



UNIVERSITÀ DI PISA

LAUREA MAGISTRALE IN
LETTERATURE E FILOGIE EUROPEE

SEMINARIO DI CULTURA DIGITALE A.A. 2017/18

La conoscenza nel nuovo web tra Wikipedia e Wikidata

Progetti e proposte per un “database dello scibile umano”

Michele Gandolfi

Matricola: 443218

Sommario

Una panoramica sul mondo della conoscenza libera e collaborativa di Wikipedia e Wikidata. Ampliare l'enciclopedia e il database per raccogliere lo scibile umano: progetti avviati e un resoconto della propria esperienza di collaboratore.

Indice

Indice	2
1. Velocità, collaborazione, libertà: sviluppo e fortuna di Wikipedia	3
2. Wikidata: il database collaborativo della conoscenza	5
3. Completare il puzzle: proposte e progetti per Wikipedia e Wikidata	6
3.1. Breve storia di un utente wiki	6
3.2. Incrementare il database tra località geografiche, stazioni ferroviarie, vescovi e sindaci	7
3.3. Un esempio specifico su Wikidata: il sindaco Ippolito Luciani	10
4. Bibliografia	11
5. Sitografia	11

1. Velocità, collaborazione, libertà: sviluppo e fortuna di Wikipedia

La divulgazione, la fruibilità e la velocità caratterizzante questi due elementi sono i concetti cardine intorno ai quali ruota la conoscenza del ventunesimo secolo. Se già dalla fine del secolo scorso si intravedevano le avvisaglie di una sempre maggiore necessità di fruizione il più possibile istantanea di dati e materiale che si avvicinasse all'utilizzatissimo dizionario enciclopedico, con l'avvento del web e soprattutto a partire dai primi anni duemila, questo fenomeno oltre che inarrestabile è divenuto progressivamente fondamentale. Nell'epoca del Web 2.0, termine coniato dall'informatica Darcy DiNucci nel 1999 e poi attestatosi stabilmente a partire dal 2004 su iniziativa di Tim O'Reilly e Dale Dougherty, il favorire costantemente l'iterazione e il libero scambio interattivo di contenuti e informazioni sono diventati la principale finalità della ricerca informatica. La concezione di web come piattaforma – «web as platform»¹ – in contrapposizione alla schematicità statica della prima forma di rete, il retroattivamente noto Web 1.0, ha permesso la formazione e lo sviluppo di blog, social network e siti che consentono una condivisione sempre maggiore di immagini o video; basti pensare a nomi ormai nettamente conosciuti ed entrati nel quotidiano quali MySpace, Facebook, Twitter, YouTube. Non a caso, proprio per la sua valenza interattiva e di aggregazione, il web 2.0 è stato anche definito *social web*, e già Tim Berners-Lee, inventore del World Wide Web, aveva scritto nel 1999 che «the Web is more a social creation than a technical one. I designed it for social effect – to help people work together – and not as a technical toy».² Non meravigliamoci dunque se tra i siti più visitati al mondo, in qualsiasi classifica mensile, Wikipedia, la nota enciclopedia online, figura sempre nelle prime dieci posizioni.

Nata nel 2001 da un'intuizione di Jimmy Wales e Larry Sanger come evoluzione della precedente loro “creatura” Nupedia, il successo di Wikipedia è dovuto principalmente nella libertà di scambio ed elaborazione dei dati online da parte della collettività. Se inizialmente nell'idea dei creatori vi era la necessità di dare vita a un'enciclopedia online che avrebbe potuto superare per contenuti e grado di dettaglio le più autorevoli controparti cartacee – il loro proposito era quello di “battere” l'*Encyclopædia Britannica* – sempre incentrata sul volontariato e la collaborazione, ma portata avanti da accademici qualificati che avrebbero potuto pubblicare su Nupedia solo dopo un attento lavoro di revisione, ben presto si avvertì l'esigenza di accelerare i tempi e snellire i passaggi, tanto che Nupedia divenne Wikipedia quando Wales comprese i vantaggi dell'utilizzo del software wiki (in hawaiano, “veloce”), un'applicazione web per la creazione e la modifica collaborative di pagine all'interno di un sito internet utilizzando un linguaggio di markup semplificato, ideato da Ward Cunningham nel 1995. Fu un successo immediato, centinaia di utenti si collegarono a Wikipedia e iniziarono a contribuire con le loro competenze fino a diventare migliaia, rendendo l'enciclopedia via via sempre più ricca. Da un punto di vista qualitativo non si raggiunse certo l'autorevolezza delle voci di Nupedia – il processo di revisione non sparisce con Wikipedia, ma avviene tra la comunità di utenti in tempi non convenzionali, possono passare pochi giorni oppure anni prima che una voce venga migliorata – mentre da un punto di vista quantitativo si poté arrivare in poco tempo a livelli che nessuna enciclopedia cartacea aveva raggiunto prima. Si potrebbe dire che l'enorme mole di dati riversata da utenti determinati a divenire parte del sistema di divulgazione culturale abbia sancito quasi del tutto l'obsolescenza degli ingombranti e limitati volumi di cellulosa. Che poi è proprio questo dopotutto a definire, come abbiamo visto, la natura delle rete 2.0: sono le persone qualsiasi a programmare il web e proprio loro – noi – rivestono sempre più un ruolo di autorità: il

1 John Battelle e Tim O'Reilly parlarono del web come piattaforma nel discorso d'apertura nella prima conferenza del Web 2.0 tenutasi il 5 ottobre 2004 a San Francisco.

2 TIM BERNERS-LEE, *Weaving the Web*, New York, HarperCollins, 1999. Ed. italiana: *L'architettura del nuovo Web*, Milano, Feltrinelli, 2001.

«potere dei consumatori», lo chiamano.³ Si socializza, si collabora e si pratica il *peering*, anche perché, come ricorda il fondatore di Socialtext per le wiki d'impresa Ross Mayfield, «la nuova rete è tutta una questione di verbi, non di sostantivi».⁴ Una sfumatura non da poco, che in realtà cela la vera rivoluzione della condivisione dei dati sul web e in particolare su Wikipedia: sul piano teorico la nuova divulgazione della conoscenza parte dal basso, controllata opportunamente dagli stessi che caricano ogni giorno nuove pagine. L'evanescenza della gerarchia piramidale che gestisce il sapere online lascia, a mio parere, ampi margini per una sana critica costruttiva a tutto il sistema, marcando ancora una volta l'assunto, forse banale, ma assolutamente mai scontato, che di per sé Wikipedia non sia altro che uno strumento neutro e potenzialmente positivo. Quello che conta è la modalità con cui questo strumento viene gestito e utilizzato: al di là delle considerazioni snobistiche su autorevolezza e qualità dei contenuti – ricordando che su Wikipedia una voce è valida solo se è corredata da fonti verificabili e quindi sta anche al fruitore comprenderne il valore – e passando oltre quelle etiche che comprensibilmente scaturiscono dalla prospettiva di un binomio domanda-risposta improntato sull'*hic et nunc*, che nel quotidiano svilisce la ricerca connotandola con istanze superficiali e puramente nozionistiche, è innegabile quanto la fortuna di Wikipedia non rimanga solamente relegata all'ambito di un servizio – ho bisogno di un'informazione, cerco su Wikipedia e la trovo – ma si estenda anche nel ridefinire il concetto di enciclopedicità (*notability*) e nello stabilire nuove strutture che oltre a disegnare formalmente l'enciclopedia stessa definiscono anche un nuovo modo di visualizzare e strutturare il sapere, che rispecchiano la mente umana quale archivio organizzato di dati.

Inoltre, non dimentichiamoci che una delle più grandi novità di Wikipedia è proprio quella di essere un'enciclopedia sottoposta ad una licenza libera, come da sempre dichiarato sin dal principio dal fondatore Jimmy Wales: «offriamo alla gente quattro gradi di libertà: libertà di copiare il nostro lavoro, libertà di modificarlo, libertà di ridistribuirlo, libertà di ridistribuirne versioni modificate. [...] Quando si parla di Wikipedia come di un'enciclopedia libera, non ci si riferisce al prezzo da pagare per avere accesso alle informazioni, bensì alla libertà di acquisirle, adattarle e utilizzarle a proprio piacimento».⁵ Anche questa libertà ha favorito il coinvolgimento di sempre più utenti che si sentono parte attiva di una missione divulgativa libera da vincoli editoriali e copyright, in uno spirito romanticamente anarchico quale era avvertito lo spirito del web sin dalle sue origini. Anche se va ricordato che vi fu un tentativo nel febbraio 2002, agli inizi dell'avventura wikipediana, da parte della Bomis – azienda che aveva sviluppato Nupedia e che all'epoca aveva ancora potere anche su Wikipedia – di vendere inserzioni pubblicitarie all'interno dell'enciclopedia: in realtà si trattò di una proposta, più che di un vero e proprio tentativo, ma il risultato fu quello di provocare un'alzata di scudi da parte degli utenti che si dichiararono indignati all'idea che qualcuno potesse guadagnare sul lavoro di volontari, minacciando di abbandonare il progetto. Soprattutto l'edizione spagnola di Wikipedia, la più attiva al tempo dopo quella in lingua inglese, espresse con vigore il proprio disappunto, arrivando a fondare una nuova enciclopedia, la *Enciclopedia Libre*, con il patrocinio dell'università di Siviglia, dove trasferì tutte le voci già scritte su Wikipedia e all'interno delle quale si spostarono la maggiorparte degli utenti: fu un avvenimento noto come lo “scisma spagnolo”, che impedì per lungo tempo all'edizione spagnola di rientrare tra le maggiori versioni linguistiche di Wikipedia nonostante il potenziale alto numero di contributori madrelingua;⁶ superata per numero di voci prima da tedesco, francese, italiano, russo, olandese e giapponese, poi

3 DON TAPSCOTT, ANTHONY D. WILLIAMS, *Wikinomics 2.0. La collaborazione di massa che sta cambiando il mondo*, Milano, Rizzoli, 2008, p. 88.

4 *ibidem*

5 JIMMY WALES, *Prefazione*, in ANDREW LIH, *La rivoluzione di Wikipedia. Come un gruppo di illustri sconosciuti ha creato la più grande enciclopedia del mondo*, Torino, Codice Edizioni, 2010, p. IX.

6 LIH, *La rivoluzione di Wikipedia*, cit., pp. 158-160.

da svedese e cebuano; soltanto negli ultimi anni è ritornata preponderante, arrivando a piazzarsi all'ottavo posto della classifica.⁷

2. Wikidata: il database collaborativo della conoscenza

Wikipedia è solo uno dei tanti progetti gestiti dalla Wikimedia Foundation, fondazione senza fini di lucro nata nel 2003 sempre ad opera di Jimmy Wales con lo scopo di incoraggiare lo sviluppo e la diffusione di contenuti liberi: tale fondazione comprende infatti, tra i principali, un dizionario online, Wikizionario (*Wikictionary*), che funge anche da thesaurus; una piattaforma dedicata all'attualità quale l'agenzia di stampa Wikinotizie (*Wikinews*); una raccolta di citazioni di autori o personalità celebri estrapolate da libri, pubblicazioni o interviste, ma anche frasi famose di film, romanzi o fumetti (*Wikiquote*); una biblioteca che raccoglie materiali originali di saggistica e manualistica sviluppati dagli utenti (*Wikibooks*) e una biblioteca che invece conserva online testi e documenti di autori ormai in pubblico dominio (*Wikisource*); una vera e propria guida turistica con informazioni su cosa vedere, dove alloggiare, mangiare e come spostarsi (*Wikivoyage*); infine, non bisogna dimenticare quello che probabilmente è il maggiore dei progetti insieme a Wikipedia, ovvero *Wikimedia Commons*, il vasto archivio di immagini, suoni e video liberamente utilizzabili che fungono da supporto grafico alle voci dell'enciclopedia. Tutti questi progetti dovevano inizialmente essere collegati tra di loro inserendo un link nelle rispettive pagine – ad esempio, collegando la pagina Pisa di Wikipedia alla categoria Commons che contiene le immagini su tale città, o al progetto Wikivoyage per informazioni turistiche o a Wikiquote per avere citazioni su Pisa – così come le varie versioni linguistiche dovevano essere collegate singolarmente – inserendo ad esempio [[en:Pisa]] sulla voce italiana per collegarla a quella inglese e viceversa [[it:Pisa]] su quella inglese per collegarla a quella italiana. Dal 2012 per tutti i progetti esiste una modalità automatica che ha facilitato e migliorato la complementarità delle informazioni: viene lanciato Wikidata. Siamo giunti nell'epoca del Web 3.0, il cosiddetto *Semantic Web*, e anche Wikipedia ha così sviluppato una modalità per strutturare e riorganizzare i numerosissimi dati di cui dispone tramite una propria sintassi formale.

Wikidata si definisce «un database libero, collaborativo, multilingue e secondario che raccoglie dati strutturati per fornire supporto a Wikipedia, a Wikimedia Commons, agli altri progetti del movimento Wikimedia e a chiunque nel mondo»⁸ e si tratta quindi di un grande archivio di deposito centrale che costituisce la base di conoscenza di tutti i progetti Wikimedia. Wikipedia, e qualsiasi altro progetto Wikimedia, prende così in automatico le informazioni da Wikidata: se inizialmente ogni singola pagina dell'enciclopedia prevedeva l'inserimento di tutte le informazioni di base direttamente su di essa, così come i collegamenti con gli altri progetti e le altre versioni linguistiche, con Wikidata la voce dell'enciclopedia può invece ridursi alla sola trattazione enciclopedica. Esempio: per le pagine su elementi geografici non ci sarà bisogno di trascrivere manualmente informazioni quali coordinate geografiche, suddivisione amministrativa di appartenenza, codice postale, fuso orario e altri dati statistici in quanto verranno importate automaticamente da Wikidata. Così come per qualsiasi voce non ci sarà bisogno di inserire collegamenti manuali in ogni singola pagina per eventuali altre versioni in lingue diverse o per altri progetti come Commons, Wikinotizie, Wikivoyage, Wikisource, basterà inserire tali collegamenti sulla sola pagina dell'elemento di Wikidata. Esattamente come Wikipedia, Wikidata è visualizzabile come un insieme

⁷ https://meta.wikimedia.org/wiki/List_of_Wikipedias#All_Wikipedias_ordered_by_number_of_articles, *List of Wikipedias* su Wikipedia. La Wikipedia in lingua italiana si trova al nono posto.

⁸ <https://www.wikidata.org/wiki/Wikidata:Introduction/it>, *Wikidata:Introduzione*, su [wikidata.org](https://www.wikidata.org)

di pagine, forma in cui i dati sono strutturati. Ogni soggetto è chiamato entità (*entity*) e ogni entità possiede una propria pagina su Wikidata. Il sistema distingue due tipi di entità: gli elementi (*items*) e le proprietà (*properties*): gli elementi rappresentano singoli soggetti o classi, mentre le proprietà sono modellate sull'esempio di quelle RDF (*Resource Description Framework*), modelli di dati standard per la rappresentazione della conoscenza. Il titolo di ogni pagina avrà come identificativo “Q” seguito da una sequenza numerica progressiva, creata in automatico, e non il nome dell'entità in una lingua specifica: quello sarà compito delle etichette (*labels*), declinabili e traducibili nelle rispettive lingue del mondo, e degli *alias*, ovvero eventuali variazioni o altri nomi con i quali è conosciuto il soggetto. Ogni entità è quindi contraddistinta da un URI: ad esempio, per il sopracitato caso di Pisa, avremo <http://www.wikidata.org/entity/Q13375>.⁹ Più avanti mostrerò come effettuare i collegamenti portando un esempio specifico.

Oltre alla funzione di archivio di deposito centrale, è importante ricordare che Wikidata è utilizzato anche per integrare e disseminare la conoscenza: fornisce infatti collegamenti con vari database e ontologie, favorendo la complementarità delle informazioni nel Semantic Web. Infatti, essendo Wikipedia uno dei siti più consultati al mondo, quasi tutti i motori di ricerca, ma in particolare Google, assimilano da Wikidata numerosissimi dati che poi rielaborano per aiutare e semplificare la ricerca di informazioni da parte degli utenti.

3. Completare il puzzle: proposte e progetti per Wikipedia e Wikidata

3.1. Breve storia di un utente wiki

Quando mi sono personalmente avvicinato al mondo di Wikipedia nel 2009, l'idea di collaborare all'enciclopedia e rendere accessibile a tutti una serie di informazioni in mio possesso in vari campi di mio interesse la consideravo – e la considero tuttora – decisamente stimolante. La prima voce che ho scritto era dedicata a un letterato poco conosciuto ma relativamente importante nella storia del volgare italiano del XIII secolo, nonché mio concittadino di età remota: Andrea da Grosseto. La seconda voce che scrissi, poco dopo, fu “Zancona”, dedicata a una piccola frazione sul monte Amiata e luogo di provenienza di parte della mia famiglia nonché meta prediletta per le mie vacanze sin dall'infanzia. Questi due motivazioni prettamente personali (un mio concittadino illustre e una località geografica a me cara) mi hanno spinto ad effettuare ricerche sempre più approfondite tanto che da due incerti abbozzi privi di qualsivoglia autorevolezza, le due voci sono diventate negli anni, anche grazie ai contributi di altri utenti, voci enciclopediche sempre più complete e accurate, dotate di fonti, bibliografia, immagini. Dopodiché mi sono chiesto: ma perché limitarsi a scrivere la sola voce su Zancona? Perché non completare la lista creando le voci anche delle altre frazioni della mia provincia che ancora non esistono su Wikipedia? Da quel momento, l'idea di “completare il puzzle”, di inserire tutto ciò che potesse essere inserito – nel rispetto delle regole e nel rispetto dell'enciclopedicità (*notability*) dei soggetti – come una sorta di almanacco dello scibile umano, aveva iniziato a farsi strada in me, tanto che nel giro di alcuni anni completai le voci inerenti alle frazioni della provincia di Grosseto, per poi passare a quelle di Pisa, altro territorio al quale sono particolarmente legato, e poi Siena e dopo Livorno, nell'ottica di voler vedere completa l'enciclopedia delle voci sui principali centri abitati della Toscana e poi, perché no, dell'Italia intera. Le statistiche evidenziano come dal 2009 ad oggi abbia creato su Wikipedia 914 voci su qualsiasi argomento di mio interesse (dalla letteratura al cinema, dall'archeologia alla politica, dalla religione

⁹ WERNER LEYH, HOMERO FONSECA FILHO, *Interlinking Standardized OpenStreetMap Data and Citizen Science Data in the OpenData Cloud*, in ISABEL L. NUNES, *Advances in Human Factors and Systems Interaction*, Springer International Publishing, 2018, pp. 87-88.

ai fumetti), con una predilezione per geografia, storia e arte riferite al mio territorio.¹⁰ Tutto questo sempre spinto da quella volontà di raccogliere e strutturare lo scibile umano riferito a questo o quell'altro ambito di mio interesse rendendolo fruibile a chiunque, e stimolare allo stesso tempo la ricerca e l'approfondimento su quel determinato soggetto.

Non a caso, anche varie istituzioni culturali si sono interessate al fenomeno Wikipedia, a partire da quelle locali, come comuni e province, che favoriscono la divulgazione anche come un modo per far conoscere e valorizzare il proprio territorio¹¹ – facendo però ben attenzione a non scadere nel localismo e nell'effetto guida turistica, in quanto per quel fine esiste appositamente il progetto gemello Wikivoyage – oppure nazionali, come è stato fatto nel marzo 2018 dall'Istituto Centrale per gli Archivi (ICAR) all'interno del progetto GLAM per biblioteche, archivi ed istituzioni culturali: attraverso i siti del Sistema Archivistico Nazionale (SAN) e del Sistema Informativo Unificato per le Soprintendenze Archivistiche (SIUSA) sono stati rilasciati con licenza libera migliaia di testi, soprattutto biografie di personalità illustri i cui archivi sono raccolti e censiti da ICAR. Nel mese di marzo è stata infatti indetta una maratona per coinvolgere gli utenti di Wikipedia ad utilizzare quei testi non più proprietà intellettuale di ICAR, iniziativa che ha registrato un buon successo di partecipazione della comunità portando alla creazione di 128 nuove voci e all'ampliamento di altre 26, alla quale ho preso parte classificandomi terzo.¹²

3.2. Incrementare il database tra località geografiche, stazioni ferroviarie, vescovi e sindaci

Numerosi sono i progetti e le strategie di collaborazione attivati tra gli utenti Wikimedia per implementare l'enciclopedia e soprattutto arricchire il database di Wikidata con nuovi elementi, portando la qualità delle dichiarazioni e delle proprietà ad un elevato grado di dettaglio. La modalità che viene spesso utilizzata, che permette in modo più veloce di immettere un grande numero di dati riferiti ad uno specifico ambito circoscritto, è quella di importare automaticamente i dati da archivi online già esistenti.

Nel 2014, ad esempio, l'archivio Wikidata è stato arricchito di centinaia di migliaia di elementi estrapolando i dati riguardanti i centri abitati italiani dagli identificativi GeoNames, grande database geografico.¹³ L'iniziativa, voluta da parte degli utenti serbi e poi incrementata dai colleghi filippini, prevedeva inoltre un collegamento di ogni singolo elemento alla rispettiva voce su Wikipedia tramite l'utilizzo di un bot: il risultato è che sulle edizioni delle enciclopedie in serbo, in serbo-croato e in cebuano si trovano voci riguardanti non solo tutti i centri abitati dell'Italia, ma anche corsi d'acqua, rilievi, aree naturali e qualunque altro elemento geografico italiano presente in GeoNames. Ovviamente, l'inserimento massiccio di una tale mole di dati non è esente da critiche, anzi, proprio queste “trasfusioni” prive di controlli più oculati spesso rendono il database caotico e la successiva sistemazione manuale, elemento per elemento, ancora più faticosa. Su Wikidata, prima di tutto bisogna accertarsi di eventuali ripetizioni e minime variazioni di etichetta: un centro abitato,

10 <https://xtools.wmflabs.org/pages/it.wikipedia.org/Alienautic?limit=1000>, *Statistiche utente*, su xtools.wmflabs.org

11 La stessa Provincia di Grosseto mi contattò nel 2014 per commissionarmi un progetto di stesura voci sui quarantasei musei della rete museale del territorio provinciale. Sempre per approfondire la divulgazione inerente al Grossetano, avevo creato nel 2011 una mia personale enciclopedia, dal poco ponderato nome “Grossetopedia”, che usufruiva del software wiki della piattaforma Wikia, nella quale potevo scrivere liberamente informazioni sulla mia città e provincia senza temere di scadere in eccessivo localismo. Vedere <http://www.it.grossetopedia.wikia.com>

12 <https://www.wikimedia.it/la-wiki-maratona-record-degli-archivi-158-voci-wikipedia-nuove-aricchite/>, *La wiki maratona record degli archivi. 158 voci Wikipedia nuove o arricchite*, su [wikimedia.it](https://www.wikimedia.it), 29 marzo 2018.

13 <http://www.geonames.org/>, Geonames.org

o elemento geografico, che è censito in GeoNames con due nomi diversi, entrambi presenti in quel database, sarà presente su Wikidata con due elementi distinti, anziché in un solo elemento con un'unica etichetta (il toponimo ufficiale o più comune) e eventuale alias per l'altro toponimo, rendendo necessaria un'unione; su Wikipedia, invece, si corre il rischio di avere un numero altissimo di voci tuttavia prive di qualsiasi valore o informazioni enciclopediche, di fatto inutili: non è possibile avere trattazioni su qualsiasi località geografica dell'Italia, perché la divulgazione enciclopedica segue altre regole, e contano prima di tutto la rilevanza del soggetto trattato e la presenza di fonti attendibili che documentano una trattazione già esistente realizzata da terzi; le voci create con il bot possiedono le stesse informazioni statistiche presenti su Wikidata, mentre mancano del tutto quelle proprie in forma discorsiva che raccontano storia, contesto geografico, eventuali monumenti, infrastrutture, eventi culturali eccetera, senza considerare il fatto che più della metà di quelle voci saranno comunque destinate a rimanerne prive anche in seguito ad un controllo capillare, considerando che non sarà possibile esprimere una trattazione enciclopedica su ogni singolo centro abitato o località censita in GeoNames, che aveva preso a sua volta i propri dati dai censimenti dell'Istat (in questo caso l'ultimo censimento del 2011), così come su ogni fosso, rigagnolo o minima altura presente in Italia! Non stupiamoci quindi se la Wikipedia in lingua cebuana è la seconda al mondo per numero di voci dopo quella inglese: la maggior parte di quelle voci sono in effetti vuote perché create in automatico e destinate a rimanere tali.

Tra gli altri esempi notevoli per mole di dati, è ragguardevole il progetto portato avanti dalla comunità ungherese, la quale si è distinta per aver caricato su Wikidata – e creato le rispettive voci sulla Wikipedia in ungherese – tutte le stazioni presenti nelle linee ferroviarie nazionali d'Europa, sempre tramite l'utilizzo di un bot automatico. Ma anche gli utenti italiani non sono da meno: attualmente è in corso il progetto di trasferire su Wikidata tutti gli elementi dell'archivio della Conferenza Episcopale Italiana del sito *Le chiese delle diocesi italiane*, che raccoglie tutti gli edifici religiosi (chiese, basiliche, cappelle, oratori) delle diocesi d'Italia.¹⁴

Un progetto non ancora avviato, ma auspicabile, è quello inerente agli elementi compresi in liste enciclopediche chiuse, quindi almanaccabili senza timore di dispersione, come nel caso delle entità geografiche: si tratta delle cronotassi dei vescovi. Per l'importazione automatica è presente l'archivio Catholic Hierarchy, che fornisce tutte le serie vescovili dal XV secolo ad oggi, e ogni personalità è contraddistinta da un proprio identificativo che potrà essere importato. Dei prelati precedenti al XV secolo, invece, le liste sono perlopiù incomplete ed è necessario ricostruire l'elenco, ma con attenzione, aiutandosi con pubblicazioni e studi storici specifici, senza cadere nell'errore di affidarsi troppo alle cronotassi tradizionali, spesso errate in quanto non attestate da documenti (è il caso delle pubblicazioni quali *l'Italia sacra* di Ferdinando Ughelli o *Le Chiese d'Italia della loro origine sino ai nostri giorni* di Giuseppe Cappelletti che hanno consegnato alla storia cronotassi spesso confuse e non convalidate da fonti documentarie certe). Personalmente ho ricostruito la serie dei vescovi della diocesi di Grosseto, creando per ognuno di loro non solo l'elemento su Wikidata, ma anche la voce biografica su Wikipedia; fanno eccezione tre vescovi (Azzo, Ermanno e Filippo Bencivenne) sui quali per il momento non sono riuscito a trovare abbastanza informazioni da permettermi di creare un abbozzo biografico soddisfacente.¹⁵

Un ultimo progetto che potrebbe partire a breve è quello inerente alle voci biografiche di personalità politiche. Secondo le regole di Wikipedia, gli unici politici che possono avere una voce dedicata

14 <http://www.chieseitaliane.chiesacattolica.it/chieseitaliane/index.jsp>, *Le chiese delle diocesi italiane*, su [chieseitaliane.chiesacattolica.it](http://www.chieseitaliane.chiesacattolica.it)

15 https://it.wikipedia.org/wiki/Diocesi_di_Grosseto#Cronotassi, *Diocesi di Grosseto. Cronotassi*, su [wikipedia.org](http://www.wikipedia.org)

sull'enciclopedia automaticamente per il solo fatto di fare politica sono di quattro tipi: 1) coloro che hanno ricoperto cariche pubbliche ai massimi livelli, quali capi di stato, capi di governo, ministri e sottosegretari, oppure coloro che sono stati membro del Parlamento; 2) coloro che sono stati a capo di suddivisioni territoriali nazionali di primo livello (nel caso italiano, presidente di regione), di suddivisioni territoriali nazionali di secondo livello (nel caso italiano, presidente di provincia) o sindaci di città capoluogo; 3) coloro che sono stati firmatari dell'atto costitutivo di un partito politico rilevante a livello nazionale; 4) coloro che hanno ricoperto il massimo ruolo direttivo di movimenti o partiti politici di rilevanza nazionale. Sofferamoci solamente su quei politici che hanno ricoperto cariche pubbliche, quindi sui primi due punti, e restringiamo il campo al solo Stato italiano: le voci potenzialmente da scrivere sull'enciclopedia sono un numero molto ampio, considerando il gran numero di parlamentari, sindaci e presidenti vari che si sono succeduti nella storia d'Italia tra monarchia e repubblica, e solo una piccola parte di esse esistono già su Wikipedia. Il lavoro è sicuramente lungo e non sostenibile da un solo utente, mentre relativamente più semplice è il trasferire i dati su quei politici nel grande archivio di Wikidata. Il primo passo potrebbe essere quindi quello di creare per tutti loro un elemento su Wikidata e per farlo iniziare dall'importazione automatica dei dati. Per i parlamentari il problema è facilmente risolto: esistono già online l'archivio storico della Camera dei deputati e quello del Senato con tutte le schede di deputati e senatori sia del Regno d'Italia che della Repubblica; con un bot si potrebbe creare in automatico ogni elemento sfruttando l'identificativo del rispettivo archivio storico. Più difficile la creazione di elementi per quei politici che non sono registrati in alcun archivio online, quali presidenti di regione, di provincia e sindaci delle città capoluogo. Esiste in effetti il sito del Ministero dell'Interno, ma possiede una mole di informazioni troppo ampia (tutti i sindaci di tutti gli 8000 comuni italiani) e soprattutto il database permette di risalire agli amministratori solamente a partire dal 1985. Per tutti coloro che hanno amministrato prima di quella data non esiste alcun archivio. Qui l'automaticità è quindi impossibile ed è necessario l'inserimento manuale: in molti casi, soprattutto per i sindaci, è necessario armarsi di pazienza ed intraprendere una ricerca più attenta. Porto un caso specifico, che mi riguarda da vicino: la serie dei sindaci della mia città. Come già ho avuto modo di dire, molti dei miei contributi riguardano storia, cultura e geografia della città e del territorio provinciale di Grosseto. In vista di un completamento per quelle serie di voci almanaccabili secondo le regole di Wikipedia, come abbiamo visto, quali sono gli amministratori delle città capoluogo, mi sono da subito interessato a ricercare e condividere le notizie circa i sindaci della mia città. Nel farlo, mi sono reso conto però che non esiste alcun elenco ufficiale di sindaci di Grosseto, né esso è riportato in qualche pubblicazione di carattere storico-politico; anche chiedendo direttamente in Comune presso l'archivio comunale, tutto ciò che ho ricevuto è stata una tapina alzata di spalle: l'unica lista che esiste è dal 1946 in poi, prima di allora il buio completo. Sui podestà fascisti non si riesce a trovare niente di niente, sui sindaci post-unitari solo qualche menzione sporadica in vari testi senza che fosse possibile ricostruire una successione sicura con date certe. In mancanza di fonti terze, ho dovuto procedere a ricostruire la lista io stesso, consultando quotidiani locali (*Ombrone, Etruria Nuova, Rinnovamento*) e svariati atti e documenti conservati all'Archivio di Stato di Grosseto. Al termine della ricerca la cronotassi dei sindaci di Grosseto era completata: quattordici amministratori si sono succeduti alla guida del comune tra il 1859 e il 1943, escludendo prosindaci, facenti funzioni, commissari regi e prefettizi. Considerando che per gli otto primi cittadini della Grosseto repubblicana quattro elementi già erano presenti perché creati da altri utenti (Pollini, Antichi, Bonifazi, Vivarelli Colonna), ne rimanevano quattro da aggiungere ai sopraccitati quattordici: in tutto, diciotto nuovi elementi da creare in Wikidata. Prima ho creato gli elementi di quei sindaci per i quali potevo scrivere anche la voce di Wikipedia in quanto in possesso di notizie biografiche: tra le più accurate, quelle di Angelo Ferri e Benedetto Ponticelli, entrambi lemmati nel *Dizionario del Risorgimento Nazionale* di Michele Rosi, e quella di Giovanni Pizzetti, biografato dal Sistema

Informativo Unificato delle Soprintendenze Archivistiche che ha censito il suo archivio; inoltre, occorre segnalare che due di essi (Ferri e Tattarini) avrebbero già potuto avere un elemento Wikidata in quanto parlamentari, se solo fosse stato realizzato il progetto del bot per l'importazione automatica dagli archivi storici della Camera e del Senato.

3.3. Un esempio specifico su Wikidata: il sindaco Ippolito Luciani

Prendiamo come esempio una delle personalità politiche compresa nell'elenco precedente e della quale ancora non esiste la voce biografica su Wikipedia: Ippolito Luciani.¹⁶ Luciani fu sindaco di Grosseto per due mandati tra il 1880 e il 1886 (il mandato era allora triennale e la carica non elettiva bensì di nomina regia). Il primo passo per creare l'elemento è quello di richiedere un nuovo identificativo per il soggetto. Selezionando l'opzione del portale Wikidata “crea un nuovo elemento” (<https://www.wikidata.org/wiki/Special:NewItem>) l'identificativo viene assegnato in automatico: nel nostro caso l'identificativo è Q56230212. Si procede quindi a stabilire un'etichetta (*label*), che sarà il nome del soggetto “Ippolito Luciani”, e una breve e concisa descrizione (*description*), nel nostro caso “politico ed ingegnere italiano”, in quanto Luciani, oltre che per essere stato sindaco, si distinse anche per il suo lavoro di ingegnere civile: tali etichetta e descrizione sono valide per la stringa inerente la lingua italiana, ma si possono inserire etichette e descrizioni per ogni lingua presente nel database; l'etichetta in questo caso, in quanto nome proprio di persona, non è soggetta a variazioni, mentre cambierà invece la descrizione (“*Italian politician and engineer*” in inglese, “*homme politique et ingénieur italien*” in francese, “*político e ingeniero italiano*” in spagnolo, ecc.); non sono necessari *alias*, in quanto il soggetto non è conosciuto o citato con altri nomi. Dopodiché si passa alle dichiarazioni (*statements*), ovvero ai dati riguardanti l'elemento, consistenti in una asserzione (*claim*), ovvero una coppia proprietà-valore: opzionali sono i qualificatori e i riferimenti che approfondiscono oppure giustificano tale asserzione. La prima asserzione sarà quella che definirà prima di tutto in quale classe, categoria o gruppo di oggetti tale elemento è compreso: alla proprietà “istanza di” (*instance of*) colleghiamo quindi il valore “umano” (*human*) che stabilisce che stiamo parlando di una persona; poi passiamo a stabilirne il “sesso o genere” (*sex or gender*) assegnando il valore “maschio” (*male*): in questo caso fare attenzione a collegare il valore corretto, cioè l'elemento “maschio” con identificativo Q6581097 da utilizzarsi per esseri umani, non l'elemento “maschio” con identificativo Q44148 che invece si riferisce agli animali. Altre dichiarazioni saranno poi inerenti ai suoi dati biografici: “luogo di nascita” (“Grosseto”), “data di nascita” (“3 ott 1841”), “luogo di morte” (“Castel del Piano”), “data di morte” (“9 ago 1895”) e “paese di cittadinanza”, in questo caso “Regno d'Italia (1861-1946)”. Un'opzione aggiuntiva può essere dichiarare eventuali parentele: Ippolito fu fratello di Giuseppe Luciani, anch'egli ingegnere e noto per i suoi edifici di gusto purista tardo-ottocentesco con qualche incursione nell'eclettismo e nel liberty realizzati a Grosseto; con la proprietà “fratello/sorella” (*sibling*) lo si collega in questo modo all'elemento del fratello Giuseppe (che dev'essere aggiunta in entrambi gli elementi). Poi “occupazione” (*occupation*): nel nostro caso “politico” (*politician*) e “ingegnere civile” (*civil engineer*). Si passa poi a stabilire la carica ricoperta dal soggetto: “carica ricoperta” (*position held*), quella di “sindaco” (con identificativo Q670106 proprio per gli amministratori dei comuni italiani). In questo caso vanno inseriti anche i qualificatori (*qualifier*), che vanno ad approfondire l'asserzione: si inserisce quindi una “data di inizio” mandato (“1880”), una “data di fine” (“1886”), il qualificatore “di” che stabilisce che è stato sindaco del comune di “Grosseto”, e infine lo si inserisce in una successione con le proprietà “ha sostituito” (“Ippolito Andreini”) e “sostituito da” (“Benedetto Ponticelli”), collegandolo così anche agli elementi degli altri sindaci grossetani. Non esistendo ancora voci su Wikipedia (in qualsiasi lingua) o pagine in altri progetti Wikimedia

¹⁶ <https://www.wikidata.org/wiki/Q56230212>, *Ippolito Luciani*, su [wikidata.org](https://www.wikidata.org/)

riferiti al soggetto, tali collegamenti in questo caso non saranno inseriti.

La pagina dell'elemento la si può giudicare completata nelle sue principali dichiarazioni: eventuali approfondimenti potranno essere aggiunti in seguito, ma i dati basilari sono tutti presenti e l'elemento è sufficientemente collegato con altri già esistenti nel database. A questo punto si può ricercare se tale soggetto è presente anche in altri archivi, aggiungendo così gli identificativi con i quali esso è classificato. Nel nostro caso, la voce di autorità "Ippolito Luciani" risulta presente nel database del Sistema Bibliotecario Nazionale, in quanto autore di pubblicazioni di carattere ingegneristico, con l'identificativo IT\ICCU\CUBV\094894; ciò potrà essere aggiunto alla pagina su Wikidata per creare un collegamento anche con lo SBN; allo stesso modo si può inserire anche l'identificativo del VIAF (307383175), il *Virtual International Authority File*, progetto internazionale dell'*Online Computer Library Center* che costituisce un database di voci di autorità controllate provenienti da vari cataloghi nazionali.

4. Bibliografia

TIM BERNERS-LEE, *L'architettura del nuovo Web*, Milano, Feltrinelli, 2001.

WERNER LEYH, HOMERO FONSECA FILHO, *Interlinking Standardized OpenStreetMap Data and Citizen Science Data in the OpenData Cloud*, in ISABEL L. NUNES, *Advances in Human Factors and Systems Interaction*, Springer International Publishing, 2018, pp. 85-96.

ANDREW LIH, *La rivoluzione di Wikipedia. Come un gruppo di illustri sconosciuti ha creato la più grande enciclopedia del mondo*, Torino, Codice Edizioni, 2010.

DON TAPSCOTT, ANTHONY D. WILLIAMS, *Wikinomics 2.0. La collaborazione di massa che sta cambiando il mondo*, Milano, Rizzoli, 2008.

5. Sitografia

<https://www.wikimedia.it/la-wiki-maratona-record-degli-archivi-158-voci-wikipedia-nuove-arricchite/>

<https://www.wikidata.org>

<https://www.wikidata.org/wiki/Wikidata:Glossary/it>

<https://www.wikipedia.org>