

Il controllo della qualità nell'informazione geografica volontaria: analisi, rappresentazione e proposte per la valutazione

Quality control in Volunteered Geographic Information: analysis, representation and proposals for its assessment

LAURA CRISCUOLO*, GLORIA BORDOGNA***, PAOLA CARRARA*, MONICA PEPE*

Riassunto

Questo lavoro analizza i contributi geografici volontari come particolare sorgente di informazione in applicazioni di carattere scientifico o professionale, approfondendone la principale criticità: la qualità. Dopo aver fornito un quadro introduttivo, si illustrano le componenti di qualità nelle informazioni geografiche volontarie, si descrivono con schemi ed esempi i principali approcci al controllo della qualità, e si introduce una metodologia per stimare l'efficacia delle strategie di controllo qualitativo sui contributi.

Abstract

This work takes into account geographical voluntary contributions as a particular source of information for scientific or professional applications, deepening the main critical issues: quality. After providing an introductory framework we illustrate the components of quality in volunteered geographic information. We then describe the main approaches to quality management with diagrams and examples, and introduce a methodology for estimating the effectiveness of strategies for quality control.

Parole chiave

Volunteered Geographic Information, qualità dell'informazione, valutazione della qualità

Keywords

Volunteered Geographic Information, information quality, quality assessment

* CNR-IREA, via Bassini 15, 20133 Milano, Tel: +39 02 236993456, [criscuolo.l, bordogna.g, carrara.p, pepe.m] @irea.cnr.it

** CNR-IDPA, via Pasubio 5, 24044 Dalmine (BG), Tel: +39 035 6224262, gloria.bordogna@idpa.cnr.it

1. L'informazione geografica volontaria: origini e caratteristiche

Grazie allo sviluppo del *geo-web* e alle sempre più diffuse tecnologie di comunicazione mobili e *smart* (*smartphone*, *tablet*, ecc.), nell'ultimo decennio numerosi utilizzatori dell'informazione geografica hanno esteso il loro ruolo da quello di semplici utenti a quello più attivo di produttori di contenuti. Le informazioni geografiche generate dagli utenti web, in assenza di strutture formali di riferimento, hanno caratteristiche di caoticità: flussi di informazioni nascono e si muovono in diverse direzioni, incrociandosi e scambiandosi contenuti, fino a perdere ogni separazione tra produzione e consumo e a giustificare le ormai diffuse nozioni di *produser* (Bruns, 2006) e di *produsage*. Nonostante la disorganizzazione e l'eterogeneità che caratterizzano queste informazioni, la loro produzione è diventata con gli anni un fenomeno talmente esteso da destare l'interesse della comunità scientifica e dei soggetti storicamente impegnati nella creazione e distribuzione di informazione geografica (Borruso, 2013). La vasta gamma di contributi geografici e cartografici prodotti da utenti non esperti è stata analizzata dai diversi autori che l'hanno identificata con concetti diversi, a seconda dei contesti di riferimento. *Neogeography* (Turner, 2006, Hacklay et al. 2008) è il termine introdotto per enfatizzare la novità che risiede nell'avvento di una geografia "popolare", non esperta, e basata sul web 2.0. Goodchild (2007) ha introdotto il termine *Volunteered Geographic Information* (VGI) per indicare quei contributi a contenuto prevalentemente geografico generati dagli utenti in modo volontario; Criscuolo et al. (2013) hanno poi proposto una classificazione dell'informazione geografica che separa la *Incidental Information* dal VGI di senso classico, nei casi in cui il *produser* contribuisca in modo inconsapevole alla attività. Il termine *Crowdsourcing* è diventato altrettanto comune in ambito *geo-web* per indicare attività di produzione di informazione di tipo collaborativo (Howe, 2006, Heipke, 2010). Recenti sono anche i concetti di *Citizen as a Sensor* (Goodchild, 2007), *Participatory Sensing* e *Urban Sensing* (Campbell et al. 2006, Lane, 2008), riferiti all'uso intensivo di dispositivi mobili interconnessi da parte di cittadini privati, che partecipano così, più o meno

consapevolmente, alla costruzione di un network osservativo globale (De Longueville et al., 2010).

La modalità più frequente in cui avviene l'acquisizione delle informazioni è attraverso dispositivi di localizzazione, la maggior parte delle volte sensori GPS e apparecchi fotografici, integrati in tablet e smartphone.

L'informazione risultante è georiferita, sia essa una immagine, un Point Of Interest o un semplice testo, provvisti o meno di metadati.

Le pratiche di acquisizione e di distribuzione su web si appoggiano spesso ad applicativi estremamente intuitivi che hanno raggiunto un elevatissimo livello di diffusione pubblica. All'interno di questo vasto campionario espressivo, alcuni contributi informativi hanno finalità ludiche o sociali (ad esempio le fotografie amatoriali pubblicate sui *virtual globes*), a volte utilità pratica (come le recensioni di luoghi e servizi) o scopi di promozione (ne sono esempio la distribuzione via web di informazioni professionali, la localizzazione di esercizi commerciali, ecc.); altri sono invece diretti ad arricchire il patrimonio conoscitivo collettivo e a contribuire alla ricerca scientifica (ad esempio segnalazioni di anomalie ambientali, osservazioni biologiche, misurazione di parametri fisici). Questi ultimi contributi sono particolarmente ricercati e apprezzati nei progetti di Citizen Science (Bonney et al., 2009).

La presenza di tali strumenti sul territorio è la caratteristica fondamentale che ne determina la potenza: con una rete fitta e distribuita di sensori interconnessi diventa possibile raccogliere una grande quantità di informazioni, anche dove gli strumenti tradizionali non riescono coprire il territorio in modo capillare. Ciò rende possibile ridurre le zone d'ombra, raccogliere e valorizzare conoscenze strettamente connesse al luogo, migliorare la frequenza di aggiornamento attraverso una presenza in situ più costante (Rocca, 2013; Giannola, 2013).

Il coinvolgimento attivo di "cittadini-sensori" – insieme produttori e fruitori dell'informazione territoriale – in campagne di raccolta dati, apre prospettive entusiasmanti per un ampio numero di soggetti: enti pubblici, istituti di ricerca, organizzazioni benefiche, associazioni civiche, privati, commercianti ed agenzie educative, che possono beneficiarne arricchendo data set geografici, documenti cartografici o raccolte di vario genere. Questo tipo di apporto non esperto è solitamente ambito

non solo per i benefici che determina in termini di numerosità dei contributi, ma anche perché riesce a raccogliere informazioni di rilievo legate a competenze locali, detenute soltanto da chi vive in prima persona i luoghi e i fenomeni oggetto di analisi. L'approccio collaborativo alle attività geografiche e cartografiche ha infine ricadute positive sul coinvolgimento degli utenti e delle parti interessate, che diventano più attive e consapevoli.

1.1 Principali criticità nell'utilizzo del VGI a scopi professionali

Accanto all'entusiasmo che accompagna lo sviluppo di questo nuovo approccio alla geografia, è nata e cresciuta di pari passo anche la preoccupazione per le conseguenze che le nuove pratiche amatoriali possono portare alle discipline geografiche e cartografiche consolidate. Numerosi sono, infatti, i problemi correlati alla creazione e all'utilizzo di informazioni geografiche di origine volontaria.

Un primo problema riguarda la metadattazione dei dati di provenienza volontaria non specializzata, cioè la descrizione delle loro caratteristiche ancillari esogene ed endogene. Infatti, il grande volume di dati raccolti risulta spesso privo, del tutto o in parte, di quelle meta-informazioni che permettono di localizzarli con sicurezza nello spazio e nel tempo e di definirne parametri fondamentali per l'utilizzo: sistema di riferimento, procedura di acquisizione, accuratezza della misura, precisione strumentale, data e ora dell'acquisizione, recapiti dell'autore, ecc. La causa di questa mancanza è da ricercare nell'inesperienza degli operatori volontari, che, pur mossi da buone intenzioni, spesso non sono consapevoli del valore di queste informazioni di complemento e le omettono. Altre volte sono gli stessi strumenti con cui avviene l'acquisizione dell'osservazione a non facilitare la compilazione di metadati particolarmente elaborati. Altre volte ancora le informazioni utilizzate per progetti scientifici o di pubblica utilità vengono selezionate all'interno di volumi di dati sì volontari, ma originariamente prodotti per altri scopi, e perciò caratterizzati diversamente (ad es. nella ricostruzione della situazione precedente a un disastro ambientale possono risultare utili fotografie panoramiche, scattate senza intenti scientifici dai cittadini, che acquisiscono rilevanza soltanto in una fase successiva allo scatto). In questi casi

risulta ancora più complesso riuscire a risalire alle informazioni di accompagnamento necessarie per un uso consapevole del dato. Esse infatti risultano spesso scarse o presenti in forme difficili da trattare (ad esempio hanno la forma di link ad altri contenuti web, *nickname*, *tag* e *geotag*, ecc.), e talvolta il gioco di rimandi e collegamenti tra le diverse pagine web risulta talmente complesso da farne perdere le tracce o elevarne il livello di incertezza.

Metadati particolari, ma di importanza rilevante, sono poi quelli che attengono alla protezione dei dati personali, la privacy, i copyright e i diritti intellettuali. Sebbene non direttamente connesse al contenuto informativo, la carenza o l'ambiguità di queste indicazioni pregiudica l'utilizzo e la disseminazione di una osservazione volontaria.

Un secondo punto critico, in parte conseguenza del precedente, riguarda la difficoltà nell'elaborare le informazioni volontarie. Può accadere infatti di riuscire ad accumulare volumi considerevoli di informazioni, ma di trovare serie difficoltà nel sovrapporre e correlare gli strati informativi. Le operazioni di sovrapposizione, intersezione, combinazione (indicate spesso in ambiente *geo-web* con i termini *mash-up* e *overlay*) sono rese difficili dalla variabilità degli attributi dei dati. Diverse precisioni strumentali, diversi sistemi di riferimento utilizzati, diverse scale geografiche, intervalli temporali non definiti, assenza di standard e schemi concettuali condivisi fanno sì che l'analisi spaziale divenga difficile e talvolta addirittura avventata. In questo senso la numerosità dei dati volontari non è sempre garanzia di efficacia nella elaborazione e dunque nella produzione di nuova informazione.

Esiste poi una terza grande questione, legata all'attendibilità dei dati volontari. La qualità di un contributo di provenienza volontaria non è infatti solo una caratteristica intrinseca del dato, connessa a fattori quali derivazione, accuratezza, consistenza logica e completezza (Moellering, 1987), ma ha anche un aspetto estrinseco, legato alla affidabilità e alla esperienza dell'autore (Hovland et al., 1953): solo la combinazione di questi due aspetti infatti può dare attendibilità ad una informazione ("*credibility*" secondo Flanagan e Metzger, 2008). Questo concetto base di attendibilità si adatta sia alla tradizionale produzione di informazioni scientifiche

che “esperte”, sia ai contributi dei *producers* volontari, ma nel secondo caso il concetto risulta ulteriormente complicato da alcuni aspetti caratteristici dell’ambiente web, come la difficile reperibilità degli autori, la mancanza di standard e di selezioni meritocratiche, l’onerosità della ricerca delle fonti. Molti lavori sono stati condotti nell’ultimo decennio per costruire modelli di attendibilità (Metzger, 2007; Kessler and de Groot, 2013) e per analizzare i flussi di *user generated content* sul web in termini qualitativi e quantitativi, discutendone caratteristiche intrinseche, accessibilità, sorgenti, contenuti, motivazioni (per esempio Eysenbach e Kohler, 2002; Coleman et al., 2009; Van Dijck, 2009), e la materia è tuttora argomento caldo di discussione. Accade con ricorrenza che i contribuenti volontari commettano errori non evidenti e che sottomettano informazioni che non sono accertabili se non attraverso un diretto controllo sul campo (è il caso, ad esempio, di una segnalazione di incendio in una certa località). Nel caso più frequente, in cui non è possibile eseguire verifiche dirette, è inevitabile affidarsi alla parola del volontario, e confidare che il suo apporto sia corrispondente a verità. Un elevato grado di incertezza si raggiunge quando i componenti del gruppo di volontari che partecipano all’iniziativa non sono selezionati, o formati preventivamente, o non sono rintracciabili, e quando non sono funzionanti sistemi di filtro o di confronto con dati autorevoli sulle informazioni in ingresso. In casi estremi alcuni volontari in malafede possono addirittura inserire premeditadamente indicazioni false, allo scopo di aumentare la confusione, screditare il progetto o mistificare l’informazione complessiva a proprio vantaggio.

L’incertezza sull’attendibilità dei contributi crea dunque un danno a tutto il sistema, compromettendo la fiducia degli utilizzatori.

Altre criticità riguardano infine l’assetto delle organizzazioni – pubbliche e private – che tradizionalmente si occupano di produrre e distribuire informazione geografica a livello professionale. Tali agenzie sono state investite dall’esplosione di queste pratiche geografiche “collaterali” ed hanno dovuto fronteggiare una virata dell’interesse del pubblico. Per questi soggetti diventa oggi una sfida imprescindibile riuscire a coniugare prodotti geografici tradizionali, di elevata qualità, con meccanismi e contributi di tipo partecipativo. Ne costituisco-

no esempi le scelte di aziende produttrici di navigatori satellitari, come Tom Tom o Garmin, o di cartografia, come Compass, che oggi affiancano prodotti cartografici professionali con segnalazioni ed aggiornamenti di provenienza volontaria, soddisfacendo così sia i requisiti di rigore, sia l’interesse degli utenti a partecipare e a mantenere aggiornata la base di dati. Altri esempi in questo campo sono forniti dalle amministrazioni pubbliche che mettono a disposizione dei cittadini le proprie risorse cartografiche e promuovono applicazioni di pubblica utilità con contributi informativi volontari (ad es. la mappa partecipativa della mobilità urbana¹ del Comune di Castelfiorentino, o lo strumento di segnalazione del degrado urbano² del Comune di Messina). Il rischio che invece corrono numerosi altri soggetti, rimasti esclusi dalla nuova tendenza, è quello di perdere seguito, oltre che voluminose quote di mercato, oppure quello di allontanarsi dal ruolo di autorità del settore (Borruso, 2010).

Le prime tre problematiche illustrate sono strettamente legate a ciò che genericamente identifichiamo come “qualità dell’informazione volontaria”. Diversi studi sono stati fatti sui meccanismi con cui i coordinatori delle attività – commerciali, amministrative o scientifiche – possono intervenire per aumentare la qualità del lavoro volontario, senza tuttavia pervenire a soluzioni uniformi e definitive.

Nei capitoli seguenti ci si propone di affrontare la questione della gestione della qualità nel VGI attraverso una descrizione delle sue componenti e delle possibili strategie di controllo, nonché di introdurre una metodologia che ne misuri la riuscita in modo quantitativo.

Le domande più delicate a cui si cerca di dare risposta sono sostanzialmente le seguenti:

- In base a quali criteri giudicare la bontà di un contributo informativo volontario?
- Quali azioni possono favorire la produzione/selezione di contributi accurati?
- Come valutare l’efficacia e la stabilità delle strategie adottate?

1 “Libero accesso...accesso libera tutti!”, <http://www.comune.castelfiorentino.fi.it/castelfiorentino/comune.jsp?IdDoc=82&IsEle=0&Where=IdDoc%3D184>, ultimo accesso 19/11/2014

2 “Segnala il degrado”, <https://crowdmap.com/map/comune-messina/>, ultimo accesso 19/11/2014

A questo fine il secondo capitolo fornisce una descrizione e una organizzazione in categorie per l'informazione geografica volontaria, la sua qualità e le relative strategie di gestione. Le strategie vengono discusse singolarmente ed analizzate in alcune loro combinazioni tramite esempi concreti. Nel capitolo 3 viene illustrato un approccio per stimare quantitativamente il livello di qualità raggiunto, che può applicarsi in maniera flessibile ai diversi casi pratici. In conclusione, nel capitolo 4, viene data una visione generale del fenomeno, accompagnata da spunti di riflessione e suggerimenti operativi.

2. Rappresentazione e gestione della qualità del VGI

Per giungere ad argomentare i metodi di controllo della qualità nei contributi geografici di provenienza volontaria, è utile dotarsi in partenza di un paradigma sintetico con cui organizzare la grande varietà dell'informazione trattata.

2.1 Elementi di informazione geografica volontaria

Le informazioni geografiche più frequentemente fornite dai volontari su web hanno la forma di commenti testuali riferiti a luoghi fisici (come eventi, segnalazioni, allarmi), fotografie geo-taggate, marcatori puntuali corrispondenti a luoghi di interesse privato, turistico o commerciale. Con sufficiente ricorrenza si osservano anche oggetti geografici lineari o poligonali (ad esempio strade ed itinerari), tracciati GPS, misurazioni effettuate con sensori diversi dal GPS o con apparecchiature particolari.

Nel presente lavoro raggruppiamo tali tipi di informazione geografica multimediale nelle categorie:

- **IMMAGINI:** fotografie, videoriprese e oggetti grafici;
- **ANNOTAZIONI:** segnalazioni prevalentemente testuali;
- **FEATURES:** entità spaziali, mono o pluridimensionali, con attributi associati (per esempio *shapefile* o file GML);
- **MISURE:** valori derivati da osservazioni, prevalentemente numerici.

Si escludono invece volutamente i contributi volontari espressi da indici di gradimento (*rating*), ovvero le va-

lutazioni da parte del pubblico nei confronti dei contenuti geografici volontari (ad esempio con metodi di *thumbs up/down* o *star ratings*). Tali espressioni hanno certamente valore informativo e sono pure ampiamente diffuse, ma sono più assimilabili a strumenti di valutazione della qualità delle informazioni che ad informazioni geografiche a sé stanti. Ai fini di questa trattazione i contributi in forma di *rating* saranno perciò inclusi nei meccanismi di controllo della qualità, e non nelle tipologie di VGI.

Ogni elemento di VGI, cioè ogni contributo informativo immesso da un utente volontario, sarà costituito da uno o più componenti, che potranno appartenere alla medesima tipologia (ad es. misure di diversi parametri fisici da un' unica stazione di misura) o essere di tipologie diverse (ad es. una fotografia con una descrizione testuale annessa).

2.2 Rappresentazione della qualità nel VGI e delle sue proprietà

Una volta descritta la forma di un elemento di VGI, è necessario rappresentarne la qualità.

La trattazione della qualità nell'informazione geografica ha una lunga storia, che parte dallo scorso secolo, si intensifica con l'avvento delle tecnologie GIS (una ampia rassegna si trova in Van Oort, 2006) e trova nuova fioritura nell'ultimo decennio, con l'avvento del *geo-web* ed il proliferare di applicazioni di *mapping* collaborativo. Infatti, benché la qualità dell'informazione geografica sia stata ampiamente discussa e trovi degli standard di riferimento nelle norme ISO 19113:2002 e ISO 19114:2003, la qualità dell'informazione geografica volontaria ha infatti caratteristiche tali da necessitare di nuovi indicatori per essere descritta e valutata in maniera soddisfacente (Van Exel et al., 2010). La qualità del VGI è in effetti una proprietà composita, che comprende non solo aspetti legati alle caratteristiche del dato, ma anche alle caratteristiche del produttore del dato e del contesto di applicazione.

Per tenere conto di questa complessità, scegliamo di descrivere la qualità del VGI attraverso tre principali categorie, prendendo spunto dal modello di Bordogna et al. (2014):

- **QUALITÀ INTRINSECA,** dipendente dalle caratteristiche del contenuto informativo;

- QUALITÀ ESTRINSECA, dipendente dalle caratteristiche del contesto, e quindi intesa come giudizio di credibilità sull'informazione e sull'autore;
- QUALITÀ PRAGMATICA, intesa come capacità di soddisfare le necessità di un utente o di un utilizzo (English, 1999).

- INTELLIGIBILITÀ, ovvero possibilità del contributo di essere compreso ed esaminato.

Sono invece proprietà relazionabili alla qualità estrinseca:

- VERIDICITÀ dell'informazione,
- ATTENDIBILITÀ dell'autore dell'informazione.

Le caratteristiche che contribuiscono a determinare la qualità intrinseca di una informazione volontaria possono essere a loro volta scomposte in proprietà elementari, quali:

- ACCURATEZZA dell'osservazione, ovvero conformità al valore reale o atteso;
- PRECISIONE dell'osservazione, ovvero ripetibilità della misurazione;
- CORRETTEZZA del contributo, ovvero assenza di errori formali;
- COMPLETEZZA, cioè assenza di omissioni significative;

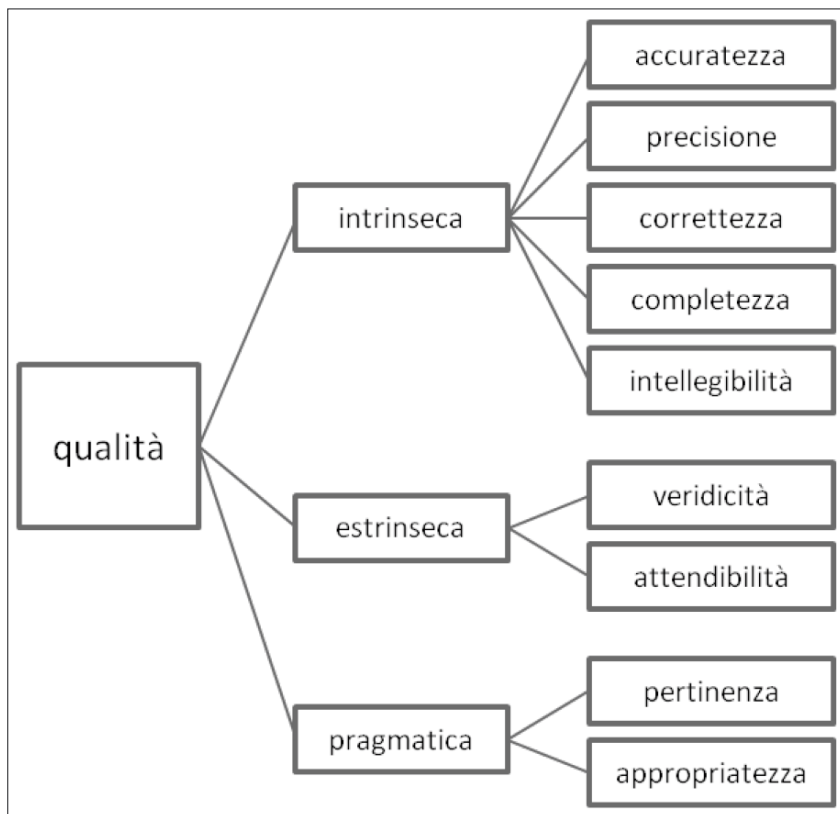
Infine la qualità pragmatica può essere descritta da:

- PERTINENZA dell'informazione,
- APPROPRIATEZZA per un dato utilizzo.

Il modello descritto è schematizzato in Figura 1.

La rappresentazione della qualità in Figura 1 è indipendente dalla forma assunta dal contributo volontario (immagini, annotazioni, misure, features), dal suo contenuto e dal suo contesto, e può pertanto essere assunta come base per valutare – ed eventualmente comparare – la qualità in qualsivoglia progetto di VGI.

FIGURA 1 – Scomposizione della qualità dell'informazione geografica volontaria in proprietà elementari, secondo il modello descritto (Elaborazione degli Autori)



2.3 Approcci al controllo della qualità

Una volta definiti gli elementi di VGI e descritte le proprietà di qualità, si raggruppano i tipi di approccio al controllo della qualità.

Dal punto di vista temporale l'approccio al controllo qualitativo potrà avvenire in via:

- preventiva, se ha luogo attraverso procedure precedenti o contestuali all'immissione di informazioni volontarie (ad esempio con schede di input che guidano il volontario, proposte di suggerimenti e normalizzazione di termini);
- correttiva, se si verifica successivamente all'immissione di informazioni nel sistema (ad es. con selezione dei contributi, rettifica automatica o manuale di alcune informazioni).

Dal punto di vista dei soggetti coinvolti, le operazioni di valutazione avvengono ad opera di:

- team di amministrazione, quando sono condotte manualmente dai coordinatori del progetto, dallo staff tecnico o da gruppi di esperti incaricati;
- comunità dei partecipanti, quando è il gruppo di volontari stesso a valutare ed eventualmente validare le informazioni immesse;
- automatico, quando uno o più componenti informatici dell'infrastruttura operano selezione dei contenuti o apportano modifiche automatizzate.

Infine, dal punto di vista dell'azione prodotta, l'informazione ritenuta non idonea potrà subire:

- segnalazione, e quindi essere pubblicata con accluso un giudizio o un avvertimento;
- rimozione, ovvero essere esclusa dalla pubblicazione e dalle successive elaborazioni.

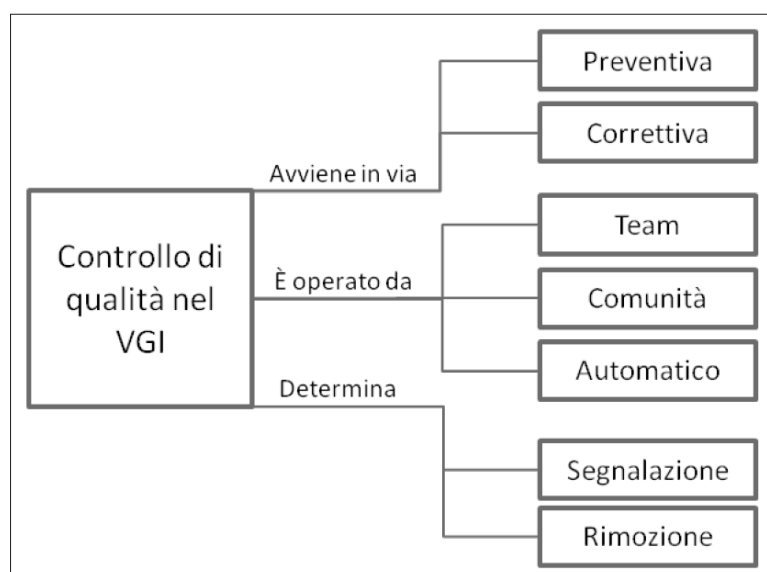
Lo schema così costruito è rappresentato in Figura 2.

Per ogni utilizzo ed ogni contesto applicativo deve essere attentamente valutata la strategia migliore, che spesso comprende l'uso di metodi ibridi. Bilanciando le potenzialità e i deficit delle diverse opzioni, si può giungere ad una gestione adeguata dei problemi di qualità.

Discutiamo di seguito vantaggi e svantaggi collegati a ciascuna opzione, presentando alcune possibili implementazioni.

I meccanismi preventivi, mirati a facilitare una corretta compilazione dell'informazione geografica a priori, possono consistere, ad esempio, in semplici manuali e guide alla collezione di osservazioni, oppure in procedure di compilazione guidata, con campi a scelta multipla o liste di termini vincolati, strumenti automatici di normalizzazione dei termini, strumenti di estrazione automatica di metadati, ontologie e *gazetteer* geografici (Popescu et al., 2009; Kuhn, 2001), o ancora nella selezione e addestramento dei contributori volontari (Galloway et al., 2006; Crall et al., 2011; Ostermann e

FIGURA 2 – Schematizzazione delle tipologie di approccio al controllo della qualità nel VGI (Elaborazione degli Autori)



Spinsanti, 2011). Tutte queste modalità facilitano sicuramente l'inserimento di dati volontari in una forma omogenea e formalmente corretta, difficilmente tuttavia riescono ad assolvere funzioni di controllo sulla veridicità dell'informazione immessa o sull'utilità della stessa all'interno del contesto in cui è stata inserita (elementi della qualità estrinseca e pragmatica).

I metodi a posteriori, di tipo correttivo, agiscono invece modificando le componenti difettose o effettuando una cernita tra i dati in ingresso con criteri qualitativi o quantitativi. Possono prevedere l'uso di algoritmi automatici, o filtri di tipo geostatistico (De Tré et al., 2010; Latonero et Shklovski, 2010), ma anche di supervisori umani per individuare errori sistematici e mantenere la consistenza dei data set (Dickinson, 2010; Huang et al., 2010), o ancora possono monitorare in tempo reale l'integrità semantica dei dati raccolti (Pundt, 2002). I metodi correttivi sono adatti ad agire sull'attendibilità e sull'efficacia dell'informazione volontaria, oltre che sulle caratteristiche intrinseche di qualità della stessa, ma poiché agiscono rimuovendo, fondendo o rimodellando contributi inappropriati, possono causare la perdita parziale o addirittura totale di informazione.

Le valutazioni di qualità effettuate dai responsabili, o gruppo di esperti, offrono una certa tutela, presupponendo in loro doti di competenza, raziocinio e imparzialità. Eppure, persino scienziati o tecnici non possono sempre godere di una padronanza completa delle variabili qualitative, e il loro giudizio non potrà essere che soggettivo e approssimato. Può capitare a volte che privati cittadini con conoscenze locali, professionisti specializzati in particolari attività, od osservatori diretti dei fenomeni possano dare giudizi più circostanziati ed attendibili di quanto siano in grado di fare gli stessi supervisori. In alcune occasioni è tuttavia rischioso assegnare a contributori volontari compiti di valutazione; essi infatti, per impreparazione, superficialità o malafede, potrebbero creare confusione ed arrecare danni all'intera collezione di dati. La combinazione dei due metodi – quello più tradizionale di tipo autoritario, o *top-down*, e quello democratico, o *bottom-up*, caratteristico delle applicazioni on-line di tipo non esperto, non solo è possibile, ma può produrre anche risultati apprezzabili. In questo contesto infatti il web può essere sfruttato come luogo di incontro e di valutazione collettiva:

l'accesso continuativo ad un prodotto web da parte di una squadra ibrida di esperti, locali, amatori, visitatori occasionali, a cui venga data facoltà di valutazione sui contenuti, può dar luogo a una sorta di valutazione di credibilità con metodi partecipativi, con un alto potenziale di selezione e giudizio (Flanagin e Metzger, 2008; Connors et al., 2012).

I meccanismi di controllo automatici, infine, possono rivelarsi estremamente utili nei casi in cui si voglia popolare un database o in cui la mole di dati sia tale da non poter essere controllata manualmente. Allo stato attuale, tuttavia, i sistemi di controllo automatici raramente raggiungono i livelli di affidabilità di un validatore umano per quanto attiene alla rilevanza del contenuto, allo stile, o alla pertinenza.

I controlli che agiscono per rimozione, come filtri, sono di grande aiuto per la conservazione dell'integrità e della consistenza di collezioni di dati volontari, ma, bloccando le informazioni ritenute inadeguate in base a parametri di qualità pre-impostati, comportano l'esclusione o la perdita parziale di informazioni che, per quanto difettose, potrebbero rivelarsi utili in altri contesti. Altri procedimenti, che prevedono di agire sulle informazioni non conformi mantenendole, pur segnalandone il difetto, non perdono alcuna informazione immessa e spingono gli utenti ad utilizzare in maniera consapevole i dati affetti da incorrettezze. Essi pure, tuttavia, hanno due risvolti sfavorevoli: dal punto di vista dell'ottimizzazione del processo agiscono in modo non efficace, destinando una parte della memoria a dati di dubbia rilevanza; dal punto di vista dell'utilizzo, offrendo un set di dati non filtrati, lasciano all'utenza l'onere di decidere sull'affidabilità dei dati che hanno subito segnalazioni.

2.4 Esempi di strategie di controllo della qualità

Analizziamo ora alcuni esempi pratici atti a comprendere come questi approcci siano applicati in maniera composita in alcuni casi concreti di successo e come ne influenzino l'esito.

“Wikimapia”³ è un ampio progetto di cartografia partecipativa che invita utenti volontari ad arricchire

³ “About Wikimapia”, http://wikimapia.org/docs/About_Wikimapia, ultimo aggiornamento 24-10-2013, ultimo accesso 19/11/2014.

una mappa base con informazioni geografiche volontarie: immagini, annotazioni, entità geografiche. La politica attuata è di tipo democratico: l'intera comunità di utenti, senza alcuna selezione, ha la possibilità di creare, modificare e rimuovere i contributi apportati. Alcune precauzioni sono tuttavia state introdotte dai coordinatori del progetto per proteggerne la correttezza e tenere il livello qualitativo il più elevato possibile. Questi meccanismi sono sia preventivi (come ad esempio i manuali d'uso, i dizionari controllati, la consulenza online, la semiprotezione di entità geografiche sensibili, la concessione di permessi aggiuntivi ad utenti esperti, ecc.) sia correttivi (ad esempio meccanismi di rimozione di contenuti offensivi, o esclusione di utenti scorretti). La comunità di utenti-contributori opera correzioni a posteriori sui contenuti, mentre il team di amministrazione coordina i meccanismi preventivi, l'esclusione di utenze, e decide sulla eventuale rimozione di contenuti inappropriati.

Le modifiche apportate ai contributi volontari possono determinare segnalazioni o rimozioni, che tuttavia non sono definitive: viene infatti conservata memoria delle modifiche apportate nel tempo, in modo che l'evoluzione di ciascun elemento mappato possa essere consultata in maniera completa e che ogni sua versione possa essere eventualmente ripristinata.

“Hai sentito il terremoto?”⁴ È un'applicazione creata dall'Istituto Nazionale di Geofisica e Vulcanologia (INGV) per raccogliere informazioni volontarie sulla percezione umana dei terremoti, ed ha la sua analoga versione in diverse nazioni (ad esempio lo statunitense “Did You Feel It?”, curato da U.S. Geological Survey⁵). Le informazioni sono raccolte in forma di annotazioni, riferite ad una località sul territorio italiano. La compilazione dei contributi avviene tramite *web form*, con campi a scelta multipla, o variamente vincolati. I dati in ingresso vengono sottoposti ad un filtro automatico di tipo statistico, ma non sono validati singolarmente prima della pubblicazione. Chiunque può partecipare, ma nessun utente o amministratore ha facoltà di agire

sui contenuti, né con segnalazioni né con correzioni. Saranno poi attività complementari in sito, non direttamente collegate all'applicazione web, a verificare l'effettiva corrispondenza delle segnalazioni più rilevanti con reali eventi sismici.

“eBird”⁶ è il più ampio progetto di raccolta di osservazioni volontarie di avifauna a scala mondiale. È supportato da una vasta comunità di scienziati e amatori, che arricchiscono con le loro osservazioni (annotazioni ed immagini) un database; le procedure di analisi associate vanno così ad aggiornare mappe di frequenza, statistiche e grafici. L'intento scientifico del progetto, l'imponente mole di dati, e la pesante infrastruttura che li elabora hanno imposto una politica severa di controllo della qualità. La precisione e la veridicità delle informazioni in questo caso è affidata non solo a meccanismi di controllo a priori, quali manuali e a protocolli web guidati per la compilazione (tassonomie, dizionari controllati, codici, *check lists*,...), ma anche a procedure correttive a posteriori, che agiscono tramite rilevamento automatico delle imprecisioni e filtraggio. I dati in ingresso vengono visionati da algoritmi automatici e, se da essi segnalati come anomali sulla base di criteri statistici, vengono sottoposti al controllo di una équipe di esperti, che ne determina la eventuale revisione o esclusione. In questo modo il database risulta popolato solo dai contributi che hanno superato i controlli di qualità.

“Tomnod”⁷ è un'applicazione web di *crowdmapping* applicato alla segnalazione emergenze. Ai volontari vengono sottoposte immagini satellitari dell'area in cui è in corso una emergenza o in cui si è di recente verificato un evento calamitoso: ad essi viene chiesto di visionare sezione per sezione la mappa e di segnalare i punti in cui ritengono di riconoscere oggetti significativi, appartenenti a categorie precisate (edifici danneggiati, valanghe, persone disperse, strade interrotte...). I contributi rientrano nella tipologia delle *features* puntuali e che vengono archiviati in base alle etichette a loro associate (*tag*). Essi vengono analizzati in maniera automatica da un algoritmo geostatistico, che li con-

4 “Hai sentito il terremoto?”, <http://www.haisentitoilterremoto.it>, ultimo accesso 19/11/2014.

5 “Did You Feel It?”, <http://earthquake.usgs.gov/earthquakes/dyfi/> ultimo aggiornamento 09/09/2014, ultimo accesso 19/11/2014.

6 “About eBird”, <http://help.ebird.org/customer/portal/articles/1055676-understanding-the-ebird-review-and-data-quality-process>, ultimo accesso 19/11/2014.

7 “Tomnod”, <http://www.tomnod.com/>, ultimo accesso 13/01/2014

fronta con quelli immessi dalla comunità di utenti e ne determina, in base alla ricorrenza e alla corrispondenza geografica, il grado di verosimiglianza. Questo controllo particolare è quindi operato attraverso una funzione automatica, ma è implicitamente sostenuto dalla comunità. L'affidabilità dei contributi di ciascun contributore volontario viene poi sommata e determina a sua volta il grado di attendibilità dell'autore. I volontari vengono così coinvolti in un meccanismo che richiama quelli tipici dei *social games*, e che li porta, attraverso il confronto con la comunità, ad accrescere il proprio status di esperienza, aumentare il livello di consapevolezza e allo stesso tempo ne incentiva la partecipazione. I contributi sono dunque sottoposti ad un controllo di qualità a posteriori di tipo iterativo, che si rinnova ad ogni nuovo inserimento di dati. Al fianco di questo meccanismo fondamentale, vengono attuate anche due tipiche strategie preventive: il vincolo sulle tipologie di oggetti da individuare su mappa (una categorizzazione dei *tag*) ed un rapido addestramento dei volontari, attraverso gallerie fotografiche di esempio.

Usando le categorie precedentemente introdotte per la rappresentazione dei meccanismi di controllo della qualità possiamo descrivere in maniera sintetica l'approccio al controllo qualitativo dei quattro progetti con il modello in Figura 3.

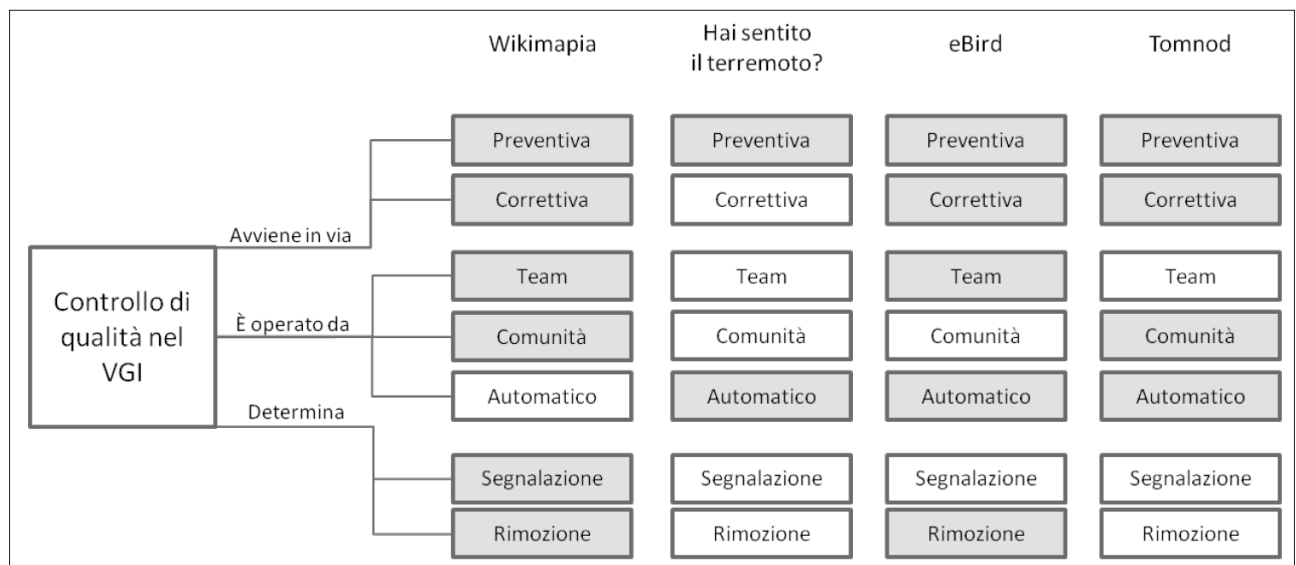
La Figura 3 mette in evidenza come possano essere scelte ed implementate configurazioni completamente diverse del controllo della qualità, sulla base delle esigenze e delle politiche legate alle diverse attività specifiche.

3. Un approccio flessibile per la stima della qualità

Quali che siano le strategie scelte per fronteggiare i problemi di qualità delle informazioni volontarie, a valle di queste è utile adottare un meccanismo in grado di valutare il livello di efficacia raggiunto, di stabilire cioè che grado di qualità viene raggiunto dai singoli contributi e dal set di dati in generale, ma anche di monitorarne l'andamento nel tempo, legato a vari fattori, tra cui l'evoluzione della comunità di contributori.

In tempi recenti sono stati fatti diversi sforzi per elaborare procedure di valutazione, concentrandosi talvolta su aspetti legati all'attendibilità (Metzger, 2007), talvolta sull'accuratezza geografica (Kessler e de Groot, 2013; Sabone, 2009; Goodchild, 2008), spesso calcolata attraverso confronti tra data set o campagne in sito (Hacklay, 2010). Tali proposte, pur efficaci nel valutare particolari aspetti di qualità, utili nel loro contesto specifico, non offrono un metodo sufficientemente gene-

FIGURA 3 – Rappresentazione sintetica delle strategie di controllo della qualità attuate in quattro applicazioni che raccolgono informazione geografica volontaria (Elaborazione degli Autori)



rale, flessibile né comprensivo dei diversi aspetti che caratterizzano la qualità dell'informazione geografica volontaria.

Per superare questi limiti, scegliamo di basarci sulla rappresentazione della qualità introdotta nel paragrafo 2.2. Appoggiandosi ad essa è possibile definire degli indici elementari di qualità, associati a ciascun componente di VGI, e successivamente aggregarli in indici composti, fino a pervenire alla ricostruzione di un indice complessivo di qualità dell'informazione, e, volendo, all'indice di qualità di un intero set di dati. Un metodo simile è stato descritto in Bordogna et al. (2014), e viene qui proposto in una forma più generalizzata, in modo da renderlo facilmente applicabile e trasferibile ad ogni applicazione di VGI.

Scomponiamo innanzitutto un contributo informativo volontario nelle sue componenti atomiche, che possono essere costituite da una o più immagini, annotazioni, misure, features geografiche. L'informazione complessiva di un contributo, che indichiamo con VGI_{TOT} , sarà dunque data dalla aggregazione delle n componenti informative VGI_i

$$VGI_{TOT} = \oplus (VGI_1, VGI_2, VGI_3, \dots, VGI_n)$$

con \oplus operatore di aggregazione

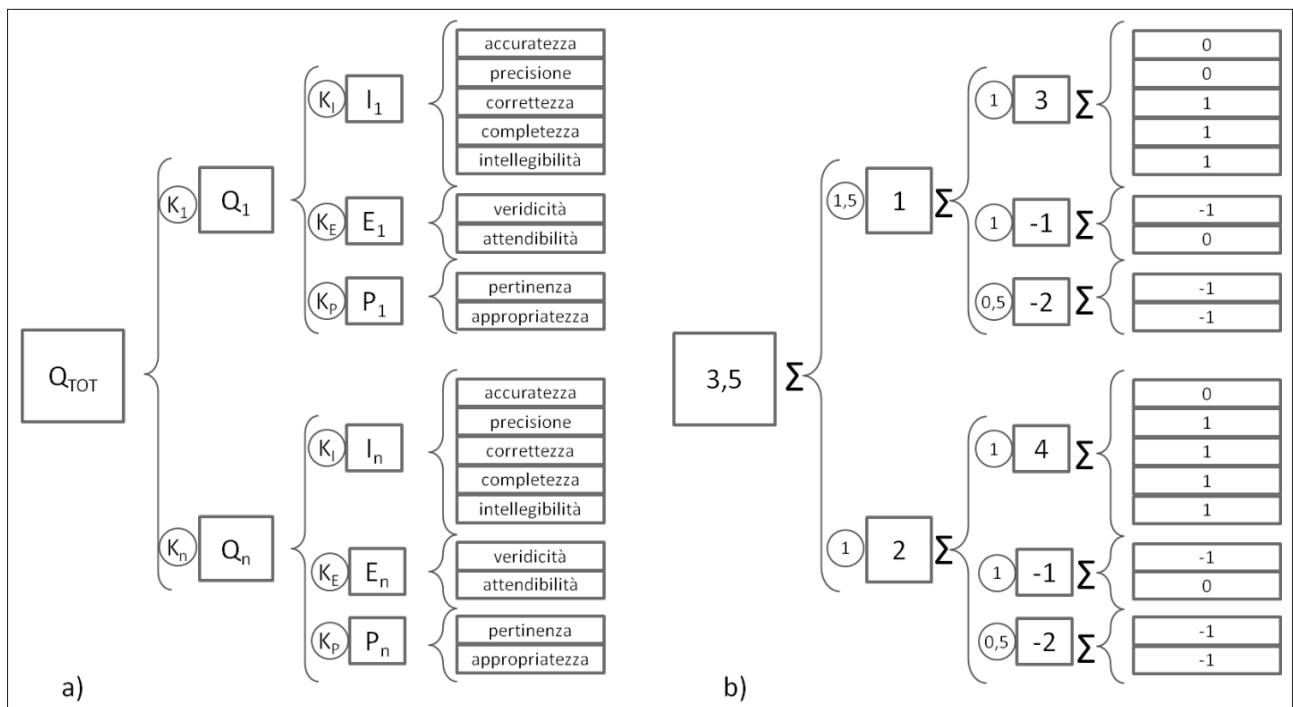
Si associa poi all'informazione complessiva VGI_{TOT} un indice di qualità Q_{TOT} , dato dall'aggregazione degli indici Q_i associati agli n componenti. Nella fase di aggregazione ad ogni indice Q_i sarà attribuito un peso numerico K_i , scelto dall'analista in funzione delle esigenze progettuali specifiche. L'aggregazione potrà avvenire con una operazione scelta dall'analista, ad esempio con una media pesata o una sommatoria.

$$Q_{TOT} = \oplus (K_1 Q_1, K_2 Q_2, K_3 Q_3, \dots, K_n Q_n)$$

Ogni indice Q_i sarà a sua volta determinato dall'aggregazione di tre indici qualitativi - I_i, E_i, P_i - relativi rispettivamente agli aspetti intrinseci, estrinseci e pragmatici della qualità.

I_i, E_i e P_i , possono avere a loro volta associati dei pesi - K_i, K_E e K_P - stabiliti in relazione all'importanza che l'analista intende attribuire ai tre aspetti della qualità. La qualità totale risulta dunque:

FIGURA 4 – Rappresentazione dello schema di aggregazione per un caso generico (a) e per un ipotetico caso operativo (b) (Elaborazione degli Autori)



$$Q_{TOT} = \oplus (K_1 * \oplus (K_{I_1}, K_{E_1}, K_{P_1}), \\ K_2 * \oplus (K_{I_2}, K_{E_2}, K_{P_2}), \dots, \\ K_n * \oplus (K_{I_n}, K_{E_n}, K_{P_n}))$$

Si possono infine scomporre i tre indici qualitativi I_i , E_i e P_i con indici di livello inferiore, rappresentativi delle proprietà elementari della qualità: accuratezza, precisione, correttezza, completezza, intelligibilità per I_i , veridicità e attendibilità per E_i , pertinenza e appropriatezza per P_i . Anche questo livello sarà legato al precedente da operazioni di aggregazione:

$$I_i = \oplus (\text{accuratezza}_i, \text{precisione}_i, \text{correttezza}_i, \text{completezza}_i, \text{intelligibilità}_i)$$

$$E_i = \oplus (\text{veridicità}_i, \text{attendibilità}_i)$$

$$P_i = \oplus (\text{pertinenza}_i, \text{appropriatezza}_i)$$

Gli indici descritti e i loro rispettivi livelli di aggregazione sono rappresentati in figura 4a.

Per meglio chiarire l'applicabilità del modello in una attività concreta di tipo VGI, esplicitiamo gli indici di qualità per un caso verosimile. Indossiamo ad esempio le vesti di analisti del progetto Wikimapia e proviamo a valutare la qualità totale di un contributo volontario, avente forma di poligono a cui è annessa una foto. Chiameremo l'indice di qualità della feature poligonale Q_1 , e quello dell'immagine Q_2 . Scegliamo innanzitutto una operazione di sommatoria per realizzare l'aggregazione e definiamo soggettivamente i pesi:

$$K_1 = 1,5 \quad K_2 = 1$$

immaginando un maggior interesse nel tutelare la qualità delle features rispetto a quella dell'immagine;

$$K_I = 1 \quad K_E = 1 \quad K_P = 0,5$$

immaginando un più modesto interesse nel garantire la qualità pragmatica dell'informazione rispetto alle qualità intrinseca ed estrinseca.

Attribuiamo dunque dei valori di esempio alle proprietà elementari di qualità, cercando di simulare una situazione plausibile per Wikimapia. Ad esempio scegliamo di assegnare valori compresi nel dominio [-1, 0, 1]:

- diamo valore nullo ad accuratezza e precisione della feature, non essendo direttamente determinabili in Wikimapia stessa, mentre diamo punteggio paria 1

a correttezza, completezza e intelligibilità, che sono valutabili ed ipotizziamo pienamente soddisfatte;

- diamo poi valore -1 e 0 rispettivamente a veridicità e attendibilità, ipotizzando che siano stati pubblicati dalla comunità dei commenti negativi relativi alla feature e che lo status di esperienza dell'autore sia quello di "neofita" (*user level 0*, o *Unregistered* in Wikimapia);
- diamo infine valore -1 sia a pertinenza che ad appropriatezza, immaginando che la feature inserita non appartenga alle categorie di interesse per l'applicazione (ad esempio riproduca l'area in cui si svolge un evento temporaneo);
- assegniamo i rispettivi valori anche alla componente immagine del contributo VGI, con ragionamenti analoghi a quanto fatto per la feature.

La situazione ipotizzata è rappresentata in Figura 4b, in analogia con il modello teorico.

Con questa procedura, che può essere svolta in modo automatico, semi-automatico, o anche manuale, si può dunque calcolare un valore di qualità per ciascun contributo volontario, il che consente di confrontare tra loro i contributi raccolti, di effettuare un ordinamento tra loro e di selezionarne i migliori applicando ad esempio un valore minimo di soglia.

4. Conclusioni

I contributi informativi geografici di tipo volontario costituiscono indubbiamente una potenziale sorgente di dati con granularità spazio-temporale molto fine. Per poterli sfruttare a scopi scientifici o professionali è necessario affrontarne la principale criticità, la qualità, e dunque individuare metodi per rappresentarla, controllarla e stimarla adeguatamente.

Questo lavoro propone una analisi articolata dell'argomento, utile per delineare in modo sintetico e sistematico i vari fattori in gioco, per guidare il coordinamento di attività di VGI e per verificarne il risultato in termini di qualità.

Dopo aver definito l'ambito di interesse e la relativa bibliografia di riferimento nel primo capitolo, si passa

a descrivere sistematicamente la qualità delle informazioni geografiche volontarie e ad analizzare le strategie che vengono comunemente adottate per controllarla. Tali strategie, che nella pratica vengono spesso combinate in approcci ibridi, sono qui descritte attraverso la loro scomposizione in proprietà atomiche. Una normalizzazione dei casi operativi complessi può dunque avvenire attraverso la rappresentazione schematica delle loro proprietà, come esposto dagli esempi operativi del paragrafo 2.4.

Proprio da questi esempi appare evidente l'estrema variabilità delle combinazioni esistenti. Ciascuna configurazione analizzata è caratterizzata da un diverso bilanciamento dei componenti, la qual cosa non è direttamente rapportabile a particolari segnali di solidità/debolezza delle applicazioni. Le diverse conformazioni risultanti per le applicazioni appaiono semplicemente adeguate ai diversi obiettivi e target.

Ciò costituisce un chiaro esempio di come, nell'ampio panorama del VGI, non esista un'unica soluzione ottimale, bensì possano coesistere configurazioni efficaci negli specifici casi d'uso. Le scelte sono determinate dallo scopo perseguito dall'attività (ludico, sociale, scientifico, professionale, sperimentale, ecc.), dal tipo e dal quantitativo di informazione attesa in ingresso (immagini, annotazioni, features, misure), dal target di volontari a cui ci si rivolge (cittadini, studenti, dilettanti, esperti, ecc.), dall'infrastruttura e dalle tecnologie a cui ci si appoggia (database geografici, client web e mobile, servizi per sensoristica, ecc.).

Questo tipo di rappresentazione può essere utilizzato non solo a posteriori, per descrivere le strategie di controllo attivate, ma può anche essere impiegato in fase di progettazione, per configurare la soluzione più efficace di gestione della qualità.

Una volta individuata ed attuata una strategia di controllo, essa necessita di essere valutata. Una stima della qualità risultante dei contributi volontari è utile per diversi scopi: per quantificare il valore apportato dal VGI, per effettuare confronti interni al set di dati o con set di dati raccolti in maniera alternativa, per selezionare i contributi di maggior valore, per monitorare l'andamento della qualità nel tempo, in relazione all'evoluzione delle altre variabili (allargarsi della comunità di volontari, variazioni nelle tecnologie o negli stru-

menti di immissione, avverarsi di eventi sensibili come cambiamenti politici e sociali, calamità naturali o emergenze umanitarie, ecc.).

La stima della qualità del VGI è una sfida che è stata affrontata con metodi differenti da diversi autori. I metodi comunemente reperibili in letteratura sono studiati tuttavia per rispondere ad esigenze specifiche, e perciò, per quanto apprezzabili ed utili nei singoli casi, appaiono affetti da uno o più dei seguenti limiti:

- mirano a quantificare l'incertezza di una singola variabile di qualità (credibilità dei volontari, accuratezza spaziale delle features, ecc.);
- si occupano di una singola tipologia di VGI (immagine, annotazione, feature, misura);
- sono adatti ad una specifica applicazione (si appoggiano ad una data tecnologia, o presumono la partecipazione di una certa quantità di volontari, ecc.).

Il contributo che si cerca di apportare con questo lavoro è quello di proporre un metodo di stima della qualità che superi queste limitazioni. Esso viene definito formalmente in termini sufficientemente operativi, per garantirne l'applicabilità, ma anche sufficientemente generali, per garantirne la trasferibilità ai diversi casi applicativi.

Il metodo proposto risulta comprensivo di tutte le proprietà di qualità e si presta a descrivere ogni contributo di tipo VGI sotto tutti i possibili aspetti. La sua forza risiede nella grande flessibilità: le operazioni di aggregazione possono essere scelte per adattarsi alla finalità dell'analisi e i pesi sono assegnati in relazione a ciascun caso specifico. Le operazioni di aggregazione, descritte qui nella loro massima complessità, possono poi essere semplificate, assegnando ad esempio peso nullo – e quindi elidendo – a proprietà ed indici che si ritengano superflui. Persino qualora non si intenda aderire alla definizione di qualità proposta (Figura 1), il metodo rimane applicabile, sostituendo indici e proprietà alternativi. Il metodo infine può fornire una guida per sistematizzare ed esplicitare i criteri di valutazione della qualità che vengono utilizzati in una applicazione di VGI.

In conclusione, le soluzioni possibili per un utilizzo controllato – scientifico o professionale – di informazione geografica volontaria non mirano all'eliminazio-

ne del problema della qualità, ma piuttosto alla gestione controllata delle variabili in gioco, estremamente dipendenti dal caso specifico.

Conseguentemente, anche i metodi di stima e di quantificazione della qualità devono adattarsi ai diversi aspetti caratterizzanti l'attività, ed essere comprensivi

delle molteplici sfaccettature che la qualità del VGI può assumere.

Infine si evidenzia come una corretta rappresentazione sia un valido strumento di supporto sia nella fase di definizione del problema, sia nella conseguente fase di individuazione delle risposte più efficaci.

Bibliografia

- BONNEY R., COOPER C.B., DICKINSON J. (2009), Citizen science: a developing tool for expanding science knowledge and scientific literacy, "BioScience", 11 (2009), pp. 977-984.
- BORDOGNA G., CARRARA P., CRISCUOLO L., PEPE M. e RAMPINI A. (2014), A linguistic decision making approach to assess the quality of volunteer geographic information for citizen science. "Information Sciences", 258 (2014), pp. 312-327.
- BORRUSO G. (2010), La 'nuova cartografia' creata dagli utenti. Problemi, prospettive, scenari, "Bollettino A.I.C.", 138 (2010), pp. 241-25.
- BORRUSO G. (2013), Cartografia e Informazione Geografica "2.0 e oltre", Webmapping, WebGIS, "Bollettino A.I.C.", 147 (2013), pp. 7-16.
- BRUNS A. (2006), Towards produsage: futures for user-led content production, Proceedings of: Cultural Attitude towards Communication and Technology, pp. 275-84, F. Sudweeks, H. Hrachovec, and C. Ess (eds). MurdochUniversity, Perth, Australia.
- CAMPBELL A. T., EISENMAN S. B., LANE N. D., MILUZZO E. e PETERSON R. A. (2006), People-centric urban sensing, Proceedings of the 2nd annual international workshop on Wireless internet , p. 18, ACM.
- COLEMAN D. J., GEORGIADOU Y. e LABONTE J. (2009), Volunteered Geographic Information: the nature and motivation of producers, "International Journal of Spatial Data Infrastructures Research", 4(2009), pp. 332-358.
- CONNORS J. P., LEI S. e KELLY M. (2012), Citizen science in the age of neogeography: Utilizing volunteered geographic information for environmental monitoring, "Annals of the Association of American Geographers", 6 (2012), pp. 1267-1289.
- CRALL, A. W., NEWMAN, G. J., STOHLGREN, T. J., HOLFELDER, K. A., GRAHAM, J. e WALLER, D. M. (2011), Assessing citizen science data quality: an invasive species case study, "Conservation Letters", 4(2011), pp. 433-442.
- CRISCUOLO L., PEPE M., SEPPI R., BORDOGNA G., CARRARA P. E ZUCCA F. (2013), Alpine Glaciology: an historical collaboration between volunteers and scientists and the challenge presented by an integrated approach, "International Journal of Geo-Information", 2(2013), pp. 680-703
- DE LONGUEVILLE B., ANNONI A., SCHADE S., OSTLAENDER N. e WHITMORE C. (2010), Digital earth's nervous system for crisis events: Real-time Sensor Web enablement of volunteered geographic information. "International Journal of Digital Earth", 3(2010), pp. 242-259.
- DE TRÉ G. e BRONSELAER A. (2010), Consistently handling geographical user data: Merging of coreferent POIs, Fuzzy Information Processing Society (NAFIPS), 2010 Annual Meeting of the North American, pp. 1-6, IEEE.
- DICKINSON J. L., ZUCKERBERG B. e BONTER D. N. (2010), Citizen Science as an Ecological Research Tool: Challenges and Benefits. "Annual Review of Ecology, Evolution, and Systematics", 41 (2010), pp. 149-172.
- ENGLISH L.P. (1999), *Defining Information Quality*, in *Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits*, John Wiley and Son, pp. 544.
- EYSENBACH G., e KÖHLER C. (2002), How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews, "Bmj: British Medical Journal", 324(2002), pp. 573-577.
- FLANAGIN A. J. e METZGER M. J. (2008), The credibility of volunteered geographic information, "GeoJournal", 72(2008), pp. 137-148.
- GALLOWAY A. W. E., TUDOR M. T. e HAEGEN W. M. (2006), West The Reliability of Citizen Science: A Case Study of Oregon White Oak Stand

- Surveys, "Wildlife Society Bulletin", 34(2006), pp. 1425-1429.
- GIANNOLA E. (2013), Il ruolo dei Web - GIS nella partecipazione civica al processo decisionale, "Bollettino A.I.C.", 147 (2013), pp. 41-52.
- GOODCHILD M. (2007), Citizens as Sensors: The World of Volunteered Geography, "GeoJournal", 4(2007), pp. 211-221.
- GOODCHILD, M. (2008), Spatial accuracy 2.0., Proceedings of the eighth international symposium on spatial accuracy assessment in natural resources and environmental sciences, pp. 1-7.
- HAKLAY M. (2010), How good is volunteered geographic information? A comparative study of OpenStreetMap and Ordnance Survey datasets, "Environment and planning. B, Planning & design" 37(2010), pp. 682-703.
- HAKLAY M., SINGLETON A. e PARKER C. (2008), Web mapping 2.0: The neo-geography of the GeoWeb, "Geography Compass", 6(2008), pp.2011-2039.
- HEIPKE C. (2010), Crowdsourcing geospatial data, "ISPRS Journal of Photogrammetry and Remote Sensing", 65(2010), pp. 550-557.
- HOVLAND C. I., IRVING L. J. e HAROLD H. K. (1953), Communication and persuasion: Psychological studies of opinion change, CT: Yale University Press, New Haven, pp. 134.215.
- HOWE J. (2006). The rise of crowdsourcing, "Wired magazine", 6(2006), pp. 1-4.
- HUANG, K. L., KANHERE, S. S., HU W. (2010), Are you contributing trustworthy data?: the case for a reputation system in participatory sensing, Proceedings of the 13th ACM international conference on Modeling, analysis, and simulation of wireless and mobile systems, pp. 14-22, ACM.
- KEBLER C. e DE GROOT R. T. A. (2013), Trust as a Proxy Measure for the Quality of Volunteered Geographic Information in the Case of OpenStreetMap, Geographic Information Science at the Heart of Europe, Springer International Publishing, pp. 21-37.
- KUHN W. (2001), Ontologies in support of activities in geographical space. "International Journal of Geographical Information Science", 15(2001), pp. 613-631.
- LANE N. D., EISENMAN S. B., MUSOLESI M., MILUZZO E. e CAMPBELL A. T. (2008), Urban sensing systems: opportunistic or participatory?, Proceedings of the 9th workshop on Mobile computing systems and applications, pp. 11-16, ACM.
- LATONERO M. e SHKLOVSKI I. (2010), Respectfully yours in safety and service: emergency management and social media evangelism, Proceedings of the 7th Int. Conference on Information Systems for Crisis Response and Management, Seattle, Vol. 1.
- METZGER M. J. (2007), Making sense of credibility on the Web: Models for evaluating online information and recommendation for future research." Journal of the American Society for Information Science and Technology", 13(2007), pp. 2078-2091.
- NATIONAL COMMITTEE FOR DIGITAL CARTOGRAPHIC STANDARDS (US), MOELLERING H, GEOLOGICAL SURVEY (US) (1987), A Draft Proposed Standard for Digital Cartographic Data, National Committee for Digital Cartographic Data Standards.
- OSTERMANN F.O. e SPINSANTI L. (2011), A conceptual workflow for automatically assessing the quality of volunteered geographic information for crisis management, Proceedings of AGILE 2011, pp. 10-14.
- POPESCU A., GREFFENSTETTE G. e BOUAMOR H. (2009), Mining a Multilingual Geographical Gazetteer from the Web, IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, Ieee, 2009, pp. 58-65.
- PUNDT H. (2002), Field Data Collection with Mobile GIS: Dependencies Between Semantics and Data Quality, "GeoInformatica", 6(2002), pp. 363-380.
- ROCCA L. (2013), I GeoBlog: strumenti per una "cartografia aumentata", "Bollettino A.I.C.", 147 (2013), pp. 17-39.
- SABONE B. (2009), "Assessing Alternative Technologies for Use of Volunteered Geographic Information in Authoritative Databases". Unpublished M.Sc.E. Thesis, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, N.B. Canada. Retrieved from <http://www2.unb.ca/gge/Pubs/TechnicalReports.html>.
- TURNER A. (2006), Introduction to neo-geography, CA: O'Really Media, Sebastopol.
- VAN DIJCK J. (2009), Users like you? Theorizing agency in user-generated content, "Media, culture, and society", 1(2009), pp. 41.
- VAN EXEL M., DIAS E., FRUIJTIER S. (2010), The impact of crowdsourcing on spatial data quality indicators, Proceedings of the 6th GIScience international conference on geographic information science, 2010, p. 213-217.
- VAN OORT P. A. J. (2006), Spatial Data Quality: From Description to Application, Wageningen: Wageningen Universiteit. PhD.