

Luca Martinelli¹

Wikisource, la ‘wikibiblioteca’ collaborativa e multilingue

Wikisource è uno dei dodici progetti ufficiali della Wikimedia Foundation, la fondazione non-profit di diritto statunitense che gestisce, fra gli altri, l’enciclopedia libera Wikipedia. Lo scopo principale di *Wikisource* è raccogliere le versioni digitalizzate di testi, documenti e libri di ogni tipologia, epoca o argomento, purché in pubblico dominio o pubblicati con una licenza libera. Il progetto è attualmente suddiviso in 65 differenti versioni linguistiche, fra cui ovviamente quella italiana, che conta oltre 105.000 testi.

Qualche cenno storico

Il progetto nacque ufficialmente il 24 novembre 2003² con il nome di ‘Project Sourceberg’, un gioco di parole fra ‘iceberg’ e ‘Project Gutenberg’, il capostipite di tutte le biblioteche libere *online*³. Lo scopo originale del progetto era quello di raccogliere tutti i testi storici importanti o utili alla stesura delle voci di Wikipedia. Nelle

¹ Wikimedia Italia, Email: luca.martinelli@wikimedia.it.

² Fino a quella data, tutti i testi storici importanti o utili alla stesura delle voci di Wikipedia venivano temporaneamente raccolti all’indirizzo <<http://ps.wikipedia.org/>>, dove il prefisso ‘p’ stava per ‘primary source’ (fonte primaria) o per ‘Project Sourceberg’. Successivamente, per evitare di occupare ‘abusivamente’ il sotto-dominio relativo alla lingua pashto, i testi furono spostati all’indirizzo <<http://sources.wikipedia.org/>> (ultimo accesso 22.06.2017).

³ L’indirizzo ufficiale è <<http://www.gutenberg.org/>> (ultimo accesso 22.06.2017).

parole di The Cunctator, l'utente che propose la creazione del progetto, «we don't want to try to duplicate Project Gutenberg's efforts; rather, we want to complement them»⁴.

Il 6 dicembre 2003 il progetto viene ribattezzato 'Wikisource', ma viene mantenuto il logo che lo contraddistingueva: un iceberg, a simboleggiare che i libri esistenti in versione digitale all'epoca fossero, per l'appunto, solo la punta dell'iceberg di tutti i libri pubblicati nella storia. La versione italiana nacque soltanto due anni più tardi, nel 2005.

Le regole di Wikisource

Wikisource, come tutti i progetti Wikimedia, è un progetto a cui chiunque può registrarsi e contribuire liberamente e in modo totalmente gratuito. Tuttavia, la pubblicazione dei testi deve seguire un certo numero di regole, che si possono così riassumere:

1. tutti i testi presenti devono essere già editi, ossia già pubblicati su supporto cartaceo⁵;
2. tutti i testi presenti devono essere in pubblico dominio oppure essere pubblicati con una licenza libera compatibile con la CC BY-SA⁶;
3. tutti i testi presenti devono essere identici all'edizione cartacea di riferimento.

⁴ «Non vogliamo cercare di duplicare gli sforzi del Progetto Gutenberg, piuttosto vogliamo essere complementari a esso». Link all'intervento ufficiale: <https://nostalgia.wikipedia.org/w/index.php?title=Wikipedia_commentary/Project_Soourceberg&direction=prev&oldid=54333> (ultimo accesso 22.06.2017).

⁵ Per quello che riguarda le pubblicazioni *digital-born*, sono ammesse solo se provengono da siti ufficiali o da editori in linea riconosciuti (cioè che hanno un obbligo di deposito legale presso la biblioteca nazionale del loro Paese).

⁶ La licenza *Creative Commons* Attribuzione-Condividi allo stesso modo o CC BY-SA è una licenza d'uso di un'opera che permette il suo libero riutilizzo in ogni forma e per ogni scopo, anche commerciale, fermo restando il rispetto di due semplici clausole: che sia citato l'autore o l'autrice originale dell'opera (Attribuzione) e che l'eventuale opera derivata sia rilasciata con la stessa licenza (Condividi allo stesso modo).

Cerchiamo adesso di analizzare più nello specifico queste regole generali. Le tipologie di testi che si possono pubblicare su *Wikisource* sono⁷:

- libri, romanzi, novelle, poemi, poesie e altri testi letterari editi;
- fonti primarie editate (diari, registrazioni, lettere, opere scientifiche, verbalizzazioni ufficiali, documenti legali, ecc.);
- documenti storici e discorsi di interesse nazionale o internazionale editi;
- canti popolari, dialettali⁸ e inni nazionali;
- traduzioni italiane editate e libere da copyright di opere in lingua straniera;
- tesi o dissertazioni approvate e pubblicate da una istituzione accademica riconosciuta.

L'insistenza della comunità di *Wikisource* riguardo la presenza di una pubblicazione 'ufficiale' per i testi presenti sul progetto si spiega con il principio del punto di vista neutrale, che è un'altra delle regole incontestabili e fondamentali di tutti i progetti Wikimedia. Per quel che riguarda *Wikisource*, questo viene declinato in due modalità fondamentali:

1. nel riprodurre i testi fedelmente e integralmente, senza enfatizzare od omettere parti del testo e mettendo a disposizione dell'utente una 'versione a fronte' originale;
2. nell'evitare commenti personali, recensioni o critiche stilistiche, letterarie, giuridiche, politiche o quant'altro da parte degli utenti.

Ovviamente tutti i testi sopra elencati devono essere fuori dall'ambito di applicazione del *copyright*, oppure essere pubblicati

⁷ Per maggiori informazioni, vedi <https://it.wikisource.org/wiki/Wikisource:Cosa_mettere_su_Wikisource> (ultimo accesso 22.06.2017).

⁸ Per i canti popolari e in dialetto, è prevista una deroga *sui generis* alla regola della pubblicazione edita. Tuttavia, resta preferibile indicare una fonte, possibilmente cartacea, che garantisca sull'affidabilità del testo.

con una licenza libera. Di norma, quindi, sono accettati su *Wikisource* tutti i materiali in pubblico dominio sia in Italia, sia negli Stati Uniti (dove risiedono i *server* della Wikimedia Foundation), ossia tutti i materiali il cui autore è morto da più di 70 anni e la cui prima pubblicazione è avvenuta da più di 95 anni.

Entrambe le condizioni sono necessarie alla pubblicazione su *Wikisource*. Qualora una delle due condizioni dovesse venire a mancare, il testo risulterebbe ancora coperto dal *copyright* secondo la legge italiana o statunitense, dunque non sarebbe possibile pubblicarlo⁹.

Nel caso in cui un'opera sia stata pubblicata in più versioni differenti – come per esempio *I promessi sposi* di Alessandro Manzoni o *La Traviata* di Giuseppe Verdi – oppure esistano differenti traduzioni della stessa opera – come per esempio il *Macbeth* di William Shakespeare o *Il vento fra i salici* di Kenneth Grahame – è possibile riportare tutte queste versioni, a patto che rispettino le regole sul *copyright* sopra esposte. In questo caso, verrà creata una pagina 'Opera' dove si presenterà brevemente la differenza fra le varie versioni, rinviando poi alle varie edizioni presenti sul progetto.

Inoltre, essendo *Wikisource* una biblioteca libera, le linee guida fanno espresso divieto agli utenti di pubblicare¹⁰:

- testi coperti da diritto d'autore;
- opere originali degli utenti non pubblicate altrove su carta;
- testi anonimi¹¹ o traduzioni anonime¹²;
- commenti ai testi scritti dagli utenti;

⁹ Per maggiori informazioni, vedi <https://it.wikisource.org/wiki/Wikisource:Rispetto_del_copyright> (ultimo accesso 22.06.2017).

¹⁰ Per maggiori informazioni, vedi <https://it.wikisource.org/wiki/Wikisource:Cosa_non_mettere_su_Wikisource> (ultimo accesso 22.06.2017).

¹¹ I testi anonimi sono ammessi soltanto se di comprovata valenza storica.

¹² In realtà, pochissime versioni linguistiche (fra cui quella inglese) ammettono la possibilità di pubblicare traduzioni realizzate dagli utenti. La versione italiana, tuttavia, mantiene espressamente questo divieto, come indicato alla pagina <https://it.wikisource.org/wiki/Wikisource:Niente_traduzioni originali> (ultimo accesso 22.06.2017).

- materiale puramente numerico oppure compilatorio;
- testi in evoluzione;
- testi pubblicitari o *curriculum vitae*.

Il caricamento dei testi

Il procedimento di caricamento e riletture dei testi è piuttosto lungo e complesso¹³, quindi si affronterà in questa sede un'analisi 'semplificata', seguendo come esempio il testo del 1574 *Trattato della neve e del bere fresco*, originariamente scritto dal medico spagnolo Nicolas Monardes e tradotto in italiano da Giovan Battista Scarampo¹⁴.

Come già accennato, è bene che il testo che si intende caricare su *Wikisource* sia già disponibile in forma scansionata, che il file relativo sia in un formato aperto, come il DjVu¹⁵, e che sia corredato di tutti i metadati necessari (autore, titolo, sottotitolo, editore, lingua di pubblicazione, fonte da cui proviene il file, ecc.).

Il file DjVu deve essere caricato, ovviamente con l'apposita descrizione¹⁶, su Wikimedia Commons¹⁷, l'archivio di oltre 32,6 milioni di immagini, suoni e altri file multimediali della Wikimedia Foundation. In questo modo, il file può essere direttamente richiamato su *Wikisource* – e, più in generale, sugli oltre 800 progetti Wikimedia.

¹³ È disponibile una guida sistematica ai singoli passaggi all'indirizzo <https://it.wikisource.org/wiki/Aiuto:Guida_alla_pubblicazione_di_un_testo> (ultimo accesso 22.06.2017).

¹⁴ Il file è disponibile all'indirizzo <https://commons.wikimedia.org/wiki/File:Trattato_della_neve_e_del_bere_fresco.djvu> (ultimo accesso 22.06.2017).

¹⁵ Per maggiori informazioni, vedi <<https://it.wikipedia.org/wiki/DjVu>> (ultimo accesso 22.06.2017).

¹⁶ Per i libri scansionati, è preferibile utilizzare il *template* <<https://commons.wikimedia.org/wiki/Template:Book>> (ultimo accesso 22.06.2017).

¹⁷ L'indirizzo ufficiale è <https://commons.wikimedia.org/wiki/Main_Page> (ultimo accesso 22.06.2017).

Una volta su *Wikisource*, bisogna creare la pagina ‘Indice’¹⁸. Un ‘indice’ è il punto di partenza per le successive operazioni di rilettura, dove è possibile indicare le informazioni di base sull’opera, la sua eventuale ripartizione in capitoli, la numerazione delle pagine, e così via¹⁹. Una volta completate queste operazioni preliminari, si può passare alla parte più importante del lavoro: la rilettura del testo.

La rilettura dei testi: la ‘versione con testo a fronte’

Cliccando su uno qualsiasi dei link rossi della pagina ‘indice’ che compaiono a destra del frontespizio, si viene condotti alla relativa pagina e dunque all’interfaccia di modifica della pagina. Da un lato, avrai la scansione della pagina e dall’altro l’*editor* di modifica: se nell’*editor* è già disponibile del testo, vuol dire che è stato già effettuato l’OCR e che, dunque, bisogna soltanto rileggere e formattare il testo originale; se, invece, l’*editor* è vuoto, dovrai essere tu a effettuare la trascrizione.

Come detto prima, su *Wikisource* la regola è quella di riprodurre i testi fedelmente e integralmente: ciò significa che, secondo le convenzioni di trascrizione del progetto²⁰, l’ortografia e la formattazione originali del testo vanno mantenute esattamente così come sono, tranne che in rarissimi casi (come la ‘s lunga’ o ‘f’, normalizzabile in ‘s’, o la ‘doppia s’ o ‘ß’, normalizzabile in ‘ss’). Queste norme si applicano a tutto il testo, ivi comprese le intestazioni, i piè di pagina, le note, ecc.

¹⁸ Le indicazioni precise su come creare un indice sono disponibili all’indirizzo <https://it.wikisource.org/wiki/Aiuto:Guida_alla_pubblicazione_di_un_testo/Creare_la_pagina_indice> (ultimo accesso 22.06.2017).

¹⁹ Un esempio di indice lo si trova all’indirizzo <https://it.wikisource.org/wiki/Indice:Trattato_della_neve_e_del_bere_fresco.djvu> (ultimo accesso 22.06.2017).

²⁰ Le indicazioni precise sono disponibili all’indirizzo <https://it.wikisource.org/wiki/Wikisource:Convenzioni_di_trascrizione> (ultimo accesso 22.06.2017).

Ovviamente, la formattazione è la parte più complessa della rilettura. Per questo, i wikisourciani hanno creato una serie di *template*²¹ e di altri strumenti automatici²² per venire incontro agli utenti che si affacciano a questo strano mondo di amanuensi. Fra questi, è importante citare i *template* TestoCitato²³ e AutoreCitato²⁴: entrambi servono a generare dei link rispettivamente ad altri testi o a eventuali personalità citate all'interno del testo che si sta rileggendo. Questa è l'unica, vera deroga al principio della riproduzione fedele del testo, per permettere al lettore di navigare all'interno della biblioteca o di approfondire determinati rimandi.

Un'altra, sebbene minore, deroga alla riproduzione fedele del testo riguarda gli eventuali errori presenti nel testo:

1. se il libro presenta una sezione di *errata corrige*, è possibile inserire direttamente nel testo la versione corretta tramite il *template* 'Ec'²⁵ in questo modo: `{{Ec|<testo sbagliato>|<testo corretto secondo le indicazioni dell'autore>}}`; in questo modo, l'errore resta tale e quale nella versione con testo a fronte, ma viene mostrato nella versione esclusivamente testuale (a cui faremo riferimento dopo);
2. se il libro non presenta una sezione di *errata corrige*, ma c'è un errore di stampa inequivocabile (p. es. Napoleone Bonaparte viene erroneamente chiamato Bonapare), allora è possibile utilizzare il *template* Ec – fermo restando

²¹ I *template* fondamentali sono disponibili all'indirizzo https://it.wikisource.org/wiki/Categoria:Template_strumenti_di_layout_per_i_testi (ultimo accesso 22.06.2017).

²² Gli strumenti d'ausilio sono disponibili all'indirizzo https://it.wikisource.org/wiki/Aiuto:Strumenti_per_la_rilettura (ultimo accesso 22.06.2017).

²³ Disponibile all'indirizzo <https://it.wikisource.org/wiki/Template:TestoCitato> (ultimo accesso 22.06.2017).

²⁴ Disponibile all'indirizzo <https://it.wikisource.org/wiki/Template:AutoreCitato> (ultimo accesso 22.06.2017).

²⁵ Disponibile all'indirizzo <https://it.wikisource.org/wiki/Template:Ec> (ultimo accesso 22.06.2017).

che, in caso di dubbio, è sempre preferibile lasciare il testo così com'è.

Lo Stato di Avanzamento del Lavoro su Wikisource

La durata della fase di trascrizione e rilettura dipende, ovviamente, dalla lunghezza e complessità del testo da rileggere e da una serie di altri fattori, come la qualità della scansione, la presenza di illustrazioni (anch'esse da riportare così come sono nella versione digitale) e la necessità di impiegare *template* 'speciali' per la formattazione.

Per indicare lo stato in cui ci si trova, anche *Wikisource* ha adottato il metodo dello Stato di Avanzamento del Lavoro (SAL), sebbene qui le indicazioni assumono un significato lievemente differente²⁶.

- SAL 25% = OCR creato automaticamente;
- SAL 50% = Testo riletto da un utente;
- SAL 75% = Testo riletto e formattato da un utente;
- SAL 100% = Testo riletto da un secondo utente.

Va notato come l'utente che genera per primo la pagina riletta possa portare il SAL soltanto fino al 75%, dal momento che, secondo le regole di *Wikisource*, perché una pagina possa essere portata al 100% serve che sia riletta e controllata da due diversi utenti.

Esiste poi un livello SAL 0%, che viene utilizzato per indicare le pagine vuote o con testi o immagini non riconducibili all'opera (come il timbro della biblioteca, annotazioni a penna o matita o simili).

²⁶ Le indicazioni precise sono disponibili all'indirizzo <https://it.wikisource.org/wiki/Aiuto:Stato_di_Avanzamento_del_Lavoro> (ultimo accesso 22.06.2017).

Il lavoro (quasi) finito: la 'versione testuale' e la seconda rilettura

Una volta completata la prima rilettura del testo, si può passare a creare la 'versione testuale'²⁷, ossia una versione del testo digitalizzato priva della versione a fronte. Questo passaggio è ormai quasi completamente automatizzato e dipende, tuttavia, dalla buona compilazione della pagina di indice.

Per generare l'intestazione e le pagine per i singoli capitoli, basta cliccare sui link rossi che sono presenti sulla colonna destra della pagina di indice: il contenuto della pagina viene generato in automatico dal sistema e basterà alla fine cliccare su 'Salva la pagina' per crearla²⁸.

In questo modo, il testo riletto della versione con testo a fronte viene automaticamente importato nella versione testuale, così che ogni modifica fatta alla prima viene automaticamente mostrata nella seconda. Inoltre, come detto poco prima, la versione testuale mostra in automatico anche tutti gli *errata corrige* dell'autore (se inseriti) o le correzioni degli utenti. La versione testuale, così come riportata, è consultabile liberamente *online* ed è inoltre gratuitamente scaricabile in formato EPUB, MOBI, PDF, RTF e TXT.

Wikisource e il 'mondo reale': collaborazioni e attività collegate

Le attività della comunità di *Wikisource* non si fermano al (di per sé impegnativo) lavoro di rilettura *online*, ma proseguono *offline* anche con la collaborazione dell'Associazione Wikimedia Italia, che promuove i progetti Wikimedia di lingua italiana sul territorio nazionale.

²⁷ Le indicazioni precise sono disponibili all'indirizzo <https://it.wikisource.org/wiki/Aiuto:Guida_alla_pubblicazione_di_un_testo/Creare_la_versione_testuale> (ultimo accesso 22.06.2017).

²⁸ Un esempio di versione testuale lo si trova all'indirizzo <https://it.wikisource.org/wiki/Trattato_della_neve_e_del_bere_fresco> (ultimo accesso 22.06.2017).

Un esempio di collaborazione sono le donazioni di testi già digitalizzati provenienti da varie istituzioni, come la BEIC, la Biblioteca Nazionale Centrale di Firenze, la Biblioteca La Vigna, la Biblioteca comunale di Trento, il MUSE e altre. La donazione più importante, in termini numerici, è quella proveniente dalla Biblioteca Universitaria Alessandrina di Roma (circa 200 esemplari in italiano, latino e spagnolo, provenienti dal Fondo urbinato), che sarà completata entro la fine di quest'anno.

L'Associazione Wikimedia Italia promuove anche la realizzazione di cosiddetti 'biblio-hackathon', ossia delle mini-maratone di rilettura condotte con l'ausilio di utenti di *Wikisource* e aperte a chiunque abbia intenzione di mettersi alla prova con il mondo wikimediano e... sappia leggere! L'associazione sta inoltre portando avanti un progetto di Servizio civile nazionale con la Biblioteca comunale di Trento, finalizzato alla digitalizzazione e trascrizione di testi in pubblico dominio in possesso della biblioteca trentina.

Infine, a partire dal 2013, Wikimedia Italia organizza ogni anno un concorso di rilettura *online* che si tiene in concomitanza con il compleanno di *Wikisource*. Il concorso si concentra su un certo numero di testi 'validi' (romanzi, saggi, testi teatrali, raccolte di poesie, trattati filosofici e così via), selezionati dalla comunità in base allo Stato di Avanzamento del Lavoro e in base alla relativa semplicità di rilettura. I premi previsti sono tre e sono dei buoni acquisto non cumulabili di 50, 30 e 20 euro rispettivamente da spendere presso librerie *online*.

L'anno scorso sono stati assegnati ulteriori tre premi all'interno di un '*mini-contest*' speciale, dedicato ai primi testi provenienti dalla donazione della Biblioteca Alessandrina.

Sì, ok, ma... chi ve lo fa fare?

Questo è il punto più 'difficile' da spiegare, perché è qualcosa di molto personale per noi wikimediani.

Ai wikimediani piace la conoscenza, piace conoscere nuove cose e leggere è, da sempre, uno dei modi migliori per conoscere cose nuove; ma per un wikimediano, leggere e imparare non basta. Un wikimediano vuole condividere la propria conoscenza, vuole garantire ad altri di avere accesso alle stesse fonti a cui lei o lui ha avuto accesso, perché fondamentalmente è un *geek*.

Il comico britannico Simon Pegg disse: «Being a geek is all about being honest about what you enjoy and not being afraid to demonstrate that affection. It means never having to play it cool about how much you like something. It's basically a license to proudly emote on a somewhat childish level rather than behave like a supposed adult. Being a geek is extremely liberating»²⁹. In un certo senso, questa è davvero la passione che ogni wikimediano ha nei confronti della conoscenza.

Per noi, quindi, è stato bellissimo aver incontrato la comunità dei bibliotecari, perché abbiamo scoperto di condividere la passione per la conoscenza con voi, abbiamo scoperto che anche voi siete – giustamente – un po' *geek* come noi. Non a caso, sono anni che Wikimedia Italia collabora con il vostro mondo, al punto di aver trovato molto, moltissimo terreno fertile e di aver convinto molti bibliotecari ad affacciarsi con sempre più coraggio al nostro strano mondo, «un mondo in cui ciascuno possa avere libero accesso a tutto il patrimonio della conoscenza umana».

²⁹ «Essere un *geek* significa essere onesti riguardo ciò che ti piace e non aver paura di mostrare quanto ti piaccia. Significa non doversi moderare nel dire quanto davvero ti piaccia quella cosa. È come avere il permesso di farsi orgogliosamente prendere dall'emozione quasi in maniera infantile, piuttosto che comportarsi come un adulto normale. Essere *geek* è profondamente liberatorio».