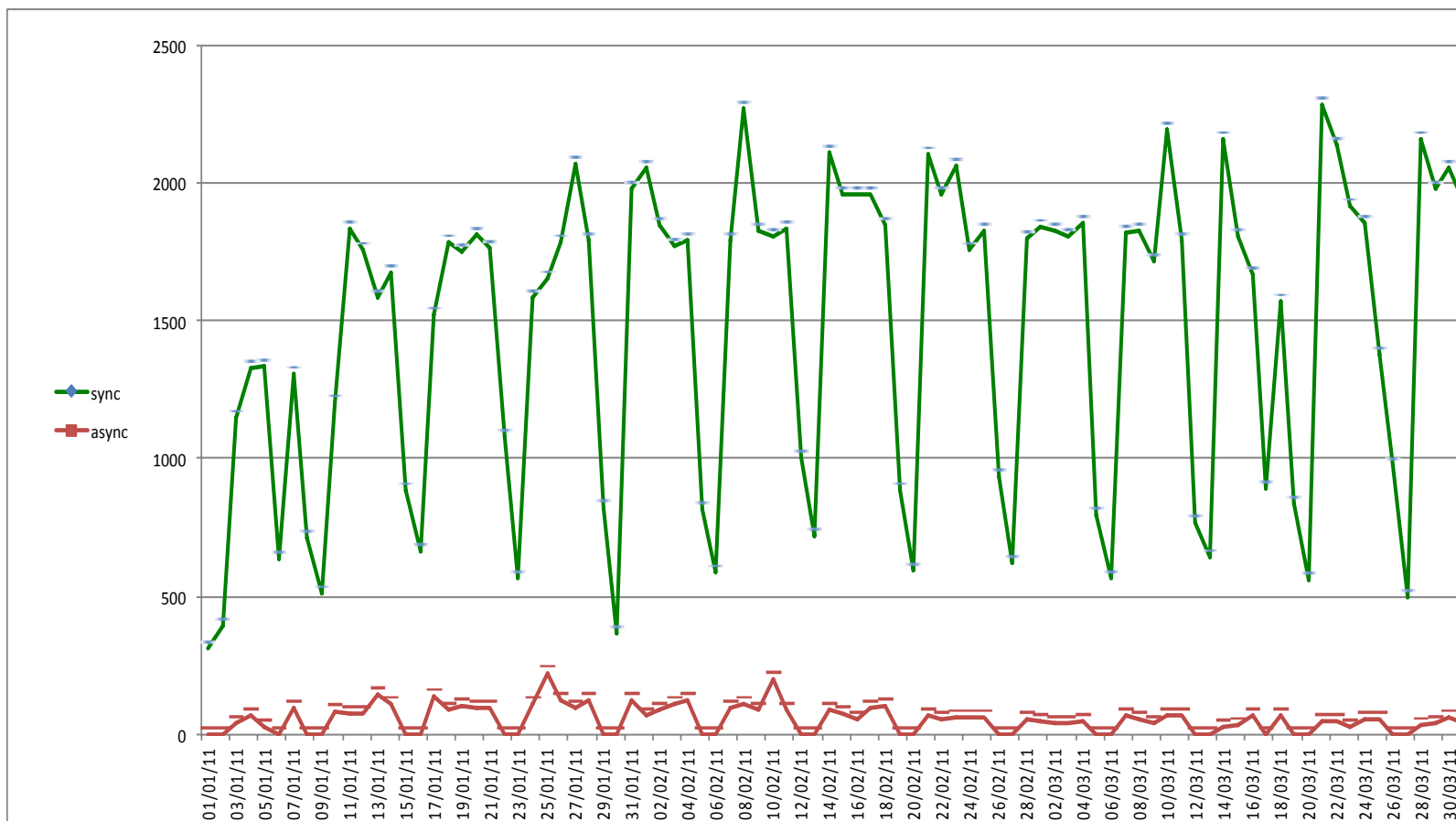
Il Registro .it: struttura organizzativa



Composto da circa 70 persone

- Organismo consultivo del Registro costituito ad aprile 2004
- Definisce il Regolamento di assegnazione dei nomi a dominio nel .it
- Composizione :
 - 6 membri proposti dalla Local Internet Community (LIC)
 - 4 rappresentanti dei Registrar
 - 1 rappresentante di ISOC
 - 1 rappresentante degli utenti
 - 1 membro del Consorzio GARR
 - 2 membri del Registro .it
 - 1 rappresentante del Ministero per lo Sviluppo Economico
- È attualmente in fase di riorganizzazione

Sincrono e Asincrono a confronto: tendenza 2011

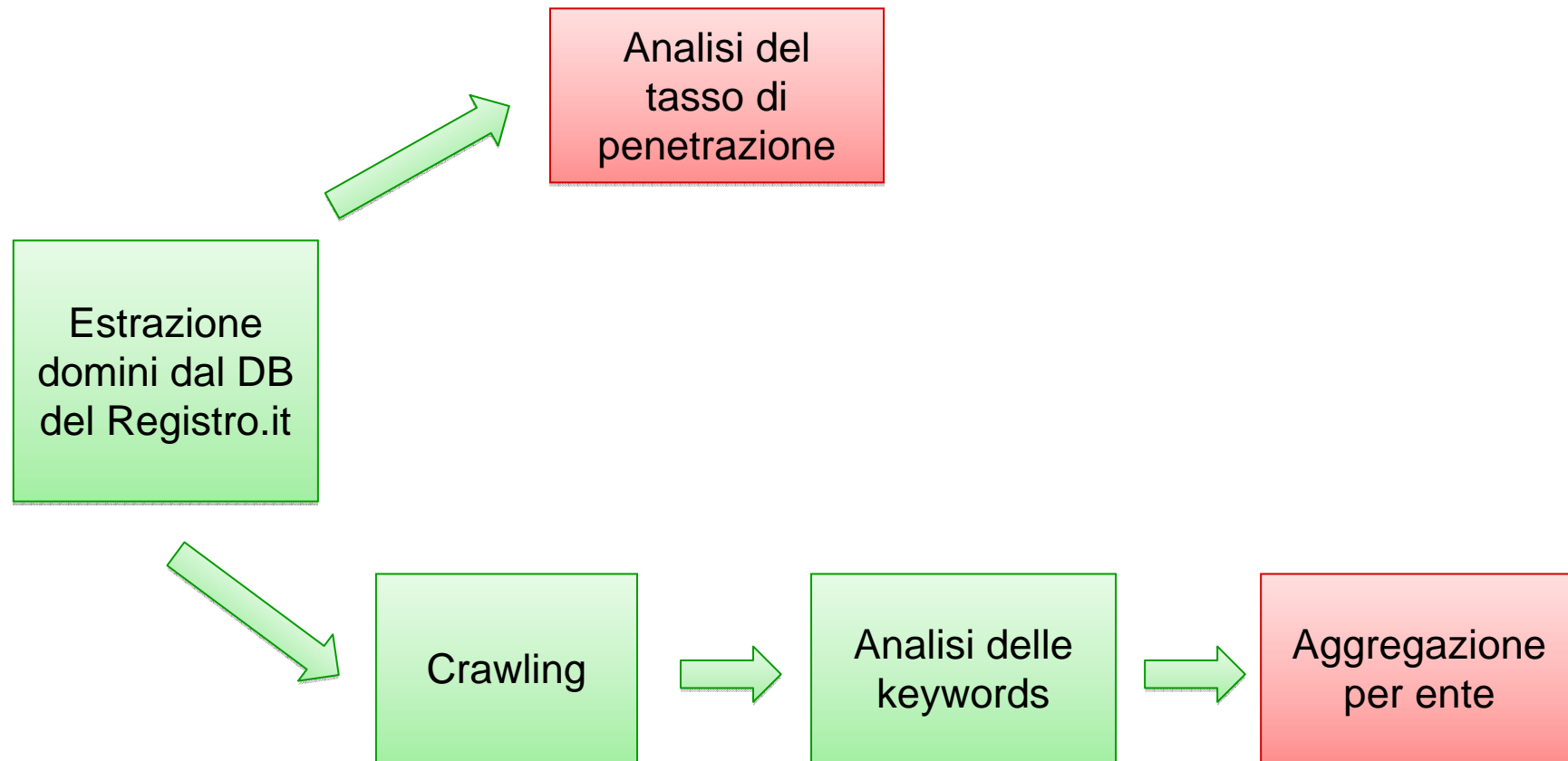


Nuove registrazioni sincrone dal 1/1/2011 al 31/3/2011: 132.217

Nuove registrazioni asincrone dal 1/1/2011 al 31/3/2011: 4.947

Totale nuove registrazioni dal 1/1/2011 al 31/3/2011: 137.164

- Analisi della diffusione di Internet nella Pubblica Amministrazione (PA)
 - metrica endogena dei nomi a dominio “.it”
- Analisi dell’aderenza dei siti web della PA alle linee guida emesse dal Ministero per la Pubblica Amministrazione e l’Innovazione
 - Crawling e analisi per keyword



VANTAGGI

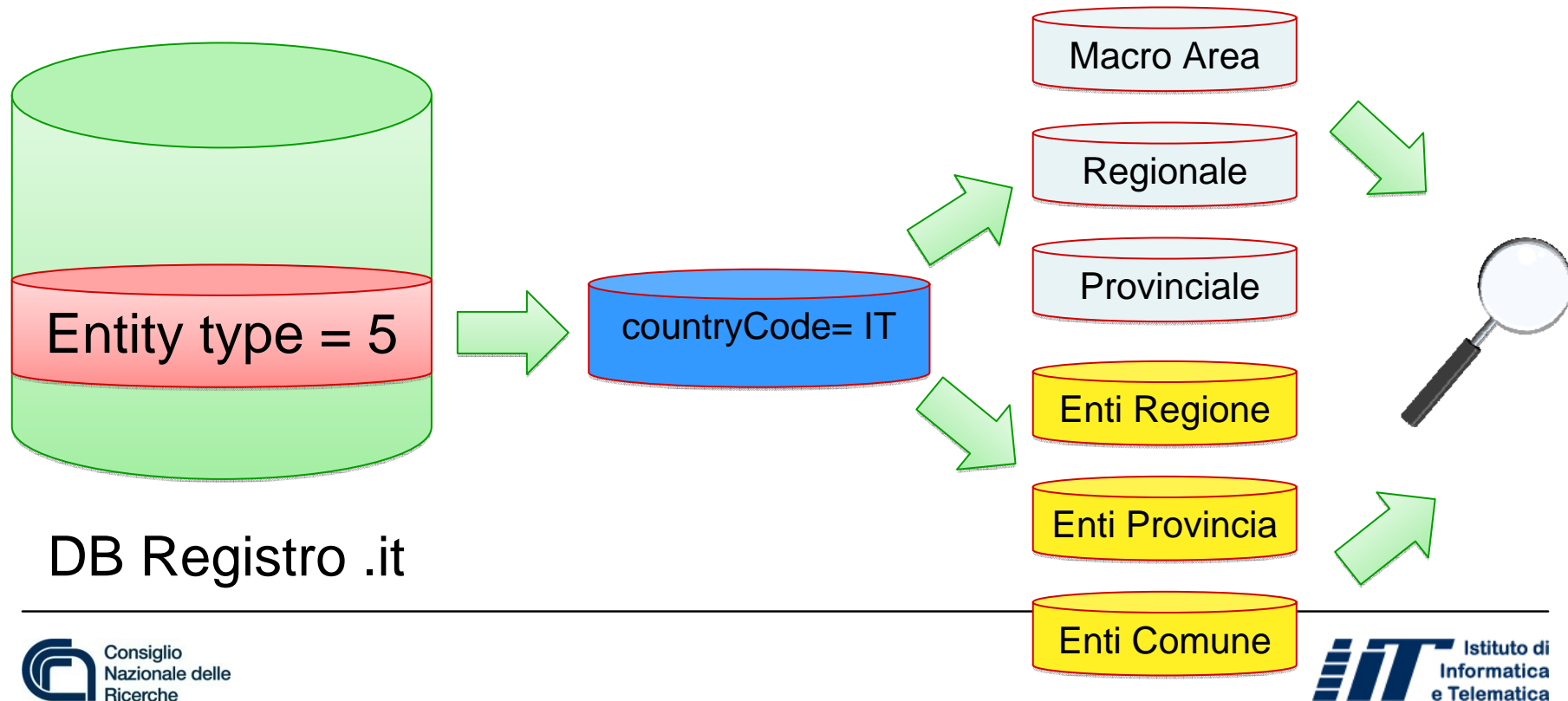
- “A domain represents a conscious decision of using the Internet in a more sophisticated manner” (Zook, 2000)
- Associazione unica con una persona o un’organizzazione
 - Caratteristiche del registrante
 - Localizzazione del registrante

SVANTAGGI

- Domain grabbing
- Sottostima della diffusione
 - Non tutti gli utenti Internet hanno un nome a dominio
 - I domini rappresentano un Lower Bound
- Sovrastima della diffusione
 - Più nomi a dominio possono essere associati allo stesso registrante

- **Caratterizzazione del Registrante**
 1. Persone fisiche
 2. Società
 3. Liberi professionisti
 4. Enti no-profit
 5. **Enti pubblici**
 6. Altri enti
 7. Soggetti stranieri
- **Localizzazione del Registrante**
 - Nazionale
 - Macro Area (Nord, Centro, Sud)
 - Regionale
 - Provinciale

- I nomi a dominio registrati dalla Pubblica Amministrazione sono stati ottenuti attraverso procedure automatiche e semi-automatiche



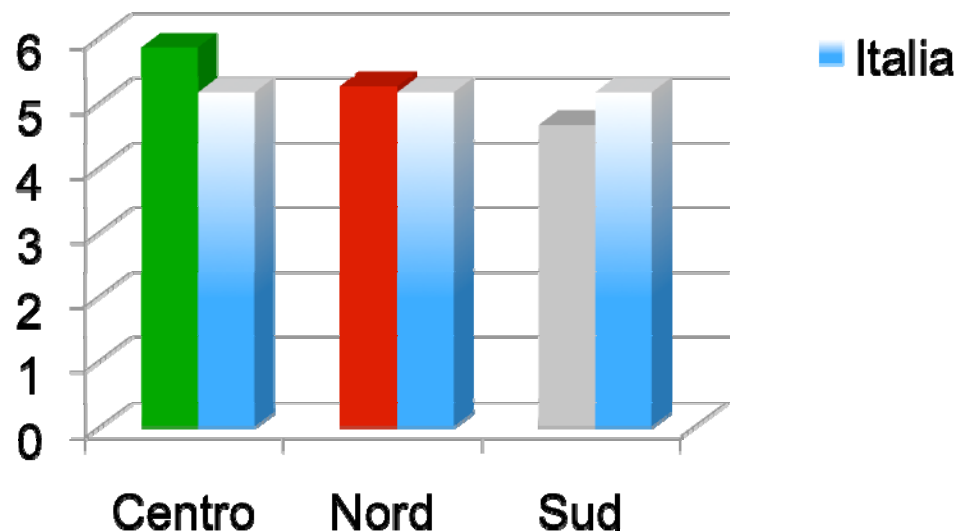
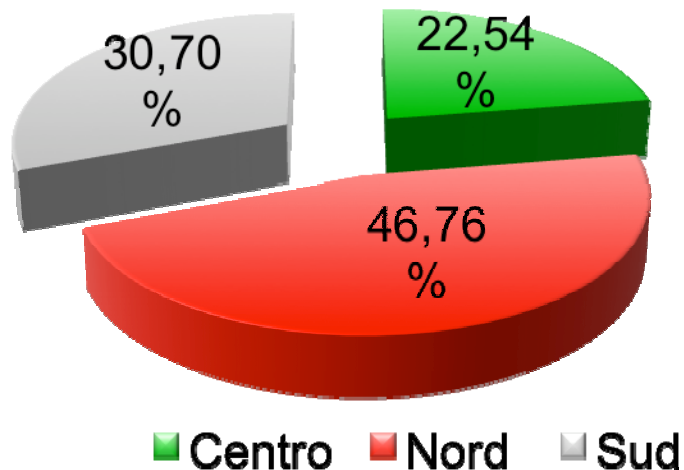
- 25.681 nomi a dominio registrati dalla PA
 - 15.279 assegnati a enti territoriali
 - 947 assegnati a 20 enti regione
 - 1.022 assegnati a 109 enti provincia
 - 13.310 assegnati a 7.923 (su 8.094) enti comune
- I nomi a dominio “unici” (esclusi gli alias) registrati dalla PA risultano 23.266
 - **Alias:** dominio al quale è associato lo stesso sito

Casi	Numero Casi in Italia	Domini Totali	Alias	Domini Senza sito	Siti validi (non alias)	Non classificabile (*)
Ente Regione	20	947	336	219	347	45
Ente Provincia	110	1.022	263	104	632	23
Ente Comune	8.094	13.310	1.350	727	11.120	113
Totale		15.279	1.949	1.050	12.099	181

() domini ai quali corrisponde un sito web con redirect errato, pagine in costruzione, ecc.*

- Per misurare la presenza di un eventuale digital divide e analizzare la diffusione di Internet a livello geografico è stato utilizzato il seguente indicatore:
 - **TP = Tasso di Penetrazione**
 - (Numero di domini registrati dagli enti pubblici/popolazione) x 10.000

Area	Domini	% domini	TP
Centro	5.789	22,54%	5,842
Nord	12.009	46,76%	5,224
Sud	7.883	30,70%	4,651
Italia	25.681	100,00%	5,152

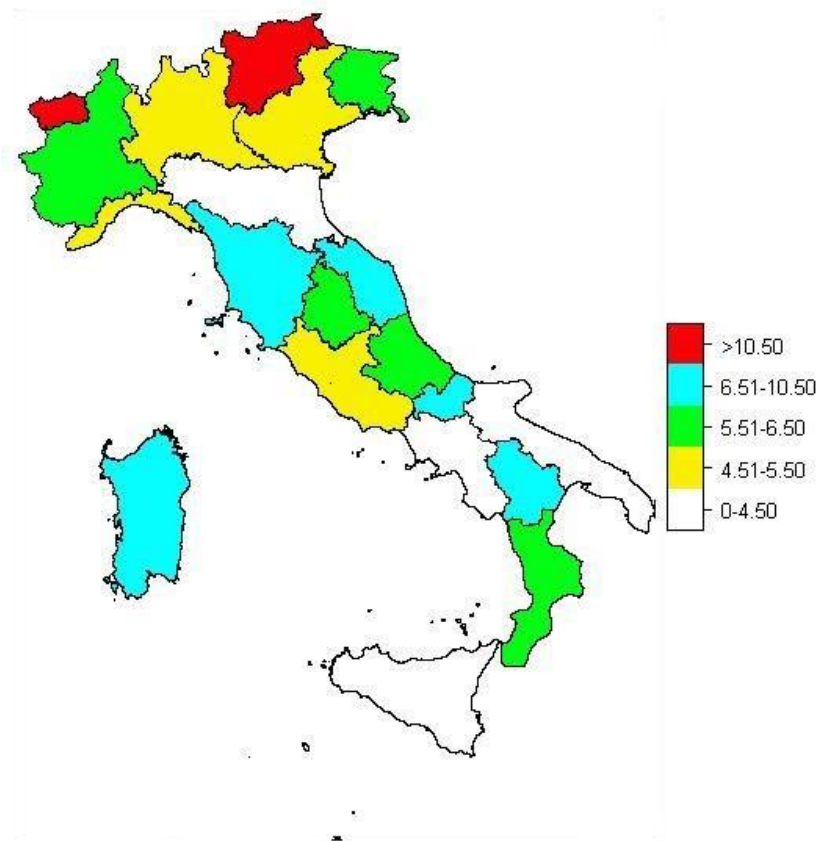


Percentuale di nomi a dominio a livello di macro-area

TP ogni 10.000 abitanti

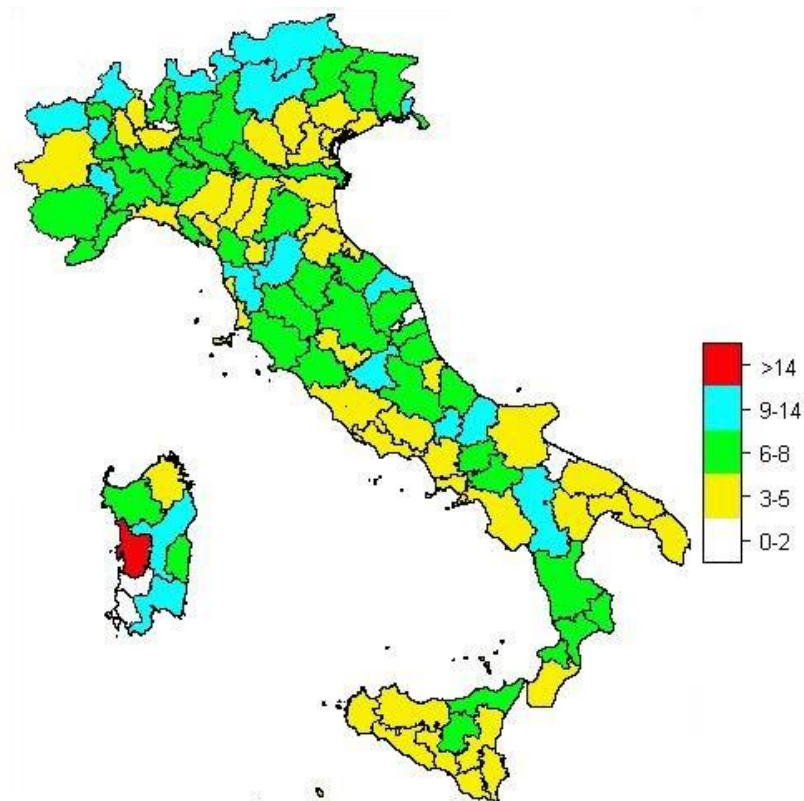
Diffusione di Internet nella PA: le prime 10 Regioni

Regioni	Numero domini	% domini	TP
Valle d'Aosta	139	0,54%	13,056
Trentino AA	921	3,59%	11,184
Molise	242	0,94%	8,967
Basilicata	406	1,58%	8,290
Sardegna	1.070	4,17%	7,561
Marche	964	3,75%	7,306
Toscana	2.236	8,71%	7,090
Friuli V.G.	677	2,64%	6,454
Piemonte	2.230	8,68%	5,932
Abruzzo	641	2,50%	5,721



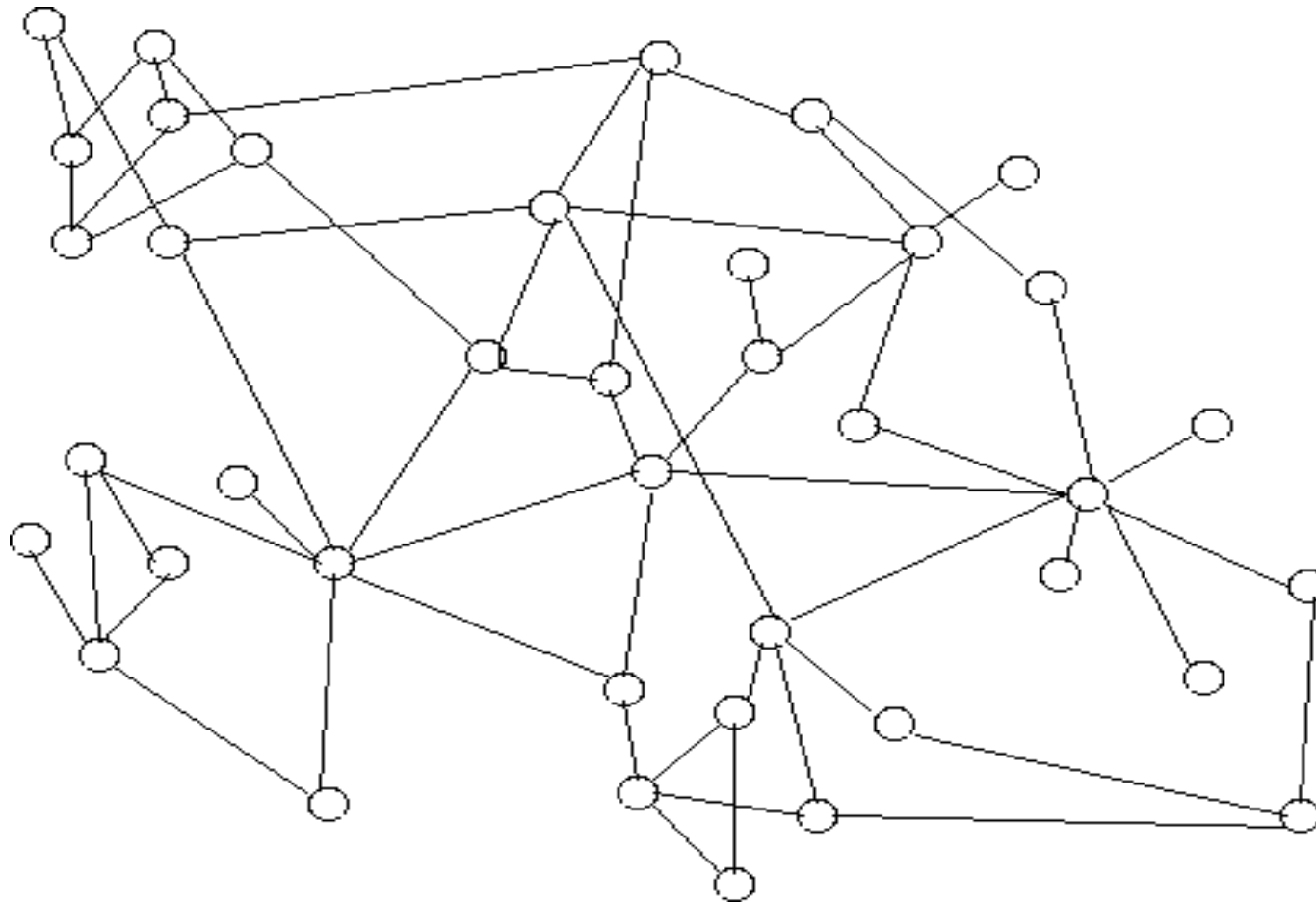
TP ogni 10.000 abitanti

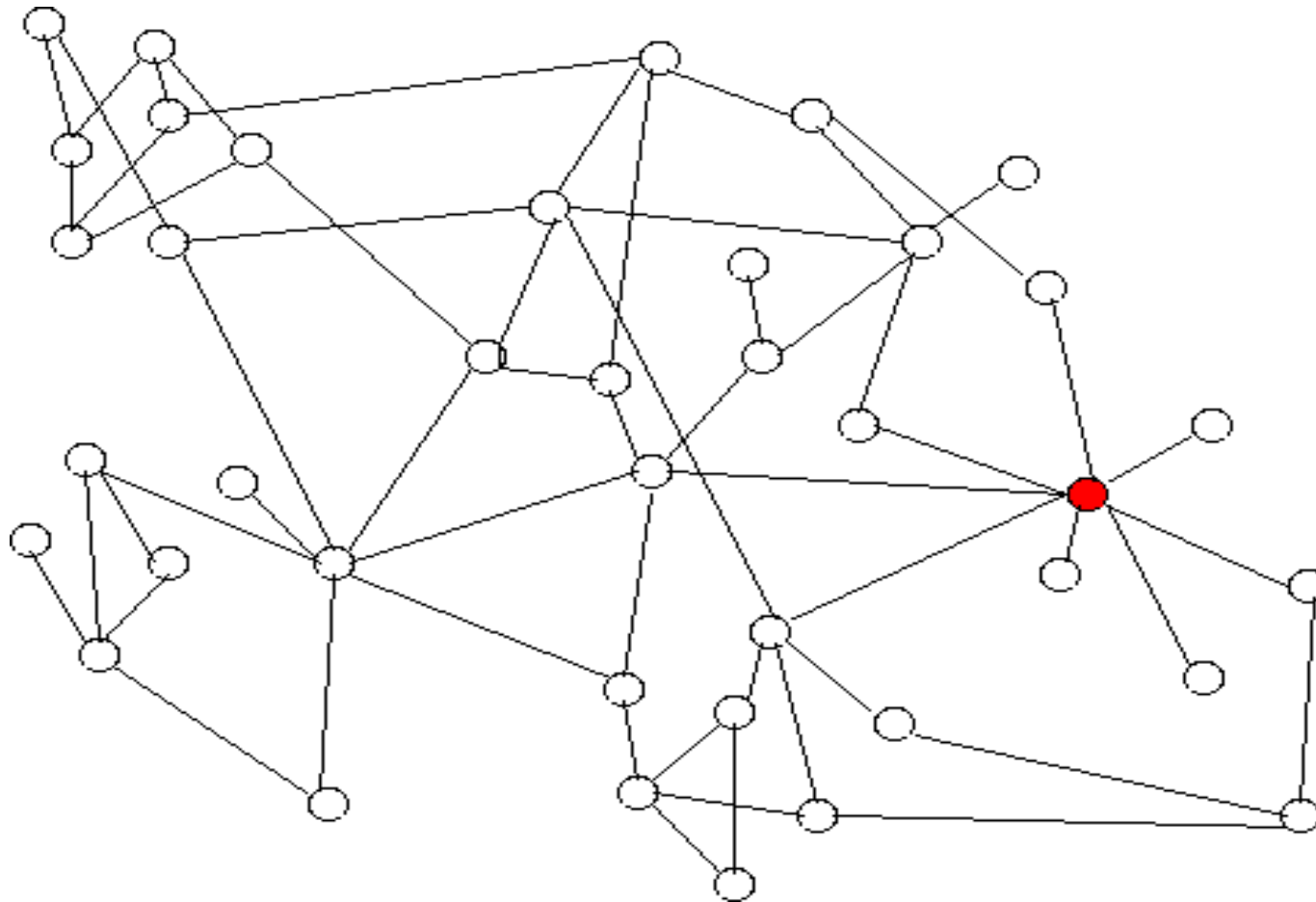
Province	Domini	% Domini	TP
Oristano	206	0,80%	14,406
Aosta	139	0,54%	13,056
Prato	267	1,04%	12,993
Nuoro	169	0,66%	12,524
Bolzano	464	1,81%	11,634
Verbania	150	0,58%	10,782
Trento	457	1,78%	10,763
Asti	193	0,75%	10,312
Potenza	323	1,26%	10,036
Cagliari	460	1,79%	9,693

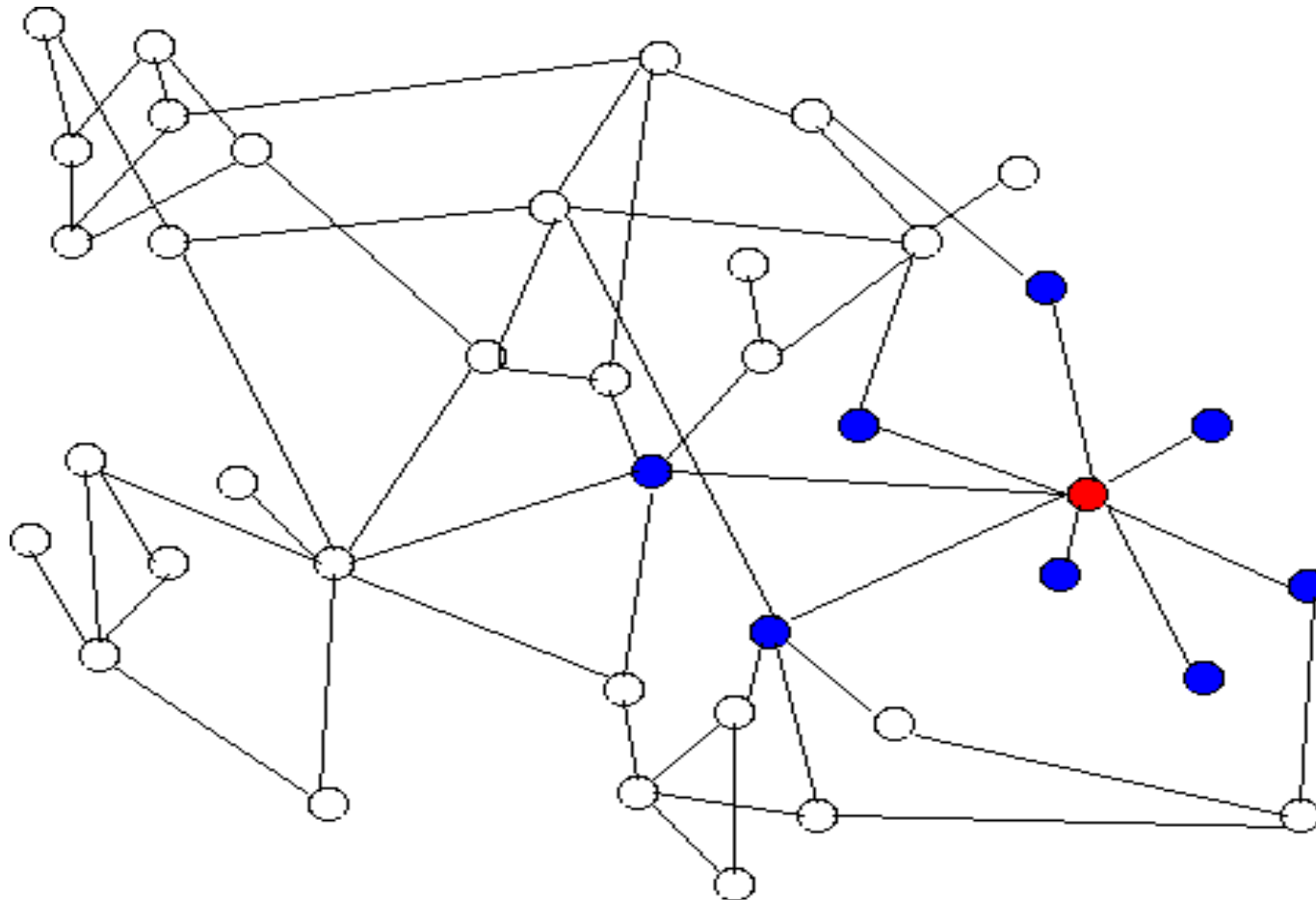


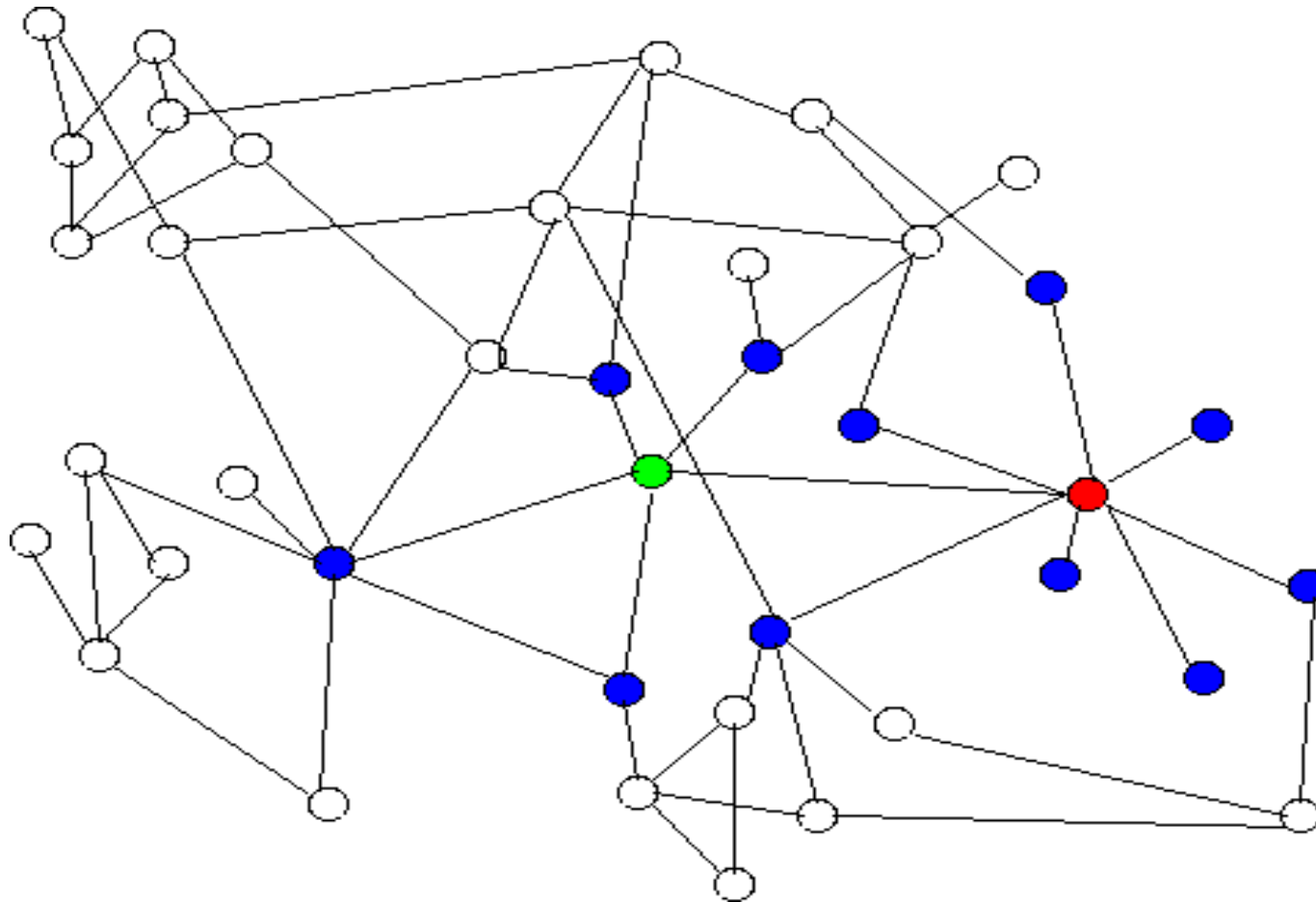
TP ogni 10.000 abitanti

- Partendo dalla lista dei nomi a dominio della PA, sono stati scaricati i siti web ad essa associati
 - Fino a 10 MB per sito non considerando immagini, fogli di stile, javascript esterni e documenti (doc, pdf, ecc.)
- Utilizzo del crawler L.webis sviluppato presso lo IIT
- Il testo in formato html è stato analizzato per individuare la presenza di “parole chiave” relative a sezioni, informazioni o servizi descritti nelle linee guida
- I dati estratti sono stati aggregati per ente







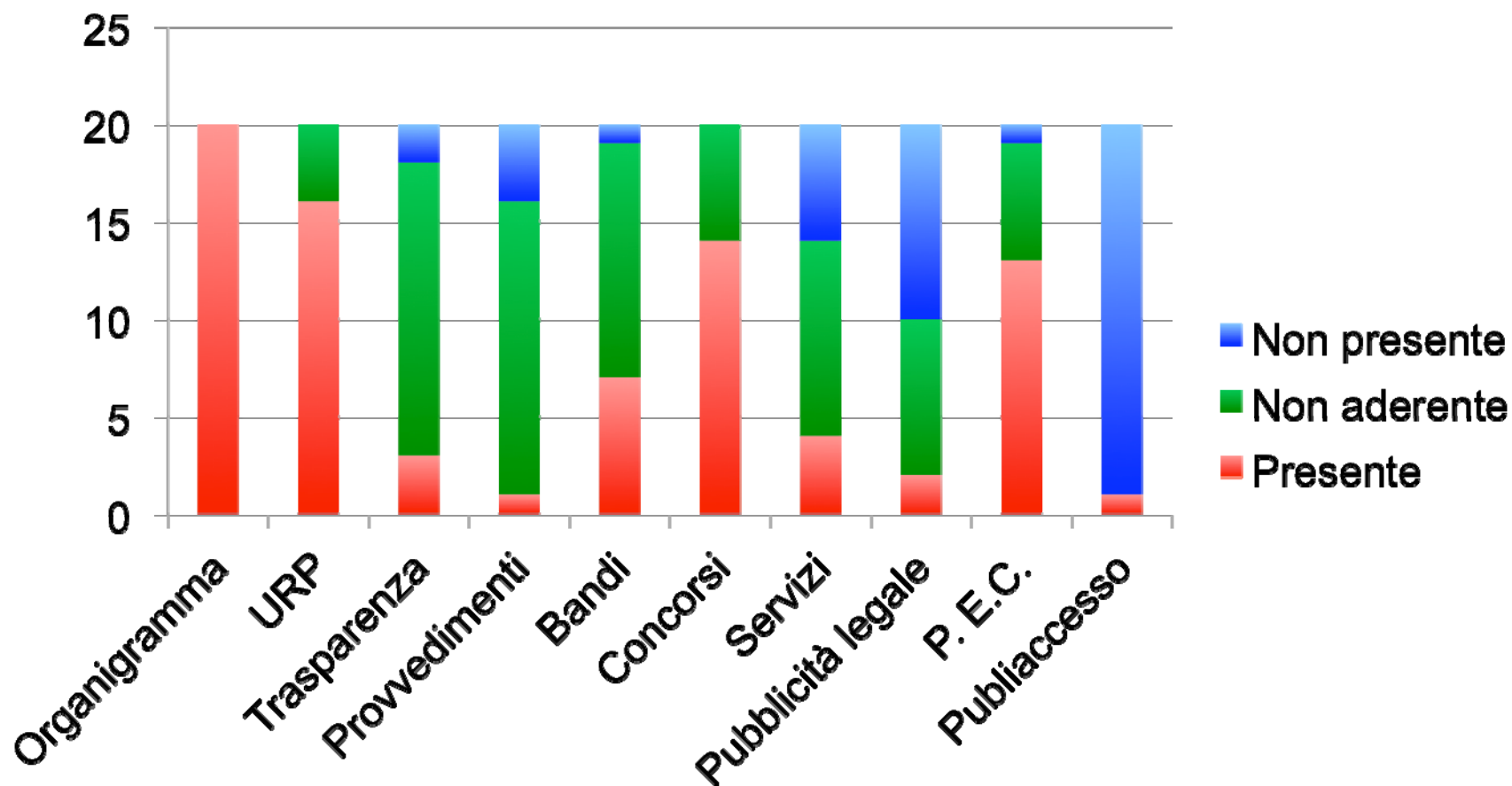


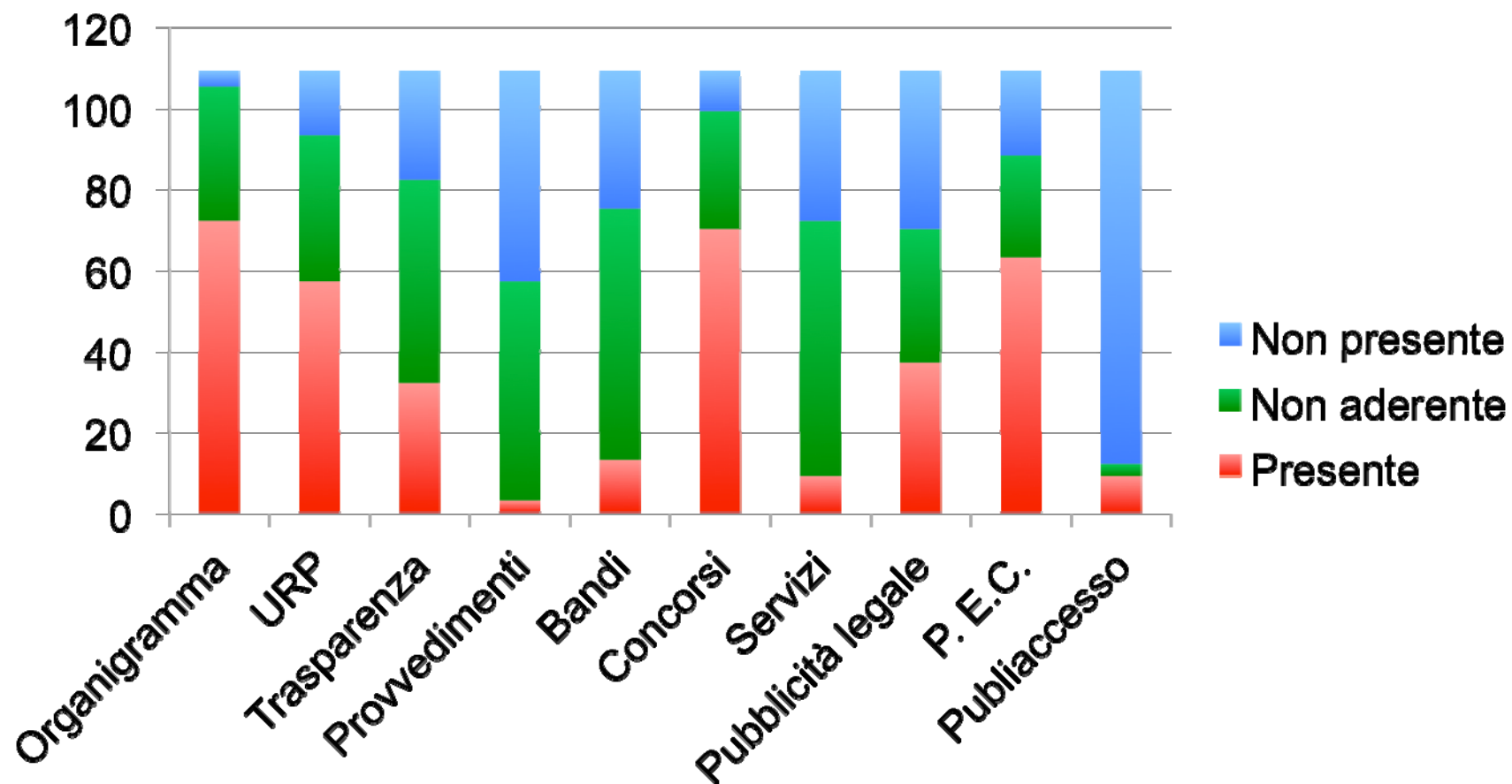
- La lista iniziale di domini delle PA ha **25.681** elementi
 - Tramite il crawling la lista viene estesa a **35.236** elementi (includendo i sottodomini)
- La fase di crawling ha impiegato **72** ore usando un cluster di 3 server
- Sono stati effettuati **10.951.256** download, di cui **8.951.731** pagine Web valide
- All'interno di tutte le pagine scaricate sono stati individuati **319 milioni di** collegamenti ipertestuali (URL)
- La mole di dati trasferiti è complessivamente di oltre **180 GB**

- Organigramma
- URP
- Trasparenza, valutazione e merito
- Procedimenti Amministrativi
- Bandi
- Concorsi
- Servizi on line
- Pubblicità legale
- PEC
- Publiaccesso (*)



- Presente: il contenuto è stato trovato ed è aderente alle linee guida ministeriali
- Non aderente: il contenuto è stato trovato ma in un formato non aderente alle linee guida
 - **Esempio**: non è un link, è un'immagine, non si trova in home page
- Non presente: il contenuto non è stato trovato in nessuna forma
 - In alcuni casi ciò è dovuto anche all'inaccessibilità del server web, al "mascheramento" dei dati, a errori ortografici, ecc.





Grazie per l'attenzione!