

Strumenti per il reperimento di informazioni rilevanti in Rete

Applicazioni nel dominio della comunicazione pubblica in
materia di immigrazione in Italia

GERARDO GIARDIELLO, CHIARA FIORAVANTI,
FRANCESCO ROMANO, MARIASOLE RINALDI*

ABSTRACT: Access to legal information is a fundamental means of guaranteeing people the protection of their rights and the knowledge of their obligations under the law. This access often takes place via thematic public web portals. A legal informatics tool aimed to facilitate the daily monitoring of institutional online information about immigration procedures and legislation has been designed and developed. In this paper we will show the information resources to be monitored, the monitoring tool features and the emerging results during the first months of use of the tool. The emerging difficulties and the ways to fix them will also be highlighted, offering ideas for future developments.

Keywords: Public platforms, Public communication, Immigration, Digital forensics, Web crawler.

1. Il contesto e i bisogni informativi

L'accesso alle risorse informative nell'odierna società dell'informazione e della conoscenza è sempre più da considerarsi un bene primario, che infatti come tale è promosso da accordi e convenzioni internazionali. L'Agenda ONU 2030, ad esempio, prevede tra i propri obiettivi anche quello di «garantire un pubblico accesso all'informazione e proteggere le libertà fondamentali, in conformità con la legislazione nazionale e con gli accordi internazionali»¹.

In un contesto, come quello attuale, in cui le informazioni sono molte e in cui si rischia che il mondo possa «diventare illeggibile per overdose di informazioni» (Galimberti 2020, 117), diviene fondamentale in ambito pubblico disporre di informazioni rilevanti, aggiornate e derivanti da fonti attendibili².

* Istituto di Informatica Giuridica e Sistemi Giudiziari del CNR. gerardo.giardiello@igsg.cnr.it, chiara.fioravanti@igsg.cnr.it, francesco.romano@igsg.cnr.it, mariasole.rinaldi@igsg.cnr.it.

¹ Sui rischi di società in cui i cittadini non siano informati e sul formarsi di élite in grado di dirigere lo Stato e la società, si veda anche (Nichols 2017, 216).

² L'articolo 53 (Siti Internet delle pubbliche amministrazioni) del Codice dell'amministrazione digitale (D.Lgs 7 marzo 2005, n. 82, ss.mm.ii.) prevede, al comma 1, che le «pubbliche

L'accesso all'informazione giuridica, in particolare, diviene poi un irrinunciabile strumento per garantire la possibilità di tutelare i propri diritti e conoscere i propri doveri, specie in società sempre più aperte e come tali soggette all'arrivo di nuovi soggetti, che hanno la necessità primaria di conoscere le norme del Paese di accoglienza³.

In particolare, la reperibilità delle norme è riconosciuta come diritto dei cittadini, i dati giuridici sono collocati tra i dati pubblici essenziali ed è ormai condivisa la necessità di prevedere azioni finalizzate a consentirne l'accessibilità telematica. Si comprende, quindi, la necessità di creare raccolte normative costantemente aggiornate al fine di fornire ai soggetti interessati uno strumento immediato di supporto al lavoro e di accesso all'informazione desiderata⁴.

Dall'esigenza di fornire un accesso unico e attendibile all'informazione giuridica in Rete nel dominio dell'immigrazione, nasce nel 2004 il Progetto PA-eSI (Pubblica Amministrazione e Stranieri Immigrati), frutto della collaborazione tra Regione Toscana, Prefettura di Firenze e Istituto di Teoria e Tecniche dell'Informazione Giuridica (ITTIG) del Consiglio Nazionale delle Ricerche (CNR), nell'ambito del Piano di azione regionale e Toscana, programma di sviluppo dell'e-government della Regione Toscana (Fioravanti e Rinaldi 2010, 93–131) (Fig. 1).

Il progetto si poneva l'obiettivo di contribuire alla diffusione e standardizzazione delle informazioni online su procedure e norme in materia di immigrazione, così da supportare il lavoro degli operatori della pubblica amministrazione e del privato sociale che assistono i cittadini stranieri nelle pratiche burocratiche necessarie per vivere e lavorare in Italia. Tale compito risulta,

amministrazioni realizzano siti istituzionali su reti telematiche che rispettano i principi di accessibilità, nonché di elevata usabilità e reperibilità, anche da parte delle persone disabili, completezza di informazione, chiarezza di linguaggio, affidabilità, semplicità di consultazione, qualità, omogeneità ed interoperabilità (...). Inoltre l'articolo 6 decreto legislativo n. 33 del 2013 prevede che «Le pubbliche amministrazioni garantiscono la qualità delle informazioni riportate nei siti istituzionali nel rispetto degli obblighi di pubblicazione previsti dalla legge, assicurandone l'integrità, il costante aggiornamento, la completezza, la tempestività, la semplicità di consultazione, la comprensibilità, l'omogeneità, la facile accessibilità (...).»

³ Sulla necessità che la comunicazione istituzionale adempia alla sua «funzione di collante sociale» si veda (Bertolo 2005, 49).

⁴ La necessità che i migranti siano in grado di conoscere le norme dei Paesi di arrivo è un'esigenza sancita in apposite convenzioni internazionali ormai abbastanza datate, come la Convenzione OIL n. 143 del 24 giugno 1975 (sulle migrazioni in condizioni abusive e sulla promozione della parità di opportunità e di trattamento dei lavoratori migranti), che all'articolo 12 invitava i Paesi ratificanti a fare in modo che i lavoratori migranti possano conoscere i loro diritti, i loro obblighi e le attività destinate a dar loro effettiva assistenza per garantire la loro protezione e permettere loro di esercitare i propri diritti. Tale diritto è stato anche ribadito nella Convenzione internazionale sulla protezione dei diritti di tutti i lavoratori migranti e dei membri delle loro famiglie del 1990 all'articolo 33. Sui diritti degli stranieri in Italia si veda anche (Bonetti 2011).

infatti, piuttosto gravoso, sia per la grande complessità e variabilità della materia, sia per il fatto che le competenze sulle diverse procedure sono suddivise fra numerose amministrazioni, nazionali e territoriali, con la conseguenza che le informazioni sono comunicate in modo frammentato nei rispettivi diversi siti⁵.

Il progetto si è concretizzato nella predisposizione di un portale web⁶, tutt'ora aggiornato e sviluppato, che fornisce informazioni di livello nazionale e territoriale (per la Regione Toscana) su procedure, norme e servizi in materia di immigrazione attraverso due distinti sotto-siti, uno indirizzato agli operatori e l'altro agli stranieri, con contenuti progettati *ad hoc* per le due diverse utenze⁷.

La realizzazione di questo punto di accesso unico ha comportato e comporta tuttora, una collaborazione continua con i funzionari che operano negli uffici territoriali degli enti responsabili dei diversi procedimenti (Questura, Prefettura, Camera di Commercio, Comuni, etc.), nonché un'attività giornaliera di monitoraggio dell'informazione istituzionale su procedure e norme resa disponibile nei siti di diversi enti.

Per quanto riguarda le informazioni sulle procedure amministrative rivolte agli operatori di settore, il Portale fornisce un database di schede informative che descrivono in dettaglio le diverse procedure che interessano i cittadini stranieri, indicando tempi del procedimento, dati del responsabile e dell'ufficio, requisiti necessari per l'attivazione, documentazione richiesta, moduli occorrenti, modalità di presentazione e ricezione del provvedimento, riferimenti normativi e domande e risposte frequenti⁸. Tali contenuti richiedono un aggiornamento costante alla luce delle nuove norme o di modifiche organizzative territoriali. L'aggiornamento e la creazione di nuovi contenuti interessano, inoltre, tutte le sezioni tematiche del Portale, a partire da quelle dedicate alle novità normative o procedurali, ai nuovi servizi, ai progetti e alle opportunità nazionali e territoriali⁹.

⁵ Sulla comunicazione amministrativa via web, la comunicazione a cittadini stranieri nei siti istituzionali, nonché sulla pubblicazione e la trasparenza delle informazioni nei siti istituzionali comunali, si vedano vari saggi in (Cattani 2018).

⁶ www.immigrazione.regione.toscana.it/ (ultima consultazione: 30/05/2022).

⁷ Sulla necessità di selezionare solo le informazioni per il proprio destinatario si veda, ad esempio, (Mason, Zilio 2019, 61) e anche (Barker 2005).

⁸ Si tratta di informazioni utili per districarsi nei meandri della burocrazia e dei servizi pubblici italiani, un tipo di informazione definita come «comunicazione di servizio» (Arena 1997).

⁹ La fruibilità e l'accessibilità dei contenuti di siti che si occupano di fornire informazioni a questo target di utenti è tema da tempo dibattuto, così come la necessità che la pubblica amministrazione acquisisca competenze comunicative di tipo interculturale. Si veda su questi temi specifici (Ducci 2012, 40–42) e anche (Ciancio 2014, 136).

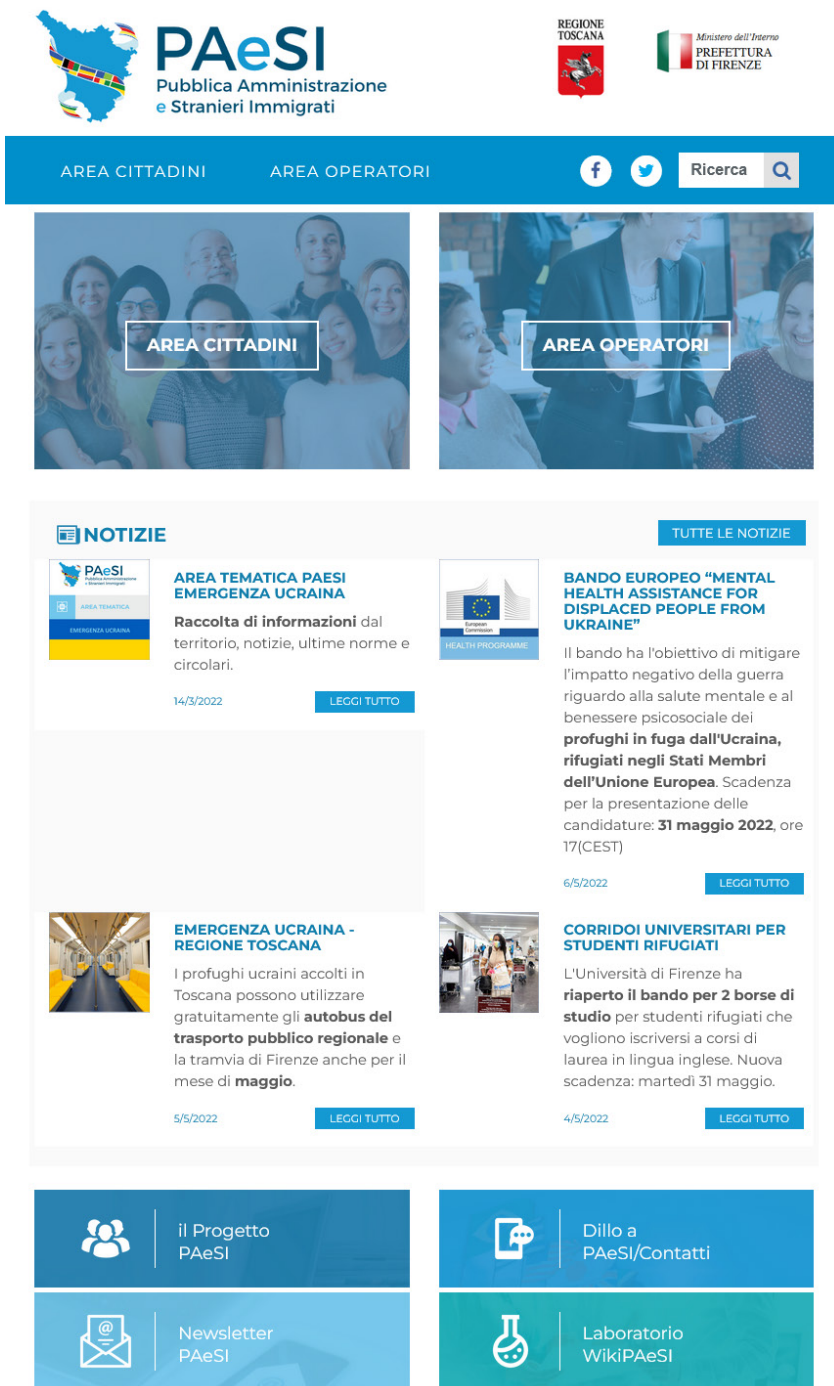


Figura 1: Immagine della home del Portale PAeSI.

Per quanto riguarda invece il reperimento delle norme che riguardano il dominio dell'immigrazione, il Portale, sin dalla sua nascita, ha dedicato una rilevante attività all'implementazione di un database di normativa nazionale, europea/internazionale, regionale e più recentemente anche comunale. Gestire l'informazione normativa nel settore dell'immigrazione presenta numerose problematiche e criticità, prima fra tutte, la necessità di rimanere aggiornati con la costante introduzione di nuove norme nell'ordinamento italiano. La tendenza della Pubblica Amministrazione a trattare, alle volte con eccessiva discrezionalità, le procedure, tanto da creare disomogeneità operativa tra gli stessi Enti presenti in province diverse, è un ulteriore elemento che complica il quadro di riferimento. Da qui deriva il sentito bisogno di offrire una raccolta normativa nazionale di riferimento il più possibile chiara e facilmente consultabile. Esiste infine una terza criticità, legata all'interpretazione normativa, che rende decisamente necessaria la pubblicazione in banca dati di circolari ministeriali esplicative, prevalentemente del Ministero dell'Interno, al fine di poter avere una chiave di lettura e un'interpretazione meno equivocabile della normativa in questione.

Il lavoro incessante di incremento e mantenimento della banca dati consiste dunque nell'introduzione, preceduta dal processo di reperimento e selezione, di nuova normativa e nell'aggiornamento di quella già presente, permettendone la visualizzazione, nelle differenti versioni del testo nel tempo¹⁰.

Tutto ciò rende pertanto fondamentale la consultazione del sito della Gazzetta Ufficiale e di molti altri siti istituzionali in cui reperire informazioni aggiornate.

Le norme in oggetto vengono selezionate nel testo originale e poi convertite in modo automatico o semi automatico in formato eXtensible Markup Language (XML) (Sartor et al. 2011). Tale linguaggio costituisce uno standard per la rappresentazione delle norme che fornisce il vantaggio di poter strutturare il contenuto di un documento utilizzando marcatori personalizzati attraverso un editore specializzato¹¹. Successivamente il testo viene analizzato con un *parser*, che individuando nel testo di legge i riferimenti testuali espliciti, crea in modo automatico link ipertestuali in base ad una sintassi predefinita¹². Inoltre, tale

¹⁰ La Direttiva (UE) 2016/2102 ha introdotto un concetto di accessibilità che va inteso non solo come «l'insieme di principi e tecniche da rispettare nella progettazione, costruzione» di siti web ma anche nella loro «manutenzione e aggiornamento». Si veda sul punto (Caporale 2019, 360).

¹¹ Si veda Circolare 22 aprile 2002 n. AIPA/CR/40 *Formato per la rappresentazione elettronica dei provvedimenti normativi tramite il linguaggio di marcatura XML*. Vedi (Ciampi et al. 2001).

¹² Si tratta della sintassi stabilita dallo standard Uniform Resource Name (URN). Lo standard serve a identificare le risorse normative in modo univoco e persistente indipendentemente dalla loro collocazione fisica in rete, si veda Circolare del 6 novembre 2001, n. AIPA/CR/35 *Assegnazione dei nomi uniformi ai documenti giuridici*. Vedi (Francesconi et al. 2010, 233-52).

standard consente anche di individuare le disposizioni di modifica presenti nel testo di una legge, procedendo al relativo trattamento. Nella banca dati dunque, i provvedimenti sono consultabili nel testo originale, in quello vigente ad una data intermedia, ovvero in una forma redazionale cosiddetta *multivigente* in cui le modifiche intervenute sul testo, fin dalla sua emanazione, sono tutte riportate e visibili nella loro interezza¹³.

Proprio per facilitare l'attività giornaliera di monitoraggio dell'informazione istituzionale online su procedure e norme, che per molti anni è stata svolta manualmente, è stato progettato e realizzato uno strumento di informatica-giuridica, che ha reso l'operazione più veloce ed efficiente ed è oggetto del presente contributo.

Nel presente articolo illustreremo le risorse informative che costituiscono il flusso da monitorare, lo strumento di monitoraggio, le sue funzionalità e i risultati emersi nei primi mesi d'uso dello strumento.

In conclusione, saranno evidenziate le criticità riscontrate e le modalità per poterle gestire, offrendo spunti per possibili sviluppi futuri.

2. Le risorse informative per il Portale PAeSI

Come detto, per mettere a disposizione dell'utenza le informazioni provenienti da diverse fonti istituzionali attraverso i contenuti del Portale PAeSI, risulta necessario svolgere giornalmente un'attività di monitoraggio dei siti degli enti e degli altri soggetti pubblici che hanno competenza in materia.

Per quanto riguarda l'informazione online utile allo sviluppo e aggiornamento dei contenuti informativi su procedure, servizi e opportunità, è necessario innanzitutto monitorare l'informazione di fonte statale. Vengono in particolare controllati i siti del Ministero dell'Interno nella sezione "Immigrazione e Asilo", del Ministero del Lavoro e delle Politiche Sociali (tema immigrazione) e del Ministero dell'Istruzione per le informazioni sull'accesso allo studio in Italia dei cittadini stranieri.

I tre Ministeri hanno inoltre realizzato nel 2012 un portale tematico (integrazioneemigranti.gov.it) per favorire l'integrazione dei migranti attraverso la diffusione di informazioni sui servizi pubblici nazionali e territoriali.

D'interesse anche i siti del Ministero della Salute per quanto riguarda la tutela della salute dei migranti, del Ministero degli Esteri per comunicazioni sui servizi consolari e le opportunità per lo studio in Italia e del Ministero della Giustizia per le informazioni sul riconoscimento di alcuni titoli di studio e il tema dei detenuti stranieri.

Il sito del Dipartimento delle Pari Opportunità della Presidenza del Consiglio dei Ministri viene controllato per verificare le iniziative in materia di

¹³ Per le modalità di marcatura di tali documenti si veda (Francesconi 2012).

contrasto alla tratta e alle mutilazioni genitali femminili ed il sito di UNAR, l'Ufficio Nazionale Antidiscriminazione Razziale, per quelle relative al contrasto alle discriminazioni.

Il sito del SAI (Sistema Accoglienza Integrazione), il sistema nazionale di protezione per titolari di protezione internazionale e per minori stranieri non accompagnati, deve essere monitorato per i report, le risorse formative, i progetti e gli avvisi su accoglienza e protezione internazionale.

Informazioni rilevanti vengono ricavate anche dal sito dell'Istituto Nazionale della Previdenza Sociale (INPS) nella specifica sezione dedicata ai lavoratori migranti e dal sito Istituto nazionale Assicurazione Infortuni sul Lavoro (INAIL) per quanto riguarda il tema della sicurezza sul lavoro.

A livello territoriale, avendo il progetto PAeSI come riferimento informativo la Toscana, sono monitorati i siti degli enti che sul territorio sono responsabili delle procedure nazionali, come le Prefetture e le Questure delle province toscane.

In ambito regionale, poi, oltre al sito della Regione Toscana, sono sottoposti a monitoraggio il sito del Centro regionale di documentazione per l'infanzia e per l'adolescenza, per le informazioni relative ai minori stranieri non accompagnati in Toscana e quello Centro Salute Globale della Regione Toscana, in tema di diritto alla salute dei migranti.

Per quanto riguarda l'accesso dei cittadini stranieri al servizio sanitario regionale, risultano d'interesse i siti delle tre Aziende usl toscane (Toscana centro, Toscana nord-ovest e Toscana sud-est). Per informazioni territoriali in materia di lavoro autonomo, vengono controllati i siti delle Camere di Commercio delle diverse province. In materia di lavoro d'interesse anche i siti delle associazioni di categoria, come Confartigianato e Confagricoltura toscana.

Il monitoraggio territoriale comprende poi i siti dei Comuni per le informazioni sulle procedure di anagrafe, stato civile e cittadinanza di loro competenza, sui servizi informativi e di assistenza alla compilazione delle pratiche per i cittadini stranieri e sulle iniziative comunali in materia di integrazione.

Le risorse informative da monitorare, dal punto di vista normativo, legate al tema dell'immigrazione sono molteplici in quanto nel sito PAeSI sono presenti risorse nazionali, internazionali, europee e locali (Toscana).

Le norme, infatti, devono essere costantemente monitorate sia per inserirne di nuove, sia per aggiornare quelle già presenti nel database.

Per questo tipo di lavoro il principale punto di riferimento è sicuramente rappresentato dal sito web della Gazzetta Ufficiale. Questa rappresenta la fonte ufficiale di conoscenza delle norme in vigore in Italia e lo strumento di diffusione, informazione e ufficializzazione di testi legislativi. Al fine di reperire norme d'interesse viene consultata la "Serie generale"¹⁴ con l'elenco delle gazzette pubblicate negli ultimi 30 giorni.

¹⁴ Questa viene pubblicata tutti i giorni feriali.

Nel sito della Gazzetta Ufficiale si trova anche la sezione relativa alla “Unione Europea” (2^a Serie Speciale) che riporta gli atti comunitari come regolamenti, decisioni e direttive.

Questi atti normativi sono reperibili anche dal sito “Eurlex”, sia nella sezione “Gazzetta Ufficiale dell’UE”, sia nella sezione “Diritto dell’UE”-“Cerca negli atti giuridici”.

Qui, attraverso diverse chiavi di ricerca, troviamo atti come i trattati dell’UE, le direttive, i regolamenti, le decisioni e la legislazione consolidata¹⁵.

Nel sito della Gazzetta Ufficiale della Repubblica Italiana si trova, inoltre, la sezione “Regioni” (3^a Serie Speciale)¹⁶; in questa pubblicazione trovano spazio tutti gli atti normativi e amministrativi di interesse nazionale emanati dalle singole Regioni¹⁷, come leggi regionali, decreti presidenziali e regolamenti.

Questi ultimi possono essere reperiti chiaramente anche nel sito della Regione Toscana nella sezione “Leggi, Atti e normative”, in cui, con modalità di ricerca avanzata, è possibile circoscrivere le risorse esistenti per materia o per parole nel testo.

Per avere un quadro completo e al fine di monitorare la produzione normativa nazionale in materia di immigrazione è opportuno consultare anche il sito dell’INPS. Nella sezione “Inps Comunica”- “Atti” sono raccolti tutti gli atti ufficiali relativi alla materia previdenziale e pubblicati nella Gazzetta Ufficiale (Atti ufficiali) nonché quelli emanati direttamente dall’INPS (Circolari, Messaggi e Normativa).

Infine, importanti risorse normative, utili a incrementare il vasto dominio dell’immigrazione, sono sicuramente rappresentate dagli atti emanati dal Ministero dell’Interno.

Nel sito web di questo ministero, infatti, nella sezione “Atti amministrativi generali” l’amministrazione pubblica decreti, direttive, circolari, programmi e ogni altro atto che dispone in generale sull’organizzazione, sulle funzioni, sugli obiettivi, sui procedimenti, o che determina l’interpretazione di norme giuridiche che riguardano o dettano disposizioni per l’applicazione.

In particolare, nella sezione “Circolari”, è possibile visualizzare le circolari di maggior interesse per gli utenti adottate dai dipartimenti del Ministero dell’Interno, dal Gabinetto del ministro e dai commissari governativi e straordinari. Altre circolari del Ministero disponibili online sono consultabili nelle pagine web delle singole direzioni o dei vari dipartimenti.

Il Dipartimento per gli Affari interni e territoriali, ad esempio, svolge funzioni di supporto alle attività di governo locale, di garanzia della regolare costituzione degli organi elettivi, del loro funzionamento e attività di collaborazio-

¹⁵ La consolidazione consiste nel proporre un documento che raggruppi l’atto giuridico di base e tutte le modifiche o rettifiche apportate successivamente.

¹⁶ Questa viene pubblicata il sabato.

¹⁷ Nel nostro caso sono d’interesse quelli emanati da Regione Toscana.

ne con gli enti locali, e vigilanza sullo stato civile e sull'anagrafe. Tra gli Uffici di diretta collaborazione per questo dipartimento, di particolare importanza per il dominio di interesse è la Direzione centrale per i Servizi demografici che svolge compiti di indirizzo, coordinamento e vigilanza in materia di anagrafe e stato civile.

Di interesse anche le circolari emesse dal Dipartimento della Pubblica sicurezza, in quanto titolare di tutte le attività connesse alla gestione dell'ordine e della sicurezza pubblica, al coordinamento tecnico-operativo delle Forze di polizia. Ed infine sono fondamentali quelle emesse dal Dipartimento per le Libertà civili e l'Immigrazione.

Le circolari di questo ufficio, pur rientrando negli atti di interpretazione, istruzione ed organizzazione destinati ai vari uffici, centrali e periferici del Ministero dell'Interno, per una corretta e puntuale applicazione delle norme, si caratterizzano, in gran parte dei casi, per la diretta incidenza sul tessuto socio-economico del Paese. In tal senso sono significative le tante circolari emanate in tema di regolarizzazione dei lavoratori extracomunitari, procedure riconducibili alle competenze dello Sportello Unico per l'Immigrazione e sulle tematiche riguardanti la concessione della cittadinanza. Tra i provvedimenti presenti in questa sezione possiamo trovare anche gli avvisi emanati dalle Autorità Responsabili dei Fondi dell'Unione europea, i bandi di gara, l'approvazione di graduatorie e gli atti di natura organizzativa tipici delle amministrazioni pubbliche.

In occasione del censimento e continuo monitoraggio dei siti web dei Comuni toscani relativamente a informazioni su servizi¹⁸ e procedimenti si è deciso di realizzare e incrementare una raccolta di atti amministrativi e normativi (in formato XML) di competenza comunale in materia di immigrazione.

Le tipologie di atti individuati (accordi, circolari, deliberazioni, determinazioni, protocolli d'intesa) relative a precise tematiche, contribuiscono a realizzare e perfezionare una sezione dedicata in cui mettere in evidenza le informazioni prodotte dagli Enti locali del territorio su servizi e procedimenti dei Comuni che spesso si trovano collocati in posizioni di non immediata visibilità nelle rispettive reti civiche e di facilitarne la fruizione.

3. Descrizione dello strumento

Al fine di supportare l'attività di reperimento dell'informazione rilevante per lo sviluppo e l'aggiornamento del Portale PAeSI, è stato progettato e realizzato un applicativo informatico in grado di effettuare le seguenti azioni:

¹⁸ <https://www2.immigrazione.regione.toscana.it/?q=procedimenticomunali> (ultima consultazione: 30/05/2022).

- monitorare regolarmente alcune fonti online di informazione e recuperare in automatico i testi pubblicati;
- sottoporre all'esperto di dominio le risorse selezionate come rilevanti;
- trarre dalle valutazioni dell'esperto, informazioni utili per correggere, quindi migliorare, le ulteriori ricerche dello strumento.

Le varie azioni saranno descritte nei paragrafi che seguono¹⁹.

3.1. *Monitoraggio e recupero dell'informazione*

Per quanto concerne la prima azione e cioè il monitoraggio delle fonti informative e il recupero delle risorse, lo strumento effettua una ricerca giornaliera (programmata in orario notturno) su tutte le fonti che l'esperto di dominio ha individuato e inserito come da monitorare. Operativamente l'applicativo effettua le stesse azioni che gli operatori del Portale PAeSI compiono con regolarità, ovvero la consultazione di siti istituzionali alla ricerca di norme e informazioni in materia di immigrazione.

La capacità di valutare se questi documenti pubblicati sono interessanti, ovvero trattano temi legati al dominio del Portale, è simulata andando a verificare la presenza di un determinato numero di argomenti citati nei loro testi. In pratica, lo strumento ricerca all'interno di tali testi certe espressioni che l'esperto di dominio ha precedentemente fornito e che, in base alla sua esperienza, sono quelle che più di frequente ricorrono nelle risorse scelte come rilevanti. Il metodo che abbiamo definito, seppur semplice, copre la casistica delle espressioni presenti nei testi di interesse, che è di fatto limitata, visto che andiamo ad analizzare pagine web che utilizzano un linguaggio istituzionale, meno soggetto a variazioni linguistiche rispetto ad altri tipi di testo. La compilazione di tale elenco di espressioni è avvenuta in vari momenti.

All'esperto di dominio sono state inizialmente fornite analisi statistiche sulle parole utilizzate nei testi estratti dalla banca dati normativa nazionale in tema di immigrazione presente sul Portale PAeSI. Tale analisi non ha evidenziato un set di parole che potesse identificare in modo univoco il dominio dell'immigrazione, ma ha permesso di individuare le parole riguardanti i temi trattati dalla normativa attuale e, con una integrazione di termini, ha consentito di stilare un primo elenco di parole da ricercare.

Successivamente è stato utilizzato tale elenco per verificare se con esso lo strumento sarebbe stato in grado di selezionare (e quindi proporre all'operatore per la sua valutazione) tutti i testi della banca dati normativa precedentemente analizzata; l'analisi dei testi non individuati dallo strumento ha

¹⁹ Lo strumento è accessibile, con le apposite credenziali, a questa URL www2.immigrazione.regione.toscana.it/spider (ultima consultazione: 30/05/2022).

permesso all'esperto di dominio di selezionare ulteriori parole significative da inserire nel suddetto elenco.

Infine, è stato utilizzato l'elenco di parole per analizzare le ultime 400 decisioni, 400 direttive e 400 regolamenti pubblicati, in lingua italiana, su EUR-lex, per individuare i testi in cui tali parole erano presenti. Gli esperti del dominio hanno poi valutato la pertinenza dei testi individuati e hanno evidenziato le norme europee in materia di immigrazione che non erano state segnalate dallo strumento. Dopo tale verifica si è potuto incrementare l'elenco delle parole precedentemente individuate per la ricerca automatica delle informazioni d'interesse. Questo processo manuale semi guidato ha permesso di affinare e convergere verso un numero limitato di espressioni che sono state, quindi, generalizzate per coprire le eventuali varianti d'uso, tradotte in espressioni regolari e fornite allo strumento (Fig. 2)²⁰.

The screenshot shows the Spider PAeSI interface. On the left is a sidebar with navigation options: 'Dashboard', 'Ricerca per parole', 'Ricerca per riferimenti', 'REGISTRO EVENTI', 'Errori', 'Segnalazioni', and 'Informazioni'. The main area is titled 'Espressioni da monitorare: 121' and contains a list of regular expressions. Each entry includes the expression, a list of example terms, and a score in a red box.

Espressione	Esempi	Peso
<code>(crisi processo contestofluss.) (im)?migratori(,)?</code>	Esempi individuali - crisi migratoria - flussi migratori - contesto migratori - processo immigratorio -	100
<code>tratta(di degli)* (esseri)* umani</code>	Esempi individuali - tratta (degli esseri umani) -	100
<code>(nullaosta nulla osta)(all per)? (lavoro matrimonio)</code>	Esempi individuali - nulla-osta al lavoro - nullacosta al lavoro -	100
<code>vist.(di per)? (corto lungo)? (soggiorno lavoro)</code>	Esempi individuali - visti per lavoro - visti corto soggiorno -	100
<code>(regolar(, mente) status permesso carta)(di)? soggiorno(olante)</code>	Esempi individuali - carta (di) soggiorno - status di soggiornante - regolarmente soggiornante - permesso (di) soggiorno -	100
<code>xenofobi(c)?alrazzisi(talmo)</code>	Esempi individuali - razzismo - razzista - xenofobia - xenofobia -	100
<code>protezione (internazionale (dei)?minor sussidiaria temporanea)</code>	Esempi individuali - protezione temporanea - protezione sussidiaria - protezione (dei) minori - (status di) protezione internazionale -	100
<code>(discriminazione motivij intolleranza integrazione origine) (raz(z)?ial. etnic.)</code>	Esempi individuali - origine etnica - intolleranza razziale - integrazione etnica - motivi etnici - motivi razziali - motivi etnici - discriminazione (etnico) razziale -	100
<code>lavoro ((in)?nero irregolare)</code>	Esempi individuali - lavoro (in) nero - lavoro irregolare -	55

Figura 2: Elenco delle espressioni regolari fornite allo strumento.

Lo strumento dispone di una ulteriore chiave di ricerca: è in grado di estrarre automaticamente la lista delle norme presenti nella banca dati normativa italiana del Portale PAeSI, e i relativi identificatori URN. Tramite un *parser* linguistico, lo strumento è in grado di ricercare nei testi analizzati gli estremi dei riferimenti ad atti normativi/amministrativi e calcolare i corrispondenti

²⁰ In precedenti progetti di informatica giuridica sono state usate espressioni regolari al fine di identificare in testi giuridici e 'marcare' in linguaggio XML porzioni di testo significative per il riconoscimento di nessi normativi quali riferimenti testuali espliciti e modifiche testuali esplicite. Vedi (Mercatali, Romano, Bolioli 2004).

identificatori URN. Testando questi identificatori è in grado quindi di ricercare nei testi le citazioni ad atti presenti nel database delle norme sull'immigrazione, trattando di fatto questi riferimenti come espressioni di selezione in Rete di testi utili per gli operatori del Portale.

L'esperto di dominio fornisce allo strumento anche una lista di indirizzi di pagine online dove sono pubblicate notizie (pagine principali o sezioni 'notizie' dei portali istituzionali) e norme (banche dati o gazzette online). Lo strumento visita regolarmente tali pagine e scarica tutti i testi pubblicati che non aveva precedentemente scaricato. Tali testi sono analizzati con l'ausilio delle espressioni regolari e la lista dei riferimenti normativi e valutati come rilevanti o non rilevanti in materia di immigrazione in base all'individuazione o meno di parole e riferimenti delle suddette liste²¹.

3.2. *Intervento dell'esperto del dominio*

Lo strumento compie il suo lavoro senza necessità di interventi da parte dell'operatore. Tuttavia, l'operatore ha a disposizione una serie di informazioni dettagliate e aggregate su pagine visitate, documenti scaricati, tempi impiegati, documenti valutati, errori che inevitabilmente possono verificarsi durante l'attività.

È possibile visualizzare un report (Fig. 3) di eventi aggregati occorsi nell'ultima esecuzione dell'attività di ricerca e analisi dello strumento.



Figura. 3: Report eventi ultima esecuzione dello strumento.

Tale report evidenzia il numero documenti (cioè i link) verificati e presenti sui siti monitorati, il numero di documenti nuovi (rispetto alla visita precedente) aggiunti al database interno dello strumento di ricerca, il numero di errori e delle segnalazioni occorse.

²¹ In un progetto precedente è stato sperimentato un approccio basato unicamente sulla ricerca dei riferimenti normativi nazionali ed europei (in materia di disciplina del trattamento dei dati personali) al fine di mantenere aggiornato il DPO (*Data protection officer*) del progetto sul modificarsi della normativa di riferimento. Vedi (Arezzini et al. 2018, 221-32).

Nel rapporto (Fig. 4) si mostrano, invece, le attività aggregate occorse in tutte le esecuzioni dello strumento: tempo impiegato nella selezione dei link dei documenti, tempo impiegato per l'analisi degli stessi, dettagliato in tempo impiegato nel prelevamento dei documenti e analisi dei testi, altro tempo impiegato dallo strumento.

L'utente operatore guardando il report con i valori aggregati, può verificare che lo strumento si sta comportando correttamente, verificando che il numero di documenti, il tipo e i tempi impiegati rimangano in trend noti/limitati.

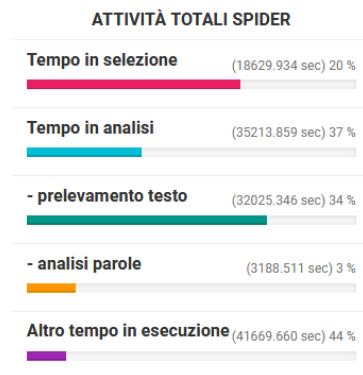


Figura 4: Report attività strumento.

Un numero di segnalazioni particolarmente elevato o la segnalazione di tempi eccessivi di esecuzione, potrebbe essere indagata andando a consultare l'elenco completo delle attività eseguite dallo strumento di ricerca (Fig. 5). Lo strumento classifica tali attività in eventi normali (informazioni), in attività che hanno generato un errore (errori) e altre attività che, pur non essendo errori, potrebbero nascondere una qualche problematica (segnalazioni): ad esempio nel caso in cui lo strumento di ricerca abbia analizzato un testo molto corto.

L'utente può, quindi, approfondire la sua analisi andando a visionare sia l'elenco completo delle attività dello strumento, sia ogni singolo documento scaricato e analizzato, andando a leggere il testo scaricato e poi analizzato e le parole individuate (evidenziate nel testo) che hanno permesso di selezionare il documento come pertinente.



Figura 5: Report log.

Il sistema mostra (Fig. 6) all'esperto di dominio anche l'attività svolta (documenti che lo strumento ha individuato e proposto all'utente e da esso valutati come pertinenti o non pertinenti) e da svolgere (documenti che lo strumento di ricerca ha proposto alla valutazione dell'esperto, ma ancora non valutati da quest'ultimo), tramite un cruscotto informativo.

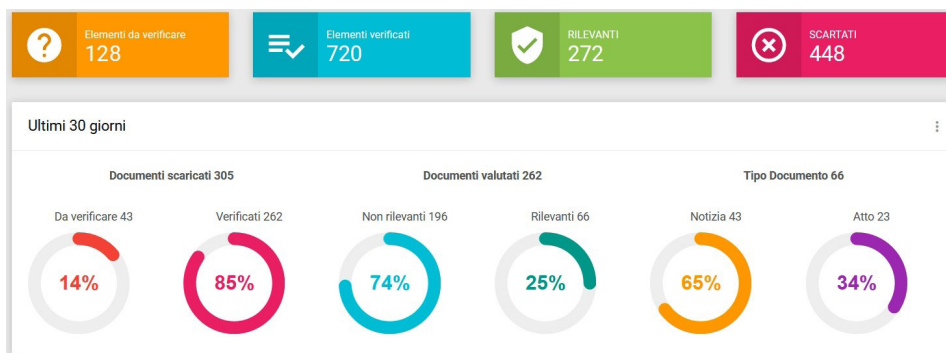


Figura 6: Cruscotto informativo.

Nel secondo riquadro è possibile avere una visualizzazione degli ultimi 30 giorni di attività (sono selezionabili anche gli intervalli di 7 e 15 giorni) in cui sono mostrati anche la ripartizione dei documenti rilevanti in notizie e norme.

Infine (Fig. 7), lo strumento indica all'utente i documenti che ha automaticamente selezionato e ritiene rilevanti come un elenco di documenti caratterizzata da una data (di reperimento del testo online), da un titolo del documento (recuperato dal titolo della pagina o dalla parte testuale del link ipertestuale che ha individuato lo strumento nella sua ricerca), dall'indirizzo della risorsa online (per visionare il documento basta cliccare sul titolo e si apre una nuova finestra all'indirizzo della risorsa) e dalla lista di parole che hanno permesso di selezionare il documento.

<p>Comunicato Stampa</p> <p>https://www.ispettorato.gov.it/it-it/notizie/Pagine/Comunicato-Stampa-20042021.aspx</p> <p>Parole segnalate: caporalato, sfruttamento lavorativo</p>	<p>21/04/2021</p> <p><input type="checkbox"/> Rilevante <input type="checkbox"/> Non rilevante</p>
<p>Lavoratori italiani all'estero: retribuzioni convenzionali 2021</p> <p>http://www.inps.it/nuovoportaleinps/default.aspx?itemdir=55261</p> <p>Parole segnalate: paesi extra UE, lavoratori extracomunitari, lavoro irregolare</p>	<p>21/04/2021</p> <p><input type="checkbox"/> Rilevante <input type="checkbox"/> Non rilevante</p>
<p>Decreto Sostegni: semplificazioni e requisiti per indennità a lavoratori</p> <p>http://www.inps.it/nuovoportaleinps/default.aspx?itemdir=55262</p> <p>Parole segnalate: lavoratori domestici, lavoro irregolare, lavoratori stagionali</p>	<p>21/04/2021</p> <p><input type="checkbox"/> Rilevante <input type="checkbox"/> Non rilevante</p>
<p>Cina: nella biblioteca di Canton nasce la sezione italiana</p> <p>https://www.esteri.it/mae/it/sala_stampa/archivionotizie/retediplomatica/cina-nella-biblioteca-di-canton-nasce-la-sezione-italiana.html</p> <p>Parole segnalate: Console Generale</p>	<p>21/04/2021</p> <p><input type="checkbox"/> Rilevante <input type="checkbox"/> Non rilevante</p>
<p>Accordo Regione-sindacati per i diritti dei rider</p> <p>https://www.toscana-notizie.it/archivio/-/asset_publisher/Lydt2ls2gGDzu/content/id/56724114</p> <p>Parole segnalate: caporalato, sfruttamento lavorativo</p>	<p>21/04/2021</p> <p><input type="checkbox"/> Rilevante <input type="checkbox"/> Non rilevante</p>
<p>MINISTERO DEGLI AFFARI ESTERI E DELLA COOPERAZIONE INTERNAZIONALE - COMUNICATO</p> <p>http://www.gazzettaufficiale.it/eli/id/2021/04/20/21A02340/SG</p> <p>Parole segnalate: Consolato generale</p>	<p>21/04/2021</p> <p><input type="checkbox"/> Rilevante <input type="checkbox"/> Non rilevante</p>
<p>Gazz. uff. dell'UE L 134/2021</p> <p>https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=OJ:L:2021:134:FULL</p> <p>Parole segnalate: cle</p>	<p>21/04/2021</p> <p><input type="checkbox"/> Rilevante <input type="checkbox"/> Non rilevante</p>
<p>ITL Ravenna Forlì Cesena: lavoro nero e omessi riposi</p> <p>https://www.ispettorato.gov.it/it-it/notizie/Pagine/ITL-Forli-Ravenna-Cesena-lavoro-nero-e-omessi-riposi.aspx</p> <p>Parole segnalate: Lavoro nero, cittadini extracomunitari , lavoro nero</p>	<p>20/04/2021</p> <p><input type="checkbox"/> Rilevante <input type="checkbox"/> Non rilevante</p>

Fig. 7: Lista documenti valutati come rilevanti.

All'utente è sufficiente selezionare il riquadro rilevante, o non rilevante, per confermare allo strumento se il documento è effettivamente pertinente.

3.3. Informazioni utili per migliorare le ulteriori ricerche

L'aggiunta di molte fonti informative da monitorare ha, come previsto, portato rumore alla ricerca. Il sistema ha iniziato a proporre, all'esperto di dominio, documenti che risultavano completamente estranei al tema dell'immigrazione.

Per ovviare a questo problema, è stata introdotta una modalità di lavoro in base alla quale l'operatore può agire sul sistema andando a pesare le espressioni di ricerca in modo tale che lo strumento di ricerca non selezioni un documento se trova unicamente o un numero limitato di occorrenze di espressioni a peso ridotto.

L'operazione di pesatura di queste parole può essere fatta *ex post*, trovando materialmente documenti proposti dallo strumento mal valutati, o *ex ante*, cercando espressioni nella banca dati interna dello strumento, che lo strumento di ricerca incrementa con i documenti automaticamente scaricati e manualmente valutati dall'esperto di dominio.

Interrogato con una parola/espressione, lo strumento visualizza i risultati raggruppando i documenti risultanti in: documenti non segnalati, documenti rilevanti, documenti scartati.

Risulta evidente che l'aggiunta di una parola/espressione che seleziona molti documenti, già valutati come da scartare, porterebbe rumore alla ricerca, andrebbe quindi inserita con un peso basso o non inserita.

Differentemente, aggiungere una parola, non presente fra quelle ricercate, che individua un numero consistente di documenti "rilevanti", potrebbe migliorare le prestazioni della ricerca dello strumento.

4. Primi risultati

Guardando i dati degli ultimi 30 giorni, un documento su quattro, tra quelli proposti dal sistema, risulta valutato come rilevante dall'esperto di dominio.

Tra i documenti valutati come rilevanti, circa un terzo riguarda documenti provenienti da siti istituzionali di produzione normativa, mentre i restanti due terzi sono selezionati da siti che pubblicano informazione qualificata (tra questi ultimi circa un terzo riguarda atti normativi e circa due terzi notizie relative al dominio).

Se si valutano le prestazioni totali, avvenute in circa quattro mesi di attività, il rapporto fra documenti valutati come rilevanti, rispetto a quelli segnalati, si attesta sul 33% (2 su 3 non rilevanti).

Inizialmente le percentuali di riconoscimento erano migliori ma il sistema monitorava quasi esclusivamente le gazzette dei siti istituzionali a livello europeo, italiano e toscano e pochi altri siti istituzionali, come quello del Governo italiano o della Regione Toscana.

Attualmente, invece, le fonti monitorate sono oltre 100 e comprendono tutte le risorse informative indicate nel paragrafo 2; il rapporto fra documenti rilevanti e segnalati si attesta sul 25% (3 su 4 non risultano rilevanti).

Ovviamente, monitorando molte più fonti informative il 'rumore', come detto già, è aumentato, ma il sistema continua a funzionare in maniera molto utile a sostegno dell'attività di aggiornamento del Portale.

5. Conclusioni

Come abbiamo sopra riferito, sono già state operate alcune modifiche alla lista delle parole e alle espressioni regolari da ricercare al fine di evitare ricerche troppo "rumorose".

Valutazioni analoghe potrebbero essere fatte sulle ricerche effettuate sulla base delle norme. Attualmente non è stato ritenuto necessario introdurre la pesatura dei riferimenti normativi, ma trovare citato un riferimento normativo, piuttosto che un altro, è un forte indicatore di rilevanza.

Ad esempio, se il testo che lo strumento di ricerca sta analizzando contiene un riferimento al Testo unico sull'immigrazione, è molto probabile che sia

più rilevante rispetto ad un testo che cita la legge di bilancio dello Stato, che conteneva al suo interno un solo comma in materia di immigrazione e che per questo motivo fa parte del nostro database.

Analogamente potrà essere perfezionato il sistema di monitoraggio per alcune fonti informative, andando a differenziare il set di parole da usare nella ricerca o la pesatura delle stesse.

Ad esempio, sul sito delle Prefetture e delle Questure ci sono moltissime notizie, di cronaca giornalistica che, usando un linguaggio molto meno controllato e trattando il tema dell'immigrazione, possono comportare la selezione di notizie che, pur contenendo le parole del nostro set di ricerca, non sono tuttavia pertinenti per l'aggiornamento del Portale PAESI.

Allo stesso modo, la selezione di informazioni da siti del medesimo tipo (ad esempio da siti di ministeri diversi) possono comportare criticità connesse all'uso di una data parola in domini diversi.

È il caso di ad esempio CIE (Centri di Identificazione ed Espulsione) e CPT (Centri di Permanenza Temporanea) che sono utilizzate anche in altri domini con altri significati.

Lo strumento attualmente è già in grado di mostrare all'utente (Fig. 8) valutazioni prettamente statistiche sulle parole individuate e quindi di aggregare, a vario livello, indici sui termini trovati o sui documenti valutati.



Figura 8: Statistiche sulle parole rilevate nei siti analizzati.

Tali informazioni potrebbero essere sufficienti a innescare il processo di ri-valutazione sui pesi dati alle varie parole, sui siti che attualmente sono mo-

nitorati e quindi potrebbero rivelarsi utili per progettare analisi più accurate, differenziando le stesse a seconda del contesto²².

Riferimenti bibliografici

- Arena, Gregorio. 1997. "Introduzione all'amministrazione condivisa". *Studi parlamentari e di politica costituzionale*, nn. 117–118.
- Arezzini, Silvia, Paolo Bosco, Sara Calderoni, Alberto Ciampa, Simone Coscetti, Stefano Cuomo, Luca De Santis et al. 2016. "ARIANNA: un Ambiente di Ricerca Interdisciplinare per l'Analisi di Neuroimmagini nell'autismo", in *Informatica e diritto*, no. 2: 221-32.
- Barker, Iain. 2005. *Information scent helping people find the content they want* https://www.steptwo.com.au/papers/kmc_informationscent/.
- Bertolo, Carla. 2005. *L'interfaccia e il cittadino. Comunicazione pubblica tra tecniche e riflessività*. Milano: Edizioni Angelo Guerini e associati.
- Bonetti, Paolo. 2011. *Diritti fondamentali degli stranieri*. <http://www.adir.unifi.it/odv/documenti/frontiere/bonetti.htm>.
- Caporale, Marina. 2019. "L'accessibilità ai siti web e alle applicazioni mobili delle pubbliche amministrazioni". *Giornale di diritto amministrativo* no. 3: 357-67.
- Cattani, Paola, e Giuseppe Sergio, cur. 2018. *Comunicare cittadinanza nell'era digitale. Saggi sul linguaggio burocratico 2.0*. Milano: Franco Angeli.
- Ciampi, Costantino, Paolo Carlo Guidotti, Mario Ragona, Loriana Serrotti, Fiorenza Socci, Pierluigi Spinosa, Andrea Marchetti et al. 2001. "Progetto NIR – Fase 2 Accesso alle norme in rete. Recupero della normativa pregressa in formato XML e standards del progetto (DTD e URNs)". *Informatica e diritto*, no. 1: 1–266.
- Ciancio, Bruno. 2014. *Sviluppare la competenza interculturale. Il valore della diversità nell'Italia multietnica. Un modello operativo*. Milano: Franco Angeli.
- Ducci, Gea. 2012. *La comunicazione pubblica e la sfida dell'interculturalità. Lo "sguardo" dei mediatori culturali nella Regione Marche*. Milano: Franco Angeli.

²² Lo strumento descritto è stato sviluppato nell'ambito del progetto Tuscany Empowerment Actions for Migrant System (TEAMS), finanziato dal Fondo Asilo, Migrazione e Integrazione (FAMI) 2014–2020 – Obiettivo Specifico 2. Integrazione/Migrazione legale. Obiettivo nazionale (ON) 2. Integrazione. Piani d'intervento regionali per l'integrazione dei cittadini di paesi terzi.

- Fioravanti, Chiara e Mariasole Rinaldi. 2010. “Il sistema informativo PAeSI: un accesso telematico unico a informazioni, norme e procedimenti in materia di immigrazione”. *Informatica e diritto*, no. 1–2: 93–131.
- Francesconi, Enrico. 2012. “Parser e Editor xmLeges”. In *Linee guida per la marcatura dei documenti normativi secondo gli standard*, a cura di Maria Pia Giovannini, Monica Palmirani, Enrico Francesconi. Firenze: European Press Academic Publishing, 2012.
- Francesconi, Enrico, Carlo Marchetti, Remigio Pietramala, and Pierluigi Spinosa. 2010. “URN-based Identification of Legal Acts: The Case of the Italian Senate”. *Informatica e diritto*, no. 1–2: 233–52.
- Galimberti, Umberto. 2020. “Se le nuove tecnologie rendono inutile comunicare”. In *Il primato delle tecnologie. Guida per una nuova iperumanità*, a cura di Carlo Bordoni. Milano: Mimesis.
- Mason, Giacomo, e Piero Zilio. 2019. *La comunicazione digitale per la PA*. Santarcangelo di Romagna: Maggioli.
- Mercatali, Pietro, Francesco Romano, and Andrea Bolioli. 2004. “Formal Models for a Legislative Grammar. Explicit Text Amendment”. In *Knowledge Management in Electronic Government (KMGov2004)*, a cura di Maria Wimmer, 210–27. Berlin: Springer.
- Nichols, Tom. 2017. *La conoscenza e i suoi nemici. L'era dell'incompetenza e i rischi per la democrazia*. Tradotto da Chiara Veltri. Roma: LUISS University Press.
- Sartor, Giovanni, Monica Palmirani, Enrico Francesconi, and Maria Angela Biasiotti, cur. 2011. *Legislative XML for the Semantic Web. Principles, Models, Standards for Document Management*. Berlin: Springer.

