

Filosofia e *open access*: l'archivio digitale *Lessici filosofici dell'età moderna*

di Annarita Liburdi e Ada Russo

1. Premessa

L'Istituto per il lessico intellettuale europeo e storia delle idee (ILIESI) del Consiglio nazionale delle ricerche è stato formalmente istituito il 12 febbraio 2001¹. L'Istituto è orientato in molteplici direzioni, muovendo tuttavia da un assunto metodologico unitario: la storia della filosofia e della scienza – o, in una prospettiva ancora più ampia, delle idee – può ricostruirsi in modo particolarmente rigoroso indagando il mutevole complesso degli strumenti espressivi di cui queste discipline, nel corso dei secoli, si sono dotate e rimanendo saldamente ancorata alla testimonianza dell'espressione scritta. Lo studio della terminologia di cultura nei suoi tecnicismi e nella sua polisemia, nei nessi sincronici e diacronici, nelle traduzioni e trasposizioni di elementi dall'uno all'altro contesto lessicale, nel lavoro di elaborazione e ricerca su *corpora* filosofici d'autore, obbliga a una lettura più analitica dei testi e permette di cogliere percorsi e rapporti talora non evidenti. Di qui le ricerche su termini e famiglie di termini, l'edizione di lessici, indici e concordanze, la costituzione di banche dati, i seminari e i colloqui internazionali su problemi metodologici e storico - linguistici, nonché l'elaborazione di programmi per un più complesso trattamento informatico dei testi e della terminologia di cultura, con particolare attenzione all'unificazione dei codici².

L'ILIESI è uno degli istituti del CNR³ ed è naturalmente finanziato con fondi pubblici per la ricerca, per questo motivo intende mettere a disposizione della comu-

ANNARITA LIBURDI, Istituto per il lessico intellettuale europeo e storia delle idee - CNR, Biblioteca, via Nomentana, 118, 00161, Roma, e-mail annarita.liburdi@iliesi.cnr.it.

ADA RUSSO, Technical collaborator, Istituto per il lessico intellettuale europeo e storia delle idee - CNR, via Nomentana 118, 00161 Roma, e-mail ada.russo@iliesi.cnr.it.

Ultima consultazione siti web: aprile 2009.

1 L'Istituto nasce dall'accorpamento dei due Centri di studio operanti presso la Facoltà di lettere e filosofia dell'Università La Sapienza di Roma: il Centro di studio per il lessico intellettuale europeo (LIE), fondato e diretto da Tullio Gregory e il Centro di studio del pensiero antico (PA), fondato e diretto da Gabriele Giannantoni fino alla sua scomparsa nel 1998. È nato così l'ILIESI, che storicamente prosegue le attività del LIE unitamente al PA, sotto la direzione di Tullio Gregory fino al novembre 2007.

2 Annarita Liburdi, *Il Lessico intellettuale europeo dal 2001 al 2006: da Centro di studio a Istituto*, Roma: Lessico intellettuale europeo, 2007.

3 Il Consiglio nazionale delle ricerche (CNR) è Ente pubblico nazionale con il compito di svolgere, promuovere, diffondere, trasferire e valorizzare attività di ricerca nei principali settori di sviluppo delle conoscenze e delle loro applicazioni per lo sviluppo scientifico, tecnologico, economico e sociale del Paese.

nità scientifica i risultati e i prodotti della propria ricerca a libero accesso. L'obiettivo è quello di aumentare sia la diffusione che la visibilità delle proprie attività di ricerca e, dal momento che la disseminazione della conoscenza è incompleta se l'informazione non è resa largamente e prontamente disponibile alla società⁴, l'Istituto, attraverso la pubblicazione in archivi aperti, intende aumentare l'impatto della letteratura scientifica prodotta. Solo una reale ed efficace disseminazione dell'informazione garantisce un reale impatto: più un articolo è liberamente scaricabile, più è letto, più viene citato. Questo favorisce la condivisione del sapere e quindi un più rapido avanzamento della conoscenza, senza barriere, in tutto il mondo. Con la pubblicazione su rivista è il ricercatore che guadagna in visibilità mentre con la pubblicazione *open access* è l'istituzione che viene messa in primo piano. Anche se queste consapevolezza sembrano oramai inconfutabili in campo scientifico, sono meno evidenti in campo umanistico ed il CNR, ente generalista che aggrega istituti di entrambi gli ambiti disciplinari, non fa ancora parte né delle 300 istituzioni accademiche e di ricerca che hanno aderito alla Dichiarazione di Berlino sull'accesso aperto alla letteratura scientifica elaborato dal Max Planck Gesellschaft nel 2003, né dei 74 atenei italiani che hanno aderito in seguito alla Dichiarazione di Messina dell'ottobre 2004⁵. Consapevole della necessità di rendere visibili e in qualche modo comprensibili le proprie attività di ricerca, l'ILIESI ha scelto quindi di avvicinarsi al mondo dell'*open access*.

2. L'archivio digitale *Lessici filosofici dell'età moderna*

L'archivio digitale *Lessici filosofici dell'età moderna* è stato finanziato negli anni 2007 e 2008 sulla base della l. 10 gennaio 2000, n. 6, intesa a favorire le iniziative per la promozione e il potenziamento delle istituzioni impegnate nella diffusione della cultura scientifica e tecnologica. Il progetto rappresenta un primo passo verso l'accesso aperto; l'archivio infatti è una delle banche dati dell'ILIESI, il responsabile scientifico del progetto è Eugenio Canone, con la collaborazione scientifica di Delfina Giovannozzi, Hansmichael Hohenegger, Annarita Liburdi, Margherita Palumbo, Ada Russo, Marco Sgarbi; mentre lo sviluppo Web e l'elaborazione dei dati è di Simona Lampidecchia. Per la realizzazione dell'archivio, al gruppo dell'ILIESI si sono affiancati alcuni collaboratori esterni soprattutto per la parte revisione e controllo. L'archivio è una biblioteca digitale che attualmente già mette a disposizione di studenti e studiosi una ventina di dizionari filosofici – sarebbe corretto dire dizionari filosofico-scientifici e di erudizione – pubblicati in latino tra la fine del XVI secolo e gli inizi del XVIII secolo⁶. Il progetto prevede inoltre la registrazione della nomenclatura dei dizionari: di tutte le entrate e, in parte, delle sottoentrate, con varie possibilità di ricerca. L'archivio è poi arricchito con schede biografiche degli autori dei dizionari scelti e con schede bibliografiche relative all'edizione riprodotta. È disponibile alla consultazione anche una sezione «Materiali» che raccoglie articoli, schede sinottiche ed altro materiale di studio sui lessici pubblicati⁷.

⁴ Vincenzo Milanese, *Accesso aperto per la ricerca negli Atenei italiani: una nuova via di accesso alla conoscenza*, Press Release, <<http://www.aepic.it/conf/viewappendix.php?id=48&ap=1&cf=1>>.

⁵ <<http://www.openstarts.units.it/dspace/bitstream/10077/2542/1/Gargiulo.pdf>>.

⁶ Cfr. <<http://www.iliesi.cnr.it/Lessici/archivio.html>>.

⁷ In questa sezione possono essere pubblicati lavori già apparsi a stampa o lavori inediti che quindi sono disponibili per la prima volta in assoluto.

In futuro si prevede di incrementare tale archivio con alcuni importanti lessici filosofici pubblicati nelle varie lingue europee. Come strumento per ulteriori ricerche, si prevede di mettere a disposizione nell'archivio una versione digitale adattata del catalogo bibliografico dei dizionari dell'età moderna di Giorgio Tonelli, del quale è stato recentemente pubblicata una nuova versione ampliata e annotata⁸.

L'archivio digitale *Lessici filosofici dell'età moderna* si inserisce nella linea di ricerca dell'ILIESI: *Studi sul lessico filosofico europeo dall'Umanesimo al XX secolo*. Il progetto dell'archivio digitale, nella configurazione attuale, è nato nei primi anni del 2000, sulla base di precedenti ricerche che risalgono alla fine degli anni Ottanta del Novecento. Va tuttavia rimarcato che lo studio dei lessici filosofici dell'età moderna è stato promosso dal gruppo di ricerca dell'ILIESI sin dall'inizio della sua attività alla fine degli anni Sessanta, quando in Germania venivano pubblicate le prime ristampe anastatiche dei lessici di Johannes Micraelius⁹, Rudolph Goclenius¹⁰ e Étienne Chauvin,¹¹ tre dei più importanti dizionari filosofici dell'età moderna.

L'ambito della linea di ricerca *Studi sul lessico filosofico europeo dall'Umanesimo al XX secolo* si incentra sull'analisi del vocabolario intellettuale e degli strumenti di lessicografia filosofica in alcuni momenti rappresentativi tra umanesimo ed età contemporanea. Nell'ambito della linea di ricerca si è pertanto individuato un *corpus* selezionato dei più importanti dizionari filosofici e scientifici dell'età moderna. La definizione del "vocabolario intellettuale" – inteso come quella sezione del lessico di una lingua formata dal complesso dei termini che si costituiscono come il veicolo di concetti rilevanti nell'arco della storia intellettuale – come si è già visto, è uno degli obiettivi dell'ILIESI sin dai primi anni della sua costituzione. Lo strumento rilevatore del lessico permette infatti di osservare da una prospettiva privilegiata alcune tematiche centrali della riflessione filosofica e scientifica occidentale, colte nella loro plurivalenza e nella connessione con diversi ambiti di ricerca.

3. Obiettivi

Gli obiettivi del progetto Lessici filosofici dell'età moderna sono quindi i seguenti:

- 1) Costituzione di un archivio di dizionari filosofici (filosofico -scientifici e di erudizione) consultabile in rete, in modalità *open source*. Tale archivio rende accessibili una serie di testi significativi, che sono in rapporto sia con la tradizionale filosofia delle scuole (per esempio la tradizione aristotelica antica e medievale) sia con gli orientamenti filosofici più recenti: dalla scolastica cattolica e protestante alla filosofia cartesiana. Va precisato che si tratta talvolta di opere difficilmente reperibili.

8 Giorgio Tonelli, *A short-title list of subject dictionaries of the Sixteenth, Seventeenth and Eighteenth centuries*. Extended edition, revised and annotated by E. Canone and M. Palumbo, Firenze: Olschki, 2006. La prima edizione del catalogo di Tonelli apparve nel 1971 a Londra nel «Warburg institute survey». Nella introduzione, lo studioso osservava: «the connexion of dictionaries with doctrinal trends contemporary to them, and their influence on the evolution and diffusion of thought should be a basic field in the history of ideas, as well as a further contribution towards a more adequate use of this kind of source in general research» (ed. 2006 cit., p. xxxi).

9 Johannes Micraelius, *Lexicon philosophicum terminorum philosophis usitatorum*, Dusseldorf: Stern-Verlag Janssen, 1966.

10 Rudolph Goclenius (Göckel), *Lexicon philosophicum, quo tanquam clave philosophiae fores aperiuntur*, Hildesheim: Olms-Weidmann, 1964.

11 Étienne Chauvin, *Lexicon philosophicum*, Dusseldorf: Stern-Verlag Janssen, 1967.

2) Costituzione di un *thesaurus* interrelato dei lessici, con la messa a punto di un indice unificato delle entrate (lemmi, sintagmi e locuzioni) dei dizionari pubblicati.

4. La raccolta

Come si è detto, la raccolta di testi già pubblicata nell'archivio dell'ILIESI comprende un *corpus* di lessici filosofici latini apparsi tra la fine del XVI e gli inizi del XVIII secolo, che rappresentano un capitolo importante nella storia della lessicografia filosofica e testimoniano l'esigenza di codificazione e sistemazione della terminologia filosofica latina, in particolare nel Seicento. Sono già a disposizione nell'archivio diversi lessici pubblicati nel XVII secolo, quello che potremmo definire «il secolo della lessicografia filosofica latina». Tra questi, il *Thesaurus eruditionis scholasticae* di Basilius Faber, la cui prima edizione apparve a Leipzig nel 1571 (nell'archivio dell'ILIESI è presente l'edizione pubblicata nel 1696 sempre a Leipzig), i lessici di Henry Louis Chasteigner (1612; 1617), Rudolph Goclenius (1609; 1613; 1615), Georg Reeb (1629), Antoine Le Roy (1649), Johannes Micraelius (1662), Theodor Hackspan (1664), Johann Adam Scherzer (1675), Étienne Chauvin (1692) fino alla seconda edizione del dizionario di Chauvin (1713) e al *Lexicon philosophicum* di Plexiacus (1716).

Si è data priorità alla pubblicazione online del *corpus* dei lessici filosofici latini secenteschi, oltre che per la rilevanza numerica delle pubblicazioni apparse in questo secolo, perché il loro esame permette alcune interessanti osservazioni. Proprio nell'età in cui la tradizione filosofica aristotelico-scolastica viene a incrinarsi, i lessici filosofici – di cui è stato ricordato il costitutivo ritardo sulla lingua – registrano una terminologia affermatasi proprio in quell'ambito storico-culturale, rispondendo all'esigenza di formalizzazione del linguaggio tecnico su un piano in cui strutturalmente si intersecano terminologia e storia della filosofia. Tali lessici, pur non essendo in senso stretto “storici”, si misurano infatti con l'uso, le accezioni, le sfumature di significato che i termini hanno assunto nel passato, non mancando di segnalare la nascita di parole nuove. Tutto ciò negli anni in cui si afferma l'uso delle lingue nazionali moderne. Alla base di alcuni di questi dizionari – come quelli di Rudolph Goclenius e di Johannes Micraelius – c'è un'idea precisa di ordine classificatorio, in qualche modo uno schema di enciclopedia delle scienze. Il *Lexicon rationale* di Chauvin si segnala invece per l'attenzione al dibattito filosofico coevo e per la ricezione della terminologia cartesiana.

5. Le edizioni

Le edizioni sono state scelte sulla base di criteri diversi: primo fra tutti quello della loro circolazione e diffusione: non sempre le prime edizioni sono quelle che hanno maggiormente circolato all'epoca della loro pubblicazione e spesso le edizioni successive contengono ripensamenti e aggiustamenti da parte dell'autore che vale la pena tenere in considerazione. A volte, quando disponibili, sono state messe a disposizione online anche le prime edizioni per consentire confronti e studi sinottici¹². Un secondo criterio è rappresentato dalla disponibilità dei testi: si tratti nella maggior parte dei casi di edizioni in possesso dell'Istituto o a stampa o in microfilm¹³,

¹² Questo è il caso delle due edizioni del *Lexicon* di Étienne Chauvin, di cui è disponibile nella sezione materiali anche una scheda sinottica: <<http://www.iliesi.cnr.it/Lessici/Chauvin-differenze.htm>>.

¹³ La biblioteca dell'ILIESI possiede una delle più vaste raccolte di microfilm di lessici e dizionari filosofici.

oppure sulla base dei permessi che sono stati acquisiti in vario modo dalle biblioteche che sono in possesso dei testi originali.

Va notato che i dizionari filosofici, pur essendo testi solitamente poco citati, sono presenti nelle biblioteche private di autori di rilievo. Si può qui ricordare che nella sua biblioteca Gottfried Wilhelm Leibniz possedeva il *Vade mecum* di Johann A. Scherzer e il *Compendium lexicum philosophicum* di Johann Heinrich Alsted (entrambi i testi figurano anche nel *Catalogus librorum* di Alexander Gottlieb Baumgarten). Hermann Samuel Reimarus possedeva i lessici di Goclenius, Micraelius e Chauvin. Christian Wolff fa spesso riferimento nei suoi scritti ai lessici latino e greco di Goclenius. Immanuel Kant possedeva il *Thesaurus eruditionis scholasticae* di Faber nell'edizione del 1696, e l'elenco potrebbe continuare.

Raramente i lessici filosofici latini del Seicento forniscono una lista delle fonti utilizzate nella redazione delle voci; un *Catalogus authorum* che registra 49 nomi – tra cui i *Conimbricenses* in relazione ai celebri commentari – si trova nella *Synopsis* di Henry Louis Chasteigner, mentre la tavola delle fonti presente nell'opera di Theodor Hackspan ne ricorda 132 e il lessico di Giovanni Battista Bernardi, *Seminarium totius philosophiae Aristotelicae et Platonicae*[...], I-III, Venetiis: 1582-1585, presenta una lista di autori e opere alla fine di ciascuno dei due tomi. Per gli altri lessici non resta che la lettura analitica delle voci per tentare di individuare la complessa stratificazione di riferimenti che caratterizza ciascun dizionario, pur nella comune ascendenza alla tradizione aristotelico-scolastica.

Come si è detto, la raccolta dei lessici filosofici latini attualmente disponibile sul sito Web dell'ILIESI contiene la riproduzione digitale delle edizioni a stampa dei lessici e la trascrizione di tutte le entrate (e, in parte, delle sottoentrate) di ciascun dizionario. Questo rende possibile effettuare ricerche sul testo, non limitandosi semplicemente alla sua lettura. È possibile infatti l'interrogazione della lista delle entrate e il posizionamento immediato sul lemma che si vuole esaminare all'interno del lessico prescelto. Nella registrazione delle entrate si è fatto ricorso ad alcuni criteri di normalizzazione per rendere più agevole la consultazione. Tali interventi vanno dal semplice adeguamento della grafia all'uso moderno a interventi più cospicui (per esempio, le entrate registrate seguono l'ordinamento alfabetico, talvolta incostante nei lessici presi in esame).

6. Le fasi operative del progetto

Come già ampiamente spiegato l'archivio digitale *Lessici filosofici dell'età moderna* come collezione di testi digitalizzati nasce con l'obiettivo di rendere accessibile online la raccolta, opportunamente organizzata, di una serie di dizionari filosofico-scientifici, a volte di difficile reperimento, e di poter effettuare ricerche terminologiche al loro interno. Per raggiungere tale obiettivo il progetto ha attraversato alcune fasi comuni ai progetti che si occupano di digitalizzazione di materiali testuali per la pubblicazione online. Sono già state indicate le motivazioni legate alla scelta di realizzare un archivio di testi digitali, la tipologia di *audience* a cui si rivolge l'archivio, la scelta dei criteri di selezione dei testi, la definizione delle edizioni scelte (provenienza dei testi, copyright).

La raccolta dei lessici filosofici latini contiene la riproduzione digitale delle edizioni a stampa di 20 lessici, per un totale di circa 13.400 pagine, e la trascrizione delle entrate di ogni lessico, per un totale di oltre 45.000 entrate (vedi tabella 1).

Tabella 1

Autore	Titolo	N. pagine	N. entrate
Alsted	<i>Compendium lexicum philosophicum</i> 1626	717	7604
Atanasio	<i>Terminorum non nullorum</i> 1609	15	10
Bernardi	<i>Seminarium totius philosophiae Aristotelicae et Platonicae</i> 1582	1210	10041
Bruno	<i>Summa terminorum metaphysicorum</i> 1609	142	135
Chasteigner	<i>Celebriorum distinctionum</i> 1612	80	265
Chasteigner	<i>Celebriorum distinctionum</i> 1617	137	265
Chauvin	<i>Lexicon rationale</i> 1692	786	1445
Chauvin	<i>Lexicon philosophicum</i> 1713	764	1603
Faber	<i>Thesaurus eruditionis scholasticae</i> 1696	2613	9843
Goclenius	<i>Lexicon philosophicum</i> 1613	1153	24
Goclenius	<i>Terminorum quorundam explicationes</i> 1609	90	1196
Goclenius	<i>Lexicon philosophicum Graecum</i> 1615	398	416
			(immagini)
Hackspan	<i>Termini, distinctiones et divisiones philosophico-theologicae</i> 1664	702	774
Le Roy	<i>Floretum philosophicum</i> 1649	392	3405
Micraelius	<i>Lexicon philosophicum</i> 1662	1534	3518
Plexiacus	<i>Lexicon philosophicum</i> 1716	328	3942
Reeb	<i>Distinctiones philosophicae</i> 1629	187	71
Richter	<i>Lexicon ethicum</i> 1627	506	442
Scherzer*	<i>Vade mecum sive manuale philosophicum</i> 1675	936	920
Volckmar	<i>Dictionarium philosophicum</i> 1675	705	426
Totale		13395	46345

I testi selezionati per la digitalizzazione sono disponibili su supporti molto eterogenei: 8 ristampe anastatiche, 6 stampe da microfilm, 3 microfilm, 3 testi già digitalizzati provenienti da raccolte online (*Gallica*, ad esempio). Di tutti i testi non già disponibili in formato digitale, 6 sono stati fotografati e 11 sono stati digitalizzati con uno scanner commerciale.

Il processo di digitalizzazione dei testi (che rientra in generale nella *digital imaging*, o grafica digitale) rappresenta una fase importante nella costituzione di biblioteche digitali, o in genere di archivi digitalizzati. Occorre quindi chiarire alcuni concetti che riguardano tale processo, come quello di immagine digitale. Innanzitutto le immagini digitali si dividono in due categorie: le immagini *bitmap* e le immagini vettoriali. Le immagini *bitmap*, o *raster*, sono una rappresentazione grafica basata su una griglia di elementi statici chiamati pixel; un pixel è ciascuno dei punti che compongono la rappresentazione di una immagine nella memoria di un *computer*, e a ciascun pixel vengono assegnati una posizione specifica e un valore cromatico. Le immagini vettoriali sono invece composte da linee e curve definite da oggetti matematici detti vettori; i vettori descrivono la grafica in base alle sue proprietà geometriche. Nel processo di digitalizzazione dei testi per immagine digitale si intende un'immagine

bitmap, bi-dimensionale, idonea per essere visualizzata su un monitor o per essere trasformata in un altro formato, come ad esempio una pagina stampata.

Il numero di pixel contenuti in una determinata unità di misura (pollice) definisce la risoluzione grafica, che viene indicata con l'abbreviazione "ppi" (*pixel per inch*), e determina il dettaglio dell'immagine. Più il numero di pixel per pollice è alto, più l'immagine è nitida. La grafica vettoriale viene convertita in pixel solo al momento della visualizzazione o della stampa; di conseguenza, la grafica vettoriale non è vincolata alla risoluzione; può cioè essere ingrandita o rimpicciolita e stampata a qualunque risoluzione, senza perdita di particolari né di nitidezza. Le immagini *bitmap*, contrariamente a quelle vettoriali, sono vincolate alla risoluzione, ossia contengono un numero fisso di pixel. Di conseguenza, si possono deteriorare e apparire sfocate o scalettate se ingrandite eccessivamente sullo schermo o stampate a una risoluzione più bassa del dovuto.

Un'immagine digitale dà origine ad un file, la cui dimensione è espressa come il numero di *bytes* ottenuti dal numero dei pixel che compongono l'immagine. Più è grande il numero di righe e di colonne della griglia di pixel che compone l'immagine, più è grande la risoluzione e più grande il file. Inoltre, ogni pixel di un'immagine aumenta di misura quando aumenta la sua profondità colore – un pixel da 8 *bit* memorizza 256 colori, un pixel da 24 *bit* memorizza 16 milioni di colori.

7. Il formato immagine

Le immagini digitali sono organizzate e memorizzate secondo metodi standardizzati chiamati formato immagine. Alcuni di questi formati usano un processo di compressione delle immagini che, attraverso particolari algoritmi, diminuisce la grandezza di un file. Per fare un banale esempio l'immagine di una pagina di un testo dalle dimensioni di 21 x 29 cm, acquisita ad una risoluzione di 300 ppi a colori produce un file di circa 25 MB di grandezza. L'applicazione di un formato di compressione, come il JPEG, produce un file di circa 1,5 MB.

Ci sono due tipi di algoritmi di compressione delle immagini, e quindi due tipi di formato: senza perdita di informazioni (*lossless*) e con perdita di informazioni (*lossy*). Gli algoritmi senza perdita di informazioni riducono la grandezza del file senza diminuire la qualità dell'immagine, anche se la dimensione del file non diminuisce in maniera sensibile come nell'altro tipo di compressione. Gli algoritmi con perdita di informazioni approfittano dei limiti dell'occhio umano e eliminano le informazioni invisibili. La maggior parte di questi algoritmi permettono vari livelli di compressione, e di conseguenza possono produrre file più o meno grandi.

Ci sono molti formati di file grafici; di seguito verranno illustrati i formati utilizzati in questo progetto, e cioè JPEG, GIF e TIFF.

Il formato JPEG (Joint Photographic Experts Group) è un formato *lossy*. Quando non è molto elevata, la compressione non degrada visibilmente la qualità dell'immagine, ma nei file JPEG le perdite si accumulano nei vari processi di compressione/decompressione, quando le immagini sono ripetutamente editate e salvate. È quindi preferibile mantenere le immagini che devono essere rieditate in formato *lossless*.

Il formato GIF (Graphics Interchange Format) ha una tavolozza di colori limitata a 256 colori. Questo rende il GIF un formato adatto per memorizzare immagini con pochi colori e senza eccessive sfumature; inoltre si è dotato della funzione "Interfaccia" che permette di far visualizzare subito l'immagine anche se con scarsa qualità, e migliora man mano che il file viene scaricato. Le immagini GIF sono compresse con una tecnica *lossless*, che riduce la dimensione del file senza degradare visibilmente la qualità dell'immagine. I formati JPEG e GIF sono comunemente usati per visualizzare le immagini in Internet.

Il formato TIFF (Tagged Image File Format) è un formato flessibile che può essere sia *lossless* che *lossy*, ma non è supportato dai *browser* Web. Il formato TIFF resta quello largamente accettato come formato fotografico standard nell'editoria.

Nel processo per la digitalizzazione dei lessici, l'acquisizione in formato digitale di un testo ha generalmente puntato alla creazione di un'immagine *master* del testo stesso, salvata come file in formato TIFF. Il file *master* del documento dovrebbe rappresentare il più accuratamente possibile l'immagine del documento originale. La funzione primaria dei file *master* è servire come archivio dei materiali e fonte per le successive elaborazioni grafiche dei file, senza dover ridigitalizzare l'oggetto originale. Un file immagine *master* può servire come surrogato per l'originale, può addirittura sostituire l'originale, o può essere usato come copia di sicurezza nel caso di perdita o danneggiamento dell'originale. I criteri seguiti per l'acquisizione delle immagini via *scanner* o macchina digitale, quando si è potuto, si è cercato di ottenere dei file *master* con la più alta qualità possibile: risoluzione a 300 ppi almeno, scansione a colori o scala di grigi, salvataggio del file immagine in formato TIFF.

I file immagine derivati sono file creati dall'immagine *master* per l'elaborazione, la conversione ad altri formati, la pubblicazione e la trasmissione in rete. Le specifiche per i file derivanti dall'immagine *master* e usati per la pubblicazione possono cambiare nel tempo e essere legate a usi specifici e molto diversi tra di loro.

8. L'elaborazione delle immagini

I testi scelti per l'archivio dei lessici sono testi del Seicento, spesso tratti da microfilm o da ristampe anastatiche; le condizioni dei testi e delle riproduzioni cartacee non sono tutte buone, spesso alcune pagine sono illeggibili. È stata quindi necessaria una fase, più o meno lunga a seconda delle condizioni dell'esemplare, di elaborazione delle immagini, che si è posta come obiettivo quello di avere dei file immagine derivati con una qualità che li rendesse il più possibile leggibili. La fase di elaborazione delle immagini (avvenuta con un comune programma di fotoritocco) ha permesso anche il taglio e l'aggiustamento delle pagine che non si adattavano bene allo schermo; sono così state divise le pagine doppie, sono stati ridotti i margini bianchi quando troppo ampi, sono state raddrizzate le immagini che per qualche motivo erano distorte. In questa fase sono stati anche scelti i formati immagine dei file intermedi (JPEG o GIF) e il livello di risoluzione (in un range da 72 a 120 ppi) in cui convertire le immagini. Il criterio di base che ha guidato la creazione dei file intermedi si è basato non tanto sulla scelta di valori definiti e comuni, quanto sul controllo visivo dei singoli testi, puntando all'ottimizzazione della lettura delle pagine sullo schermo, e solo in secondo luogo all'ottimizzazione per la stampa. La differenza di scelte per i vari lessici ha prodotto un totale di file immagini suddiviso, su un totale di 13.395 immagini, in 37,5% di file jpeg e 62,5% di file gif. La prevalenza di un formato sull'altro è legata al criterio di preferire, quando possibile, file di dimensioni contenute, che possono essere caricati sulla pagina Web in tempi rapidi, ed essendo comunque testi in bianco e nero, non perdono in qualità anche in un formato con limitazione dei colori come il gif. Il controllo di qualità seguito a questa fase si è basato soprattutto sulla verifica della resa sullo schermo delle pagine dei singoli lessici, nonché dei tempi di visualizzazione in rete dell'immagine; sono stati di volta in volta apportati gli eventuali ritocchi.

I file immagine derivati ottenuti sono stati nominati con criteri che ne facilitassero la gestione automatica, e sono stati organizzati in *directory* create con lo stesso criterio. In genere, ad ogni lessico è stata assegnata una sigla, che è servita per nomi-

nare la *directory* contenente le immagini relative al testo; la stessa sigla, insieme al numero della pagina corrispondente, è stata usata per creare il nome del file immagine della pagina.

9. La realizzazione degli indici

Accanto al lavoro di digitalizzazione dei lessici, si è svolto il lavoro di trascrizione, *editing* e realizzazione degli indici. Gli indici preparati sono di due tipi: uno è semplicemente l'indice del volume a stampa, che permette di puntare a sezioni determinate del lessico; l'altro, più complesso, è l'indice delle entrate. Nell'indice delle entrate sono state registrate tutte le voci dei lessici, anche le eventuali voci vuote o di rinvio, con l'indicazione della pagina del lessico in cui sono registrate. In alcuni casi, sono state registrate come entrate anche termini che figurano nel lessico come sottoentrate. Inoltre, a volte, è stata operata una normalizzazione dell'entrata per renderla più congrua rispetto all'ordinamento alfabetico, indicando comunque l'entrata originale tra parentesi quadre. Nell'indice le entrate sono ordinate alfabeticamente, anche se non è sempre così nei lessici. È quindi possibile accedere alle entrate per ordine alfabetico o per ricerca di parole. Ogni entrata, opportunamente codificata, è trascritta in un file di testo, accompagnata dal numero della pagina, o delle pagine in cui compare; una serie di controlli formali producono un file dalla struttura consistente, che contiene le entrate raggruppate per ogni pagina, e che viene trasformato in un file di istruzioni in linguaggio sql per l'*upload* nella tabella dati.

Un approccio diverso è stato seguito per il *Lexicon philosophicum graecum* di Goclenius. In questo caso, infatti, l'elenco delle entrate in greco è stato ottenuto, anziché effettuando la trascrizione delle entrate, che avrebbe richiesto tempi di immissione e controllo più lunghi, attraverso la costruzione di una mappa grafica ottenuta dagli indici originali del lessico. Così, l'elenco delle entrate è in questo caso direttamente l'indice originale del volume, con le entrate greche che puntano all'immagine digitale della pagina corrispondente. I tempi di preparazione dell'indice hanno compreso la costruzione della mappa grafica e la produzione, attraverso *script perl*, dei file indice in html; ovviamente in questo caso non è possibile effettuare ricerche nelle forme.

Per poter permettere la ricerca delle entrate, l'accesso alle pagine dei lessici e lo scorrimento delle pagine, è stato necessario creare una serie di metadati relativi ad ogni pagina (immagine) dei lessici; attraverso tali metadati è stato possibile memorizzare le informazioni idonee ad accedere alle pagine di uno specifico lessico, come l'autore e il titolo dell'opera, il numero della pagina, l'entrata contenuta nella pagina. Successivamente i metadati sono stati organizzati in una tabella di tipo relazionale, con un indice autoincrementale; questa tabella viene poi usata dalle interfacce di interrogazione per la ricerca e la lettura dei lessici.

10. Visualizzazione e lettura dei lessici

La realizzazione dell'ambiente Web di lettura dei lessici è stata guidata da due criteri principali: la semplicità d'uso delle interfacce e la velocità di visualizzazione delle pagine. Nel seguire questo secondo criterio si è incontrato qualche problema, perché le condizioni di alcuni originali erano tali che immagini ad una risoluzione troppo bassa ne avrebbero resa difficoltosa la lettura. In genere comunque è stato possibile applicare questo criterio perché è stata utilizzata un'unica tipologia di immagine in visualizzazione. Non sono infatti presenti per le immagini delle pagine né miniature (inutili per la navigazione perché avviene sostanzialmente attraverso le entrate), né immagini a risoluzione maggiore, più lente da visualizzare e quindi non comode per la consultazione (si è cercato di ottimizzare la lettura del testo ad una dimensione

e risoluzione unica). Se non sembra costituire un fattore influente la presenza di miniature delle pagine, probabilmente può costituire una carenza la mancanza di un'immagine di alta qualità della pagina, non tanto per la consultazione dell'archivio, quanto per la completezza della raccolta digitale. Pur non rientrando negli obiettivi del progetto quello di garantire l'accesso agli originali o a copie digitali il più possibile vicine agli originali, potrebbe forse essere ragionevole programmare un aggiornamento dell'archivio con l'aggiunta delle stesse immagini ad una risoluzione più alta, da usare a scelta per ottenere immagini più simili alle originali. Secondo il criterio di semplicità, lo schermo è stato organizzato con l'elenco delle entrate sulla sinistra, e la pagina del lessico sulla destra. È possibile navigare tra le entrate in ordine alfabetico (scorrendo l'elenco o cambiando lettera) o effettuando la ricerca di un termine o della parte iniziale di un termine. Cliccando sull'entrata viene visualizzata la pagina corrispondente. Si può andare alla pagina successiva, a quella precedente o ad una pagina determinata; si può stampare la pagina visualizzata. Queste sono le sole funzioni sviluppate, ma che permettono di ottenere tutta la documentazione necessaria per la lettura del lessico. C'è poi un tasto Scheda opera che fornisce una scheda bibliografica estesa sul lessico in consultazione, e un tasto Scheda autore per avere notizie sull'autore e sulle sue opere.

Successivamente al caricamento delle immagini e dei metadati nell'archivio e alla loro pubblicazione in ambiente Web, avviene una seconda fase di controllo qualità. In questa fase viene controllata la congruenza del puntamento delle entrate alle pagine relative, e tra le varie pagine e parti del volume, apportando le opportune correzioni e integrazioni; inoltre vengono riviste anche le entrate, apportando gli eventuali aggiustamenti ad esse e, se ancora necessario, alle immagini.

11. Conclusioni

L'archivio digitale *Lessici filosofici dell'età moderna*, in conclusione, rappresenta un tentativo di conciliare uno strumento come l'*open access*, tipico dell'ambito scientifico di ricerca con l'ambito filosofico. La decisione di arricchirlo con schede biografiche degli autori dei lessici e con schede di descrizione dell'opera nasce dalla esigenza di mettere a disposizione dell'utente meno preparato gli strumenti necessari alla comprensione dell'opera stessa, dell'autore e della importanza che questi hanno rivestito nell'ambito della storia del periodo. In effetti le biblioteche degli enti di ricerca sono state finora delle biblioteche specializzate, ossia «granai» chiusi al pubblico, con tradizionali compiti di catalogazione e conservazione di un sapere altamente tecnico e parcellizzato, esclusivamente rivolto ad un'utenza selezionata, quasi sempre locale e soprattutto istituzionale in quanto costituita in massima parte da ricercatori oltre che da studiosi e giovani in formazione che a vario titolo collaborano all'attività di ricerca¹⁴. Ora però che la cosiddetta utenza remota è diventata potenzialmente numerosa, è assolutamente necessario mettere a disposizione anche dell'utente più giovane e qualche volta meno preparato gli strumenti che gli consentano una efficace possibilità di fruizione dei documenti. In effetti non è possibile calcolare il numero né tantomeno la tipologia di persone che sono messe in condizione, attraverso la rete Web, di visitare i siti degli istituti di ricerca e verosimilmente di consultare e ricercare documenti resi disponibili a tutti e quindi si rende necessario dotare gli archivi di tutti i supporti che consentano un efficace consultazione.

¹⁴ Daniela Giustini – Annarita Liburdi, *Tra ricerca scientifica e divulgazione: un nuovo ruolo per le biblioteche degli enti di ricerca?*, «Biblioteche oggi», 21 (2003), n. 9, p. 19-24.

Come detto questi strumenti sono costituiti principalmente da schede biografiche (Scheda autore) e da schede bibliografiche (Scheda opera). La scheda autore non si limita a riportare i dati salienti sulla vita dell'autore ma inquadra la sua figura nel tempo mettendo in evidenza le novità del suo pensiero, mentre la scheda opera oltre alla descrizione bibliografica del volume fornisce una descrizione del contenuto del dizionario e della sua rilevanza. Come già detto sopra per dare risposta al bisogno di chiarezza l'archivio digitale *Lessici filosofici dell'età moderna* è stato arricchito da una sezione «Materiali» che raccoglie una serie di documenti destinati a migliorare la comprensione dei lessici filosofici. Si tratta in realtà di un «open archive» dove i collaboratori dell'archivio depositano liberamente i lavori scientifici sui lessici man mano che si procede nell'implementazione dell'archivio e dove, in prospettiva, si potranno versare anche «materiali» frutto del lavoro di studiosi provenienti da altre istituzioni che decideranno di contribuire liberamente all'archivio digitale *Lessici filosofici dell'età moderna*.

Philosophy and open access: the digital archive

Lessici filosofici dell'età moderna

by Annarita Liburdi and Ada Russo

The digital archive *Lessici filosofici dell'età moderna* (Philosophical dictionaries of the modern age) funded in the years 2007-2008, represents a first stage towards open access.

The archive, that is one of the data bases of the ILIESI (Institute for intellectual lexicon and history of ideas), is a digital library that currently provides students and scholars with about twenty philosophical-scientific and scholarship vocabularies, published in Latin between the end of the 16th and the beginning of the 18th centuries.

The project also foresees the recording of the dictionaries' nomenclature: of all the entries and, to some extent, the sub-entries, with various research possibilities. The archive is also enriched with biographical files on the authors of the dictionaries chosen and with bibliographical files regarding the edition reproduced. A "Materials" section is available for consultation. It contains articles, synoptical files and other study material on the lexicons published.

The digital archive forms part of the research series of the ILIESI: *Studies on European philosophical lexicon of Humanism in the 20th century*, focussed on an analysis of intellectual vocabularies and tools of philosophical lexicography in some representative moments between humanism and modern times.

The objectives of the project *Lessici filosofici dell'età moderna* are:

- 1) Preparation of an archive of philosophical dictionaries (philosophical -scientific and scholarship) that can be consulted on line, in *open source* mode. This archive makes a series of texts accessible: they are texts that are related both to the traditional philosophy of schools (for example ancient and medieval Aristotelian tradition) and to more recent philosophical orientations: from Catholic and Protestant scholastics to Cartesian philosophy. It should be noted that these are works that are often very difficult to find.
- 2) Preparation of a interrelated *thesaurus* of the lexicons, with a combined index of entries (headwords, syntaxes and locutions) of the dictionaries published.

The digital archive *Lessici filosofici dell'età moderna* demonstrates an attempt to reconcile a typical tool of the scientific research environment such as open access with the philosophical environment.

ANNARITA LIBURDI, Istituto per il lessico intellettuale europeo e storia delle idee - CNR, Library, via Nomentana, 118, 00161, Roma, e-mail annarita.liburdi@iliesi.cnr.it.

ADA RUSSO, Collaboratore tecnico Enti di ricerca, CNR-Istituto per il Lessico intellettuale europeo e storia delle idee, via Nomentana 118, 00161 Roma, e-mail ada.russo@iliesi.cnr.it.