

Andrea Trentini¹, Pierlauro Sciarelli²

LA “SMART”-CITTADINANZA ATTRAVERSO LA COMPLIANCE DEI SITI WEB DELLA PUBBLICA AMMINISTRAZIONE: IL CASO DI STUDIO DEI COMUNI ITALIANI

Abstract

In Italy there are in effect many laws and guidelines that specify a set of compliance rules for municipal websites. This paper presents a survey regarding the accessibility, naming and registration compliance of municipal Web sites to these norms.

The study was implemented by: 1) programmatically crawling websites to check the compliance to the WCAG 2.0 guidelines adopted by the Stanca Act; 2) programmatically analyzing naming registration and resolution. The goal is to measure the degree of awareness and adoption of these laws and guidelines by the local administrations.

Alas, the results are somewhat disappointing with only 10% of sites compliant to accessibility while more promising in the naming area.

1. Introduzione

Il termine “smart”-cittadinanza ha più di un significato (De Cindio e Trentini 2014). È da intendersi:

1) sia come disponibilità di servizi tecnologici “informativi” quali opendata, accountability (Trentini 2014) e “partecipativi” quali bilanci partecipati, e-petitioning, ma anche i più “banali” servizi anagrafici, finanziari, ecc.

2) che come possibilità di creare nuovi servizi tecnologici a cura di terze parti (semplici cittadini, communities, terzo settore, etc.) basati sui servizi di base forniti dalla P.A. (Pubblica Amministrazione)

Il #1, soprattutto il contesto “partecipativo” e di servizio, vorrebbe vedere la massimizzazione della partecipazione cittadina, l’estrema facilità d’uso e l’ottimizzazione della raggiungibilità e accessibilità di tutti i servizi digitali.

Il #2 è facilitato ovvero impedito in funzione della accessibilità, della adeguatezza e della omogeneizzazione dei servizi offerti dalla P.A. stessa.

¹ Dipartimento di Informatica, Università di Milano, docente di “Cittadinanza Digitale e Tecnocivismo”, andrea.trentini@unimi.it (par. 1.3; 1.4; 2.1; 2.2; 2.4; 2.5; 3.2; 3.3)

² Barcelona Supercomputing Center (BSC), Universitat Politècnica de Catalunya (UPC), pierlauro.sciarelli@bsc.es (par. 1.1; 1.2; 2.3; 2.6; 3.1;

In entrambi i casi il rispetto delle linee guida determinate dalle norme vigenti da parte dei siti web statali, in gran parte fuori regola (Barricelli et al. 2017), incentiverebbe l'impiego da parte di cittadini e organizzazioni non governative (o, più in generale, di tutti gli enti terzi) dei dati e dei servizi messi a disposizione dalla Pubblica Amministrazione.

Questo articolo presenta uno studio effettuato sulla compliance (rispetto alla normativa Stanca e alle linee guida attuali) dei siti web dei Comuni italiani, mettendo in evidenza le evidenti mancanze trovate. Mancanze sia prettamente tecnologiche (e.g., inadempienza verso standard di accessibilità) che più generali e procedurali (e.g., non rispondere alle email, amministratori locali che ignorano lo stato del proprio sito web, ecc.). La ratio dello studio è che misurando l'attenzione delle P.A. comunali verso l'inclusione tecnologica (accessibilità in ogni senso possibile) si possa misurare il grado di progresso digitale degli enti locali più vicini alla cittadinanza.

Lo studio è stato realizzato implementando un meccanismo semi automatizzato di analisi dei circa 8000 siti web dei Comuni italiani. Il sistema software effettua principalmente due tipi di operazioni:

1) scaricare almeno una pagina di ogni sito e valutarne l'aderenza alle normative mediante strumenti automatici (cfr. Sezione 3.1);

2) effettuare analisi sulla localizzazione (cfr. Sezione 3.2) del sito e sulla registrazione del nome del sito stesso (cfr. Sezione 3.3), quindi senza entrare nel merito del contenuto.

Le sottosezioni seguenti descrivono l'impianto "normativo", sia in termini di legge effettivamente in vigore sia in termini di linee guida e *best practice* per la realizzazione dei siti web della P.A.

1.1 Web Accessibility Initiative

Prevedendo l'ascesa del World Wide Web, verso la fine degli anni '90 il W3C (WWW Consortium) diede vita alla Web Accessibility Initiative (Web Accessibility Initiative and others 1999) con l'obiettivo di redigere una lista di linee guida atte a facilitare la fruizione dei contenuti web da parte degli utenti con disabilità, in seguito allargata con linee guida per la fruizione sulle piattaforme più disparate (compresi smartphone).

Le WCAG 2.0 (Web Content Accessibility Guidelines 2.0), risalenti al 2008 (Reid e Snow-Weaver 2008), sono ad oggi alla base di gran parte delle legislazioni nazionali concernenti l'accessibilità. Le linee guida possono riassumersi nei seguenti punti:

1) Alternative testuali: a qualsiasi informazione di natura non testuale, immagini o altri contenuti di tipo multimediale, bisogna associare alternative testuali in modo da permetterne universalmente la fruizione.

2) Adattabilità: i contenuti devono essere presentabili in modalità differenti senza perdita di informazione indipendentemente dal tipo di dispositivo hardware o software utilizzato per la consultazione.

3) Distinguibilità: i vari elementi che compongono le pagine web devono essere chiaramente distinguibili.

4) Accessibilità da tastiera: tutte le funzioni accessibili tramite i sistemi di puntamento, quali mouse e touchpad, devono essere rese disponibili anche da tastiera.

5) Disponibilità di tempo: l'utente deve avere la possibilità di beneficiare dei contenuti per un tempo ragionevole, ad esempio evitando il logout automatico a pochi istanti dal login o l'aggiornamento continuo della pagina.

6) Crisi epilettiche: bisogna minimizzare la probabilità che una pagina causi crisi epilettiche evitando combinazioni di colori potenzialmente rischiose.

7) Navigabilità: i contenuti dei siti devono essere resi facilmente raggiungibili sfruttando al meglio lo spazio disponibile e non fornendo all'utente un elevato contenuto informativo.

8) Leggibilità: il contenuto testuale deve essere comprensibile e leggibile oltre che, possibilmente, fruibile in più lingue.

9) Prevedibilità: il fine delle varie operazioni effettuabili dall'utente deve essere chiaramente prevedibile e i componenti con le stesse funzionalità devono essere identificati in modo uniforme.

10) Assistenza nell'inserimento di dati e informazioni: l'utente deve essere supportato nell'inserimento dei dati prevenendo eventuali errori e fornendo chiare spiegazioni sulle corrette modalità di input. Inoltre, bisogna dare la possibilità di revisionare e modificare i dati inseriti prima di eventuali sottomissioni dei form.

11) Compatibilità: deve essere garantita la massima compatibilità con la gran parte dei browser.

1.2 Legge Stanca

La legge Stanca (Galliani 2008), risalente al 2004 ma recentemente riformata sulla base delle linee guida WCAG 2.0, illustra gli standard di accessibilità che dovrebbero soddisfare tutti i siti web di enti pubblici, aziende partecipate e semi-partecipate dallo stato.

La normativa, che in sostanza ricalca le disposizioni suggerite dal W3C, impone delle regole basate su quattro importanti principi:

Percepibilità: le informazioni e i componenti dell'interfaccia utente devono essere fruibili attraverso i canali sensoriali.

Utilizzabilità: i componenti dell'interfaccia devono essere utilizzabili da chiunque senza vincoli o disagi.

Comprensibilità: gli utenti devono poter facilmente comprendere le modalità di funzionamento dell'interfaccia.

Robustezza: il contenuto deve essere interpretabile da una vasta quantità di programmi (browser).

La legge, ad oggi, risulta diffusamente non rispettata. Tale mancanza può essere identificata come conseguenza dell'assenza di organi preposti al controllo dell'effettiva applicazione della normativa nonché di eventuali sanzioni.

1.3 Geo-compliance

Un aspetto “non funzionale” (legato cioè alla modalità più che al tipo di servizi offerti) dell'implementazione dei siti web (anche se non limitatamente a questi) della P.A. è il cosiddetto vincolo della “geolocalizzazione dei dati”. Tale normativa prevede che tutti i server governativi (ivi compresi i siti web) devono essere collocati in Italia o quantomeno in territorio interno alla Comunità Europea. La normativa è varia e disseminata, ad esempio si veda (Consiglio dei Ministri 2014):

DECRETO DEL PRESIDENTE DEL CONSIGLIO DEI MINISTRI 3 dicembre 2013 lettera b), 35, comma 2, 36, comma 2, e 71, pubblicato nella Gazzetta Ufficiale 21 maggio 2013, n. 117;

Regole tecniche in materia di sistema di conservazione ai sensi degli articoli 20, commi 3 e 5-bis, 23-ter, comma 4, 43, commi 1 e 3, 44, 44 -bis e 71, comma 1, del Codice dell'amministrazione digitale di cui al decreto legislativo n. 82 del 2005.

...

2. Fatto salvo quanto previsto dal decreto legislativo 22 gennaio 2004, n. 42, in ordine alla tutela, da parte del Ministero dei beni e delle attività culturali e del turismo, sugli archivi e sui singoli documenti dello Stato, delle regioni, degli altri enti pubblici territoriali, nonché di ogni altro ente ed istituto pubblico, i sistemi di conservazione delle pubbliche amministrazioni e i sistemi di conservazione dei conservatori accreditati, ai fini della vigilanza da parte dell'Agenzia per l'Italia digitale su questi ultimi, prevedono la materiale conservazione dei dati e delle copie di sicurezza sul territorio nazionale e garantiscono un accesso ai dati

presso la sede del produttore e misure di sicurezza conformi a quelle stabilite dal presente decreto.

O anche il Decreto Legge n. 90/2014, convertito in Legge n. 114/20 (Repubblica Italiana 2014) e le regole dell'AGID (AGID 2017).

Il motivo alla base della normativa è che la legislazione applicabile ai dati immagazzinati in un server è quella del paese in cui si trova fisicamente il server stesso (Pizzetti 2016). In particolare è importante che le norme sul trattamento (e.g., modalità di accesso, con le dovute garanzie giuridiche, da parte delle autorità) e la conservazione (e.g., la durata massima, la c.d. *data retention*) dei dati sensibili applicate siano quelle italiane o al limite quelle europee, in entrambi i casi molto più garantiste (sebbene con ampi margini di miglioramento) di quelle statunitensi o cinesi (per fare due esempi a garanzia "civica" decrescente).

1.4 Naming compliance

I "nomi" (cioè gli URL, i cosiddetti domini registrati) dei siti dei Comuni dovrebbero seguire un *pattern* standardizzato, ad esempio "www.comune.milano.it", in modo da non dover necessariamente ricorrere ad un motore di ricerca per trovarli. Invero non esistono normative stringenti, solo alcune linee guida e regolamenti dell'AGID (Agenzia per l'Italia Digitale - <http://www.agid.gov.it>)

Il regolamento "Assegnazione e gestione dei nomi a dominio nel SLD gov.it" (AGID 2014) riserva una serie di nomi a dominio che possono essere utilizzati (sono preclusi ad altre entità) dai Comuni per la registrazione:

5.3.3 Comuni

L'elenco dei nomi a dominio con cui sono identificate le denominazioni dei Comuni italiani è contenuto nell'Appendice C del Regolamento "Assegnazione e gestione dei nomi a dominio nel ccTLD .it" emanato dal Registro ".it", che viene mantenuto aggiornato a seguito di ogni modifica di carattere istituzionale intervenuta. Tali nomi a dominio non sono assegnabili e sono parte integrante della struttura geografica predefinita. Essi permettono la registrazione e il mantenimento dei nomi a dominio al di sotto di essa.

...

Nomi a dominio riservati e non assegnabili

I nomi delle Regioni, delle Province e dei Comuni italiani non sono assegnabili e sono parte integrante della struttura geografica predefinita. Essi consentono la registrazione ed il mantenimento dei nomi a dominio al di sotto di essa. L'elenco dei nomi a dominio con cui sono identificate le denominazioni delle Regioni, delle Province e dei Comuni italiani è contenuto nell'Appendice A (Regioni), nell'Appendice B (Province) e nell'Appendice C (Comuni) del documento "Assegnazione e gestione dei nomi a dominio nel ccTLD .it" del Registro ".it". Tutti i nomi riportati nell'Appendice D del documento "Assegnazione e gestione dei nomi a dominio nel ccTLD .it" del Registro ".it" non possono essere assegnati.

Nello stesso documento si specifica:

Appendice C - Elenco dei nomi corrispondenti alle denominazioni dei comuni italiani L'elenco dei comuni, per semplicità, non è allegato ma è disponibile su un apposito file presente sul sito web del Registro.

Tale appendice purtroppo non è così facile da trovare, sulla pagina web del regolamento non è referenziata ("linkata") direttamente. Alcune iterazioni tramite motore di ricerca portano ad affinare le parole chiave per arrivare a cercare "Nomi riservati dei comuni italiani" e trovare finalmente sul sito del NIC (Network Information Center, l'ente che regola la nomenclatura sul web) il file <http://www.nic.it/sites/default/files/docs/Elenco%20comuni%20italiani.pdf> da cui estrarre, con un po' di elaborazione perché non è pensato per essere trattato "programmaticamente" essendo un PDF, l'elenco dei nomi riservati.

2. Tecniche di raccolta e analisi dei dati

2.1 Open data vs. "web-scraping"

I contenuti "open data" possono essere usati, modificati e condivisi liberamente da chiunque per qualunque scopo. [da: <http://opendefinition.org>]

"Publicity is justly commended as a remedy for social and industrial diseases. Sunlight is said to be the best of disinfectants; electric light the most efficient policeman."
[frase attribuita al giudice statunitense Louis D. Brandeis (1856-1941)]

L'open data, nel suo significato più ampio, è un movimento politico che tenta di spingere i governi, le aziende, le pubbliche amministrazioni e in generale qualunque ente tratti dati di interesse generale, verso la totale (sebbene con le dovute cautele) trasparenza delle informazioni. Secondo i principi di questo movimento, ogni entità organizzata (a qualunque livello, ma soprattutto - e ovviamente - quelle finanziate da fondi pubblici) dovrebbe pubblicare in rete ogni informazione sul proprio funzionamento interno. Informazioni come (si immaginino i contesti più disparati): spese, ruoli delle persone, attività pubbliche, *workflow* interni, verbali di assemblea, ordinanze e circolari, dati di pubblica (appunto!) utilità come orari degli uffici, geo-localizzazioni dei servizi (es. fermate autobus), informazioni commerciali (sedi imprese e informazioni correlate).

Quando disponibili, gli open data sono di solito raggiungibili attraverso siti web. Molte organizzazioni (le meglio... organizzate) addirittura creano siti dedicati al solo accesso degli open data (si vedano ad esempio data.gov USA, data.gov.uk UK, dati.gov.it Italia) invece di avere una ridotta sezione all'interno del proprio sito istituzionale. Questo perché l'accesso ai dati viene aiutato da piattaforme dedicate allo scopo con strumenti software per l'estrazione di sottoinsiemi di dati, l'organizzazione e finanche la visualizzazione diretta in forma di grafici.

Gli open data vengono categorizzati in classi in base al formato in cui vengono resi disponibili e alla licenza con cui sono distribuiti (Berners-Lee 2009). Inoltre (Davies 2010) stila una classifica ulteriore basata sugli aspetti funzionali dei dati oltre che al loro formato intrinseco. Solitamente gli open data sono scaricabili in *batch* (a lotti, sotto forma di più o meno grandi file compressi da espandere una volta trasferiti sul proprio computer), ma sempre più le organizzazioni che li rendono disponibili offrono anche un accesso diretto via API (Application Programming Interface), cioè fornendo un accesso remoto ai database che possono essere quindi consultati mediante programmi che accedono ai *dataset* direttamente via rete, evitando di doversi scaricare quantità ingenti di dati (a prezzo di dover essere connessi permanentemente alla rete).

In alcuni casi, per fortuna rari, organismi particolarmente "gelosi" dei propri dati, ma obbligati a renderli comunque pubblici a causa di legislazioni che spingono verso la trasparenza, optano per una pubblicazione "zoppa". Cioè rendono disponibili sì i dati, ma inserendo, sul cammino per raggiungerli, alcuni ostacoli (chiamati *webstables* da "web" + "obstacles") che aggiungono difficoltà artefatte a chi li voglia scaricare/utilizzare. Esempio lampante è il caso di ARPA Lombardia che permette(va) il download dei dati storici sullo stato dell'aria attraverso un *form* da compilare a mano, data la presenza di un CAPTCHA (Von Ahn et al. 2003) per ogni richiesta (limitata ad una singola stazione, un singolo inquinante e su un periodo massimo di un anno), costringendo gli utenti a migliaia di richieste manuali per ottenere tutti i dati. Di fronte a questi *webstables* qualche volta è possibile aggirare le limitazioni al download utilizzando tecniche cosiddette di *web scraping* (si veda nel seguito) cioè

tecniche di accesso software a pagine web standard, progettate per fruizione umana, con estrazione di dati che vengono poi aggregati programmaticamente a formare un dataset elaborabile, il cosiddetto “obtorto collo data” (Trentini 2014).

Nota bene: a fine 2017 ARPA ha finalmente rilasciato tutti i suoi *dataset* attraverso il portale *opendata* di Regione Lombardia, rimuovendo quindi ogni tipo di ostacolo alla fruizione.

2.2 Ottenere la lista dei siti da analizzare

La lista degli url dei siti web dei Comuni italiani si può recuperare dal sito web dell’ANCI (Associazione Nazionale Comuni Italiani). Si può produrre “a mano”, partendo da <http://www.ancitel.it/link/siti/index.cfm> e compilando il *form*, oppure implementando uno script di scraping. Lo script realizzato è basato sullo strumento *wget* che è, assieme a *curl*, il “coltellino svizzero” di chi fa questo tipo di analisi su web. Serve a scaricare “programmaticamente” una (o più, ricorsivamente) pagina web, senza dover usare un browser, per poterne analizzare, sempre programmaticamente, il contenuto.

Il risultato ottenuto è un file di testo di circa 8000 righe (tanti sono i comuni italiani) da utilizzare per gli *script* successivi. Piccola nota di colore: il sito dell’ANCI ha una sezione *opendata* nella quale gli autori hanno cercato, senza successo, l’elenco già pronto degli url dei siti comunali. Tra i *dataset*, ben 10 (<http://ckan.ancitel.it/dataset> ad agosto 2017), si trova un’unica voce promettente:

Comuni italiani

Dati territoriali e demografici dei Comuni italiani.

Elaborazione Ancitel su dati Istat (2016)

[file CSV]

Ma il dataset contiene solo i seguenti campi:

Comune, ISTAT, Provincia, SiglaProv, Regione, AreaGeo,
PopResidente, PopStraniera, DensitaDemografica,
SuperficieKmq, AltezzaCentro, AltezzaMinima,
AltezzaMassima, ZonaAltimetrica, TipoComune,
GradoUrbaniz, IndiceMontanita, ZonaClimatica,
ZonaSismica, ClasseComune, Latitudine, Longitudine

Tra i quali, come si può ben notare, non c’è l’URL del sito ufficiale. Mancanza notevole se si pensa che uno studio Ancitel (Comuniverso 2017)

riporta una statistica sui siti web dei comuni secondo cui praticamente tutti lo hanno, con una media del 99.71%.

Dopo ulteriore ricerca in rete, è stato trovato un altro dataset generale di enti pubblici (includente anche i Comuni, stavolta con gli URL dei siti) su <http://www.indicepa.gov.it/documentale/index.php>, dall'elenco è stata estratta, non senza qualche difficoltà tecnica dovuta alla cattiva strutturazione del file, la lista dei siti per effettuare un confronto con la lista ottenuta dal sito ANCI.

2.3 Scaricamento homepage siti comunali

La lista dei siti è stata impiegata per scaricare una pagina web per ogni comune su cui effettuare i test di accessibilità. Sebbene fosse possibile portare avanti analisi più approfondite, ad esempio scaricando più pagine presenti nelle sitemap o seguendo i link interni, si è ritenuto sufficiente campionare unicamente le pagine iniziali in quanto tipicamente punto di partenza per la navigazione degli utenti.

Partendo dall'elenco degli URL dei siti dei comuni è stato possibile effettuare il download delle pagine iniziali tramite un altro script, sempre basato su *wget*. La homepage di ogni comune è stata salvata in un file separato per poi passarlo allo strumento di analisi (cfr. Sezione 2.6). Per questioni di efficienza, tutti i siti con cui non è stato possibile stabilire una connessione entro 10 secondi - quindi molto probabilmente irraggiungibili - sono stati scartati.

2.4 Geo-compliance

Per indagare sulla geo-localizzazione di un sito l'unica informazione utilizzabile è il suo indirizzo IP (gli indirizzi "macchina" dei server in rete, ad esempio 8.8.8.8 di uno dei server DNS di Google). Attraverso strumenti standard come *geoiplookup* (Falahrastegar et al. 2014) è possibile, con alcuni gradi di incertezza, individuare la nazione in cui si trova un sito web.

L'incertezza è dovuta a due fattori:

1) il sistema di localizzazione è basato su meta informazioni redatte a cura dei provider;

2) i siti web potrebbero avvalersi delle cosiddette CDN (Content Delivery Network), reti "overlay" di distribuzione dei contenuti per cui il sito (i suoi dati) si potrebbe trovare in una nazione ma il suo frontend appare in un'altra.

In ogni caso è comunque interessante fare questo studio, pur impreciso, per capire le risposte dei gestori comunali, alcuni dei quali (come si vedrà più avanti) non sanno neppure dove risiede il proprio sito (sic!).

2.5 Naming compliance

L'analisi sulla registrazione dei nomi consta in realtà di varie parti:

- 1) il Comune ha un sito?
- 2) il Comune ha un sito il cui nome segue il pattern standard?
- 3) dato che i nomi registrabili sono vari (ad esempio: con e senza 'www'; 'comune.mi.it' vs. 'comune.milano.it'), il Comune ha registrato tutte le varianti e, se sì, tutte "risolvono" (puntano a) lo stesso contenuto?

Gli strumenti utilizzati in questo caso sono sostanzialmente solo due: *dig* e *nslookup*. Questi strumenti permettono l'interrogazione del DNS (Domain Name System). Il DNS è il meccanismo di database distribuito che gestisce la traduzione tra il nome simbolico di un sito web - e.g. www.google.it - e il suo indirizzo IP, che è poi l'informazione che serve al computer per contattare il server. I due strumenti, combinati in alcuni *script*, permettono la verifica dell'esistenza e i vari controlli di concordanza.

2.6 Strumenti per la validazione delle pagine web

Il sito del W3C mette a disposizione una lista comprendente un centinaio di software utilizzabili per la validazione di pagine web che, pur essendo originariamente pensati per fornire supporto ai programmatori in fase di sviluppo, sono anche direttamente fruibili dagli utenti finali. Gran parte di tali strumenti consente la verifica dei contenuti web rispetto alle linee guida (Reid e Snow-Weaver 2008) mentre alcuni, meno generici, consentono di verificare la conformità rispetto a varie normative nazionali quali (BITV), (RGAA), (JIS), (INITAG), (Section508) e (Galliani 2008).

Per quanto riguarda la normativa italiana, in particolare, la lista comprende due validatori commerciali - DYNO Mapper e Monsido - e tre gratuiti - Achecker, Monsido e Vamolà - questi ultimi due basati a loro volta su Achecker.

Achecker è uno strumento sviluppato all'Università di Toronto a partire dallo scorso decennio, (Gay e Li 2010) è un validatore libero e *opensource* in grado di testare la rispondenza di una pagina web rispetto a buona parte delle linee guida indicate nella normativa Stanca. Per tale motivo, è stato scelto come strumento di riferimento per analizzare le pagine iniziali dei siti comunali i cui risultati sono consultabili nella prossima sezione.

Risulta doveroso segnalare l'assenza di strumenti "ufficiali" sviluppati o quanto meno suggeriti dallo Stato per la verifica dei contenuti web. Tale mancanza non è "scusata" dall'esistenza di Achecker, a maggior ragione se si considera la sua discreta "vecchiaia": sviluppato in PHP 5.5 (oggi, nel 2017,

siamo alla versione 7.1.9), ufficialmente deprecato e non compatibile con le versioni successive.

3. Stato dei siti comunali

3.1 W3C compliance

Dalla validazione effettuata con Achecker è emersa la quasi totale indifferenza dei comuni al rispetto della legge Stanca: i siti web conformi alle regole previste dal W3C, infatti, sono risultati complessivamente meno del 10% del totale (Figura 1).

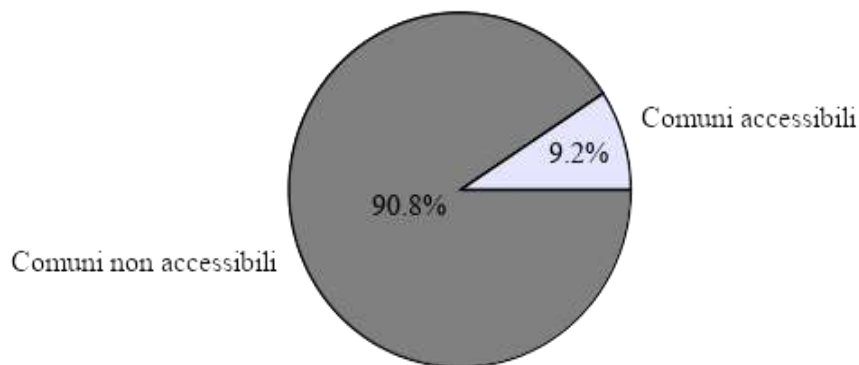


Figura 1: ITALIA – 8049 comuni - 738 accessibili

Come mostrato in Figura 2 la regione con la maggior percentuale di portali accessibili è il Trentino Alto Adige con il 36.5% di siti conformi allo standard (comprensibile, anche alla luce della standardizzazione, si veda in Sezione 3.2, del sistema di gestione dei contenuti “RIS-Kommunal”). La regione risultata più inaccessibile, con ben il 98.3% di siti che violano le WCAG 2.0, è la Calabria seguita dall’Umbria.

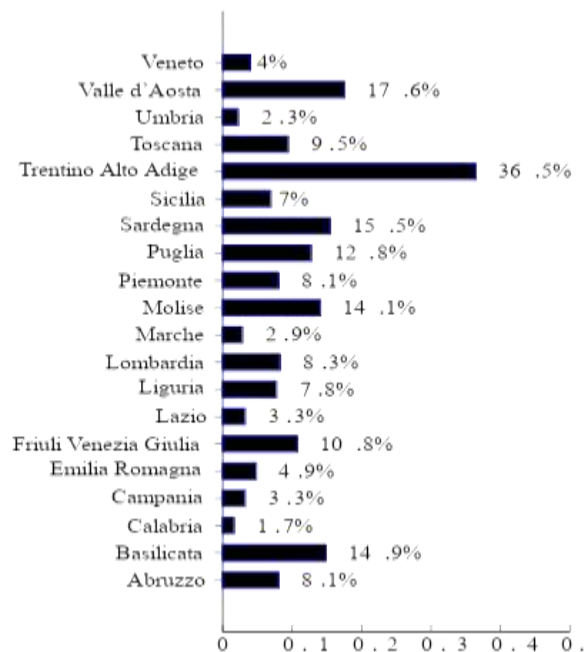


Figura 2: Comuni accessibili divisi per regione

La miglior provincia è risultata essere Bolzano con ben 101/116 siti web accessibili mentre le peggiori, con complessivamente 0/409 siti web conformi sono Agrigento, Crotone, Forli-Cesena, Latina, Lucca, Pesaro-Urbino, Reggio Emilia, Siena, Treviso e Trieste. In Figura 3 una suddivisione per latitudine dei soli siti accessibili.

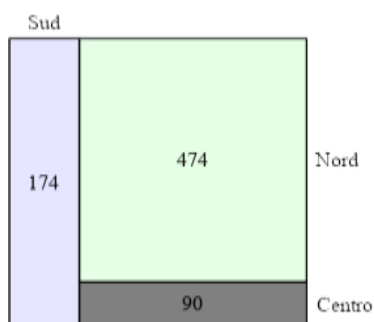


Figura 3: Suddivisione su base territoriale dei 738 siti accessibili

3.2 Geo-compliance e risposte dei comuni

Il risultato dell'esecuzione di *geoiplookup* sulla lista dei circa 8000 siti ha prodotto un *dataset* che contiene record simili ai seguenti:

Site;Location

http://castellaneta.gov.it;IT,Italy

http://comuneasuni.gov.it;IT,Italy

http://comuneazzanosanpaolo.gov.it;IT,Italy

...

http://comune.ivano-fracena.tn.it;DE,Germany

...

http://www.casteldelrio.provincia.bologna.it;DK,Denmark

...

http://www.comune.alicebelcolle.al.it;US,UnitedStates

...

La distribuzione per nazione di provenienza è la seguente:

7286: IT, Italy

212: US, United States

194: DE, Germany

126: FR, France

42: IE, Ireland

31: GB, United Kingdom

23: IP Address not found

19: CH, Switzerland

17: NL, Netherlands

13: RO, Romania

9: CA, Canada

5: JP, Japan

4: EU, Europe

3: AT, Austria

1: SG, Singapore

1: PT, Portugal

1: DK, Denmark

1: CZ, Czech Republic

In grassetto le provenienze extra EU, sono circa 250 siti su circa 8000, cioè troviamo il 3% dei siti risultanti apparentemente (ricordiamo che *geoiplookup* non è infallibile) inadempienti.

La componente “umana” dello studio è stata effettuata tentando di contattare i sindaci e i referenti dei siti non conformi per capire se, come e quanti di essi avrebbero risposto alla nostra “segnalazione civica” di non conformità. Sono state inviate le mail a ‘sindaco@<dominio>’ e ad altri indirizzi standard (‘web@’, ‘admin@’, ‘info@’, ecc.) in caso di mancata risposta (o mancata ricevuta di ritorno).

Purtroppo circa la metà degli invii a ‘sindaco@’ hanno generato un “*user unknown error*” (utente sconosciuto, sic!), gli altri indirizzi invece sono stati quantomeno inoltrati (“letti effettivamente” è un altro discorso).

Aneddoto tristemente umoristico, un paio di invii a ‘sindaco@’ hanno generato un “*mailbox full*” (casella piena, sic!), cioè l’utente esiste (è definito nel sistema), ma non legge regolarmente la posta elettronica.

Il “successo” dell’iniziativa ha generato “ben” 23 risposte, arrivate con la seguente cronologia:

1) stesso giorno, 5 risposte: Villa Lagarina, Pescasseroli, Binasco, Malgrate, Maniago

2) giorno successivo, 9 risposte: Prato allo Stelvio, Conca dei Marini, Azzate, Malnate, Voghera, Capranica, Vadena, Ronco Scrivia, Filago

3) i gg seguenti, altre 9 risposte: Ponte San Pietro, Sesto (BZ), Egna, Sondrio, Baveno, Noceto, Centola, Poncarale, Malles Venosta

Il tono della maggior parte delle risposte era un generico “Grazie per l’informazione, stiamo verificando la segnalazione con il nostro reparto IT...”, pochi di questi hanno fatto seguito con una mail ulteriore che segnalava l’esito della verifica.

Qui di seguito l’esito delle verifiche come “falsi positivi”:

1) Maniago, sembrava US, confermato UK;

2) (N.B. tutti questi appartenenti al dominio ‘bz.it’) Egna, Prad, Malles, Gais, Sixten, Vadena. Sembrava US, confermato EU;

3) Villalagarina, sembrava US, confermato UK.

Molte risposte provenienti dalla regione di Bolzano hanno fornito una risposta tecnica e concorde, tutti i Comuni della zona utilizzano un sistema CMS (Content Management System) chiamato “RIS-Kommunal” (<http://info.riskommunal.net>) e una CDN (Content Delivery Network) che fa apparire il sito come proveniente da fuori EU, ma garantiscono che i dati sono locati in EU.

I seguenti invece hanno confermato (ringraziando della segnalazione) che i siti erano fuori EU e che era in corso il processo di trasferimento in EU: Malgrate, Binasco, Voghera, Sondrio, Noceta, Conca.

3.3 Naming compliance

L’elenco dei nomi riservati estratto dall’“Appendice C” (cfr. Sezione 1.4) consta di poco meno di 40 mila nomi, esattamente 39.754. A fronte dei circa 8000 Comuni italiani può sembrare troppo, ma si pensi che sono state previste tutte le varianti brevi e lunghe (e.g., abbadia-lariana.lc.it, abbadialariana.lc.it, abbadialariana.lecco.it, abbadialariana.lecco.it, abbadia.lc.it, abbadia.lecco.it). Di questi, 24.315 non sono registrati quindi scendiamo a 15.439 registrati, alcuni dei quali sono riferiti allo stesso sito effettivo. A questo punto dovremmo indagare quanti dei circa 8000 Comuni utilizzano un nome standardizzato. Dall’elenco estratto dal sito ANCI (cfr. Sezione 2.2) si può verificare che solo circa 7000 sono

presenti nell'“Appendice C”, cioè ci sono 1000 siti comunali che usano nomi non standard, il 12.5%.

Un tema collegato alla registrazione del nome del sito è quello della corrispondenza (per quanto riguarda i contenuti mostrati) fra l'URL con e senza preposto 'www'. Consuetudine vuole che il vero nome di dominio registrato sia il nome senza 'www' mentre l'URL con 'www' venga configurato come *alias*.

Purtroppo, durante il lavoro di analisi, si è scoperto che una percentuale elevata di siti web risolvevano addirittura indirizzi diversi e/o mostravano contenuti molto diversi a seconda che si utilizzasse il 'www' o meno.

In particolare, per fortuna, circa l'80% dei nomi risolve allo stesso modo con e senza 'www' preposto. Quindi sembra che esista un “buon” 20% di siti che mostrano contenuti diversi a seconda dell'URL utilizzato. In realtà, dato che è normale per un web-server presentare contenuti lievemente diversi in funzione dell'URL utilizzato per raggiungere il sito, bisogna esaminare un po' più approfonditamente le differenze di contenuto. Sempre attraverso strumenti standard (*diff* in questo caso) è stato possibile analizzare la dimensione in righe della differenza fra il contenuto presentato dal sito con e senza 'www'.

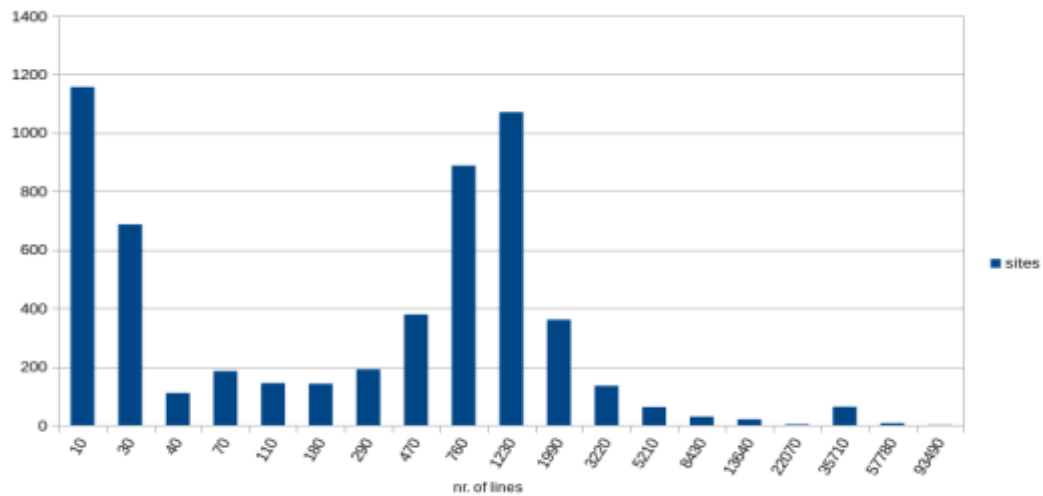


Figura 4: Differenza “www vs. senza-www” in numero di linee

In Figura 4 viene mostrata la distribuzione delle differenze, posizionando la soglia di “attenzione” sulle 100 linee di differenza si trova che circa il 36% dei siti web comunali mostra contenuti molto diversi a seconda che lo si raggiunga tramite un URL con o senza 'www'. Per confermare la cifra del 20% di cui sopra bisogna selezionare una soglia di diversità nell'intorno delle 300 righe. La scelta della soglia è arbitraria, per capire se effettivamente il contenuto è diverso andrebbero esaminate le pagine “a occhio umano” o con strumenti algoritmici di valutazione della somiglianza grafica.

Alcuni siti, tra quelli “molto diversi”, senza ‘www’ forniscono semplicemente una *homepage* di *default*, ad esempio (a Dicembre 2016) i seguenti:

<http://cassinadepecchi.gov.it> - Microsoft IIS 8

<http://cittadicapri.it> - Microsoft IIS 8

<http://comune.licodiaeubea.ct-egov.it> - Microsoft IIS 8

<http://comune.ragalna.ct-egov.it> - Microsoft IIS 8

<http://comune.trecastagni.ct.it> - Microsoft IIS 8

<http://comune.bonemerse.cr.it> - It works! (Apache)

<http://comune.capoliveri.li.it> - It works! (Apache)

<http://comune.castellina.si.it> - It works! (Apache)

<http://comune.castelsaraceno.pz.it> - IT WORKS! (Apache)

<http://comune.forlimpopoli.fc.it> - It works! (Apache)

<http://comunegorlamaggiore.it> - BGweb it works! (Apache?)

<http://comune.lanuvio.rm.it> - It works! (Apache)

<http://comune.mozzate.co.it> - It works! (Apache)

<http://comune.palazzo.pz.it> - IT WORKS! (Apache)

<http://comune.pessinacremonese.cr.it> - It works! (Apache)

<http://comune.pisticci.mt.it> - IT WORKS! (Apache)

<http://comune.pomarico.mt.it> - IT WORKS! (Apache)

<http://comune.quartusantelena.ca.it> - It works! (Apache)

<http://comune.roccadipapa.rm.it> - It works! (Apache)

<http://comune.valsinni.mt.it> - IT WORKS! (Apache)

L’ultima nota eclatante ce la fornisce il comune di Binasco. L’URL <http://www.comune.binasco.mi.it> risolve e indirizza il sito web del Comune, ma l’URL senza ‘www’, <http://comune.binasco.mi.it>, risolve e indirizza il sito di una azienda privata... e nemmeno quella che ha realizzato il sito (sic!).



Figura 4: Binasco senza 'www' (archive.org, 30 Aprile 2016)

4. Conclusioni

La situazione dei siti web dei Comuni italiani non è rosea. In questo studio sono stati analizzati alcuni aspetti tecnici:

- 1) accessibilità e conformità alle normative e linee guida internazionali;
- 2) gestione della registrazione dei nomi di dominio e corretta configurazione del web-server;

Il primo aspetto è quasi totalmente negletto (10% conforme, 90% difforme, cfr. Sezione 3.2). Il secondo invece mostra una statistica tutto sommato positiva:

- 1) solo il 3% dei siti non risiede in EU (Sezione 3.2);
- 2) solo il 12.5% dei siti non usa nomi standard (Sezione 3.3);
- 3) una percentuale intorno al 20% offre contenuti diversi tra “con www” e “senza www” (idem).

Con alcune note aneddotiche al limite del ridicolo, specie in relazione alla configurazione dei siti (“con www” vs. “senza www”). Paradossalmente, la situazione è migliore nel campo, il *naming*, in cui la normativa è meno stringente e più difficile da reperire (cfr. Sezioni 1.3 e 1.4).

Perché l’aderenza agli standard citati (ma ne esistono anche altri applicabili alla struttura del sito, ai servizi digitali che dovrebbero essere presenti, ecc.) dovrebbe avvantaggiare la “smart”-cittadinanza? Perché un sito

comunale facilmente individuabile (senza ricorrere a motori di ricerca) è facilmente raggiungibile da chiunque ed è facilmente identificabile come sito ufficiale, viviamo pur sempre nell'era del *domain squatting* (Illia 2003). Se poi il suo contenuto è "accessibile" (nel senso W3C) è anche "accessibile" (nel senso della lingua italiana) da un maggior numero di utenti (disabili o meno) che arrivano sul sito tramite i *device* più disparati potendo fruire dei servizi offerti senza difficoltà: usare un sito non adattivo su uno *smartphone* è una esperienza molto frustrante e spesso scoraggia definitivamente dall'uso del servizio online (Venkatesh, Ramesh, and Massey 2003).

Bibliografia

- AGID. 2014. "Assegnazione E Gestione Dei Nomi a Dominio Nel SLD Gov.it." Presidenza del Consiglio dei Ministri. <http://www.agid.gov.it/notizie/2014/05/14/aggiornato-il-regolamento-lassegnazione-del-dominio-govit>.
- AGID. 2017. "Il Sistema Di Conservazione Garantisce Autenticità, Integrità, Affidabilità, Leggibilità E Reperibilità Dei Documenti Informatici, Come Previsto Dal CAD (Art.44)." Presidenza del Consiglio dei Ministri.
- Barricelli, Barbara Rita, Pierlauro Sciarelli, Stefano Valtolina, and Alessandro Rizzi. 2017. "Web Accessibility Legislation in Italy: A Survey 10 Years After the Stanca Act." *Universal Access in the Information Society*. Springer, 1–12.
- Berners-Lee, Tim. 2009. "Linked Data." <http://www.w3.org/DesignIssues/LinkedData.html>; WWW Consortium.
- "BITV." n.d. Germany. <http://www.bitvtest.eu>.
- Comuniverso. 2017. "Dati Regionali Sui Comuni Con Sito Web." http://www.comuniverso.it/index.cfm?Dati_regionali_sui_Comuni_con_sito_web&menu=244.
- Consiglio dei Ministri, Presidente del. 2014. "DECRETO Del Presidente Del Consiglio Dei Ministri 3 Dicembre 2013." <http://www.gazzettaufficiale.it/eli/id/2014/03/12/14A02098/sg>.
- Davies, Tim, and Practical Participation. 2010. *Open Data, Democracy and Public Sector Reform: A Look at Open Government Data Use from Data. Gov. Uk. Practical Participation*.
- De Cindio, Fiorella, and Andrea Trentini. 2014. "A Layered Architecture to Model Digital Citizenship Rights and Opportunities." In *Conference for E-Democracy and Open Governement*, 403.
- Falahrastegar, Marjan, Hamed Haddadi, Steve Uhlig, and Richard Mortier. 2014. "Anatomy of the Third-Party Web Tracking Ecosystem." arXiv Preprint arXiv:1409.1066.
- Galliani, D. 2008. *L'accessibilità Ai Siti Internet Delle Pubbliche Amministrazioni E La Cd. "Legge Stanca"*. Giuffrè.
- Gay, Greg, and Cindy Qi Li. 2010. "AChecker: Open, Interactive, Customizable, Web Accessibility Checking." In *Proceedings of the 2010 International Cross Disciplinary*

- Conference on Web Accessibility (W4a), 23:1–23:2. W4A '10. New York, NY, USA: ACM. doi:10.1145/1805986.1806019.
- Illia, Laura. 2003. "Passage to Cyberactivism: How Dynamics of Activism Change." *Journal of Public Affairs* 3 (4). Wiley Online Library: 326–37.
- "INITAG." n.d. Ireland. <http://universaldesign.ie/Technology-ICT/Irish-National-IT-Accessibility-Guidelines/>.
- "JIS." n.d. Japan. <http://waic.jp/docs/jis2016/>.
- Pizzetti, Franco. 2016. *Privacy E Il Diritto Europeo Alla Protezione Dei Dati Personali: Dalla Direttiva 95/46 Al Nuovo Regolamento Europeo*. G Giappichelli Editore.
- Reid, Loretta Guarino, and Andi Snow-Weaver. 2008. "WCAG 2.0: A Web Accessibility Standard for the Evolving Web." In *Proceedings of the 2008 International Cross-Disciplinary Conference on Web Accessibility (W4a)*, 109–15. W4A '08. New York, NY, USA: ACM. doi:10.1145/1368044.1368069.
- Repubblica Italiana. 2014. *Misure Urgenti Per La Semplificazione E La Trasparenza Amministrativa E Per L'efficienza Degli Uffici Giudiziari*. Gazzetta Ufficiale.
- "RGAA." n.d. France. <https://references.modernisation.gouv.fr/rgaa-accessibilite/>.
- "Section508." n.d. USA. <https://www.section508.gov/>.
- Trentini, Andrea. 2014. "Lombardy EPA Obtorto Collo Data and Anti-Pollution Policies Fallacies." *Journal of E-Learning and Knowledge Society* 10 (2).
- Venkatesh, Viswanath, Venkataraman Ramesh, and Anne P Massey. 2003. "Understanding Usability in Mobile Commerce." *Communications of the ACM* 46 (12). ACM: 53–56.
- Von Ahn, Luis, Manuel Blum, Nicholas J Hopper, and John Langford. 2003. "CAPTCHA: Using Hard Ai Problems for Security." In *International Conference on the Theory and Applications of Cryptographic Techniques*, 294–311. Springer.
- Web Accessibility Initiative, and others. 1999. "Web Content Accessibility Guidelines 1.0." World Wide Web Consortium. Available Online at: <Http://Www.w3.org/TR/Wcag10/> (Accessed 28 September 2005).

Sommario

L'Italia è stata una delle prime nazioni ad includere nel proprio ordinamento una serie di leggi contenenti linee guida per la conformità dei portali istituzionali. Il presente articolo espone i risultati di uno studio sullo stato attuale dei siti municipali in termini di accessibilità, denominazione e distribuzione geografica con l'obiettivo di verificare l'effettivo grado di adozione delle normative vigenti. Le statistiche presentate sono state raccolte in due fasi distinte: in primo luogo i portali comunali sono stati scaricati e validati secondo le specifiche della cosiddetta legge Stanca; in seguito, sono state analizzate le caratteristiche dei nomi di dominio e la localizzazione delle relative risorse. Per quanto riguarda l'accessibilità, lo studio ha evidenziato un alto tasso di non conformità – solo il 10% del totale dei siti municipali - mentre si sono mostrati promettenti i risultati riguardanti denominazione e localizzazione geografica dei dati comunali.

Keywords

Accessibility, Compliance, Government Web sites, Web Content Accessibility Guidelines, Government regulations, Accessibility standards, Naming

Parole chiave

Accessibilità, conformità, siti istituzionali, accessibilità web, normative, linee guida, standard di accessibilità, nomi di dominio