

## Archivi Fotografici per la Storia dell'Arte e Semantic Web. Problemi, Risorse e Linee di Ricerca

Marilena Daquino<sup>(a)</sup>

a) Digital Humanities Advanced Research Centre (DHARC); Università di Bologna, <https://orcid.org/0000-0002-1113-7550>

---

Contact: Marilena Daquino, [marilena.daquino2@unibo.it](mailto:marilena.daquino2@unibo.it)

Received: 18 December 2018; Accepted: 13 February 2019; First Published: 15 May 2019

---

### ABSTRACT

In this paper are presented problems, resources, and research lines related to the application Semantic Web technologies in art historical photo archives. In detail, (1) benefits of the usage Linked Open Data for art historians are illustrated, (2) a research line shared between cataloguers and art historians is proposed, namely the definition and retrieval of authoritative artwork attributions in art historical photo archives, (3) a methodology for formalising such a problem is described, and (4) limits and future work related to the valorisation of photo archives catalogues are discussed.

### KEYWORDS

Photo archives; Art History; Semantic Web; Hermeneutics.

### CITATION

Daquino, M. "Archivi Fotografici per la Storia dell'Arte e Semantic Web. Problemi, Risorse e Linee di Ricerca." *JLIS.it* 10, 2 (May 2019): 37-47. DOI: [10.4403/jlis.it-12533](https://doi.org/10.4403/jlis.it-12533).

## Introduzione agli archivi fotografici per la storia dell'arte

La fotografia è stata protagonista di un cambiamento radicale nella creazione, rappresentazione e condivisione della memoria. Modificando drasticamente «the alphabet/image ratio» (Schwartz 1995, 41), ovvero la comunicazione tramite immagini, la fotografia ha influenzato profondamente l'organizzazione della conoscenza.

Ad oggi, biblioteche, archivi e musei sono saturi di collezioni fotografiche, dove rivestono un ruolo fondamentale nel supporto alla ricerca storica. Nonostante ciò, la storia della sedimentazione e della preservazione delle collezioni fotografiche non è stata lineare (Lien e Edwards 2014). L'originalità e l'affidabilità dell'oggetto fotografico sono state spesso messe in discussione (Marien 2006, xiv), essendo oggetto vulnerabile alla storicizzazione, soggetto quindi a riscritture (Lien e Edwards 2014). Pertanto la storia della preservazione dei fondi fotografici si è evoluta su traiettorie differenti rispetto a quelle delle collezioni librarie e museali. Ad esempio, il modello della «social biography» (Gosden, Larson e Petch 2007), che ben si presta alla ricostruzione di collezioni museali – dove la storia degli oggetti culturali viene tracciata su una linea temporale –, poco si adatta alla descrizione di oggetti seriali come le fotografie, le cui traiettorie non sono sempre chiaramente identificabili, tanto meno lineari.

Negli anni Settanta la ricerca di narrative e nuove forme di rappresentazione si manifestò in un ritorno degli istituti dei beni culturali alle collezioni fotografiche. Una tendenza che vide la catalogazione e l'archiviazione come punti cardine della produzione fotografica, data la rinnovata necessità di conservare la memoria storica, scientifica e sociale. Non di meno, la letteratura scientifica di quegli anni è traboccante di continue contraddizioni sulla natura dell'oggetto fotografico – tra informazione, documento e rappresentazione (Schlak 2009, 86). La natura documentaria della fotografia è stata largamente analizzata dalla letteratura archivistica (Edwards and Hart 2004; Franceschini and Mazzucco 2014; Belovari 2013), specialmente per il suo rapporto con gli archivi fotografici per la storia dell'arte (Caraffa 2011). La riproduzione fotografica è infatti uno degli strumenti principali per la ricerca storico-artistica, in quanto *aides-mémoires*, testimonianza di oggetti scomparsi o inaccessibili, fonte per la comparatistica e l'insegnamento.

Similmente, l'archivio fotografico per la storia dell'arte è stato il centro di eccellenza per la ricerca. Tradizionalmente storici, critici d'arte e figure di spicco del mercato dell'arte vi si sono ritrovati per scambiarsi opinioni e registrare i risultati della propria ricerca storico-filologica sulla documentazione fotografica – sotto forma di annotazioni, rimandi a letteratura e consulenze agli archivisti.

Tra queste figure spiccano i *connoisseur*, storici dell'arte che hanno dedicato i loro studi al riconoscimento della paternità delle opere d'arte. Attraverso le loro preziose annotazioni sul verso delle fotografie si conservano le loro testimonianze, l'evoluzione del dibattito storico-artistico e il loro metodo scientifico. Tra le note degli storici vi si trovano preziose informazioni sulla storia attribuzionistica, la storia del collezionismo, gli interessi economici relativi all'opera d'arte (con riferimenti a cataloghi d'asta) e la storia del restauro. Altrettanto preziosi sono i costanti riferimenti bibliografici, registrati dai conoscitori al fine di supportare le proprie tesi attribuzionistiche e dare credito ad altre scuole di pensiero (Mambelli 2018).

In tal senso si è parlato di «potenziale epistemologico» delle collezioni fotografiche per la storia dell'arte (Caraffa 2011), il quale si fonda proprio sulla dibattuta storicizzazione della fotografia (intesa

come fonte storica non neutrale), così come emerge dal rapporto con l'archivio in cui sedimenta. Il fruitore dell'archivio fotografico può scoprire le trame della storia dell'arte grazie non solo all'analisi filologica alla riproduzione dell'opera d'arte, ma in virtù del rapporto che l'oggetto fotografico intrattiene con il suo contesto archivistico, la collezione fotografica e le altre collezioni (librerie e museali) ad essa collegate.

Pertanto, l'archivio fotografico per la storia dell'arte si presenta come una tra le casistiche più interessanti nel panorama dei beni culturali, dove le intersezioni tra discipline (storico-artistica, archivistica, libraria e museale) danno linfa a numerose linee di ricerca e dove i ricercatori possono contribuire in prima persona a sostenere il lavoro di catalogazione e inventariazione.

## Le problematiche sollevate dal Semantic Web

Con l'affermarsi delle tecnologie dell'informazione quale strumento privilegiato per condurre le fasi di ricognizione, sviluppo e condivisione della ricerca scientifica, il rapporto tra archivi per la storia dell'arte e fruitori è mutato radicalmente. Se tra i vantaggi della riproduzione fotografica c'è la possibilità di comparare immagini di opere preservate in luoghi disparati, la digitalizzazione delle immagini e la diffusione sul Web offrono un servizio quantitativamente imparagonabile. L'archivio fisico va svuotandosi di quella componente interlocutoria tra ricercatori, esperti ed esponenti di settori affini, ma il «potenziale epistemologico» persiste e i cataloghi online diventano i nuovi strumenti privilegiati per la valorizzazione del patrimonio fotografico e il supporto alla ricerca storico-artistica. Ad oggi, per valorizzare appieno questi strumenti, aumentando la visibilità degli istituti e fornendo servizi più espressivi ai propri utenti, anche gli archivi fotografici per la storia dell'arte si muovono verso l'adozione di tecnologie per il Semantic Web. Come in altri settori del dominio dei beni culturali, tra gli obiettivi primari c'è l'integrazione di dati tra fonti affini, con l'idea di realizzare vasti cataloghi del patrimonio culturale europeo e internazionale. Di rilievo in questo senso è l'iniziativa promossa dal consorzio PHAROS,<sup>1</sup> che vede riunirsi 14 prestigiosi archivi fotografici per la storia dell'arte in Europa e Nord America. Obiettivo del consorzio è la creazione di un catalogo online per la ricerca storico-artistica, in cui dettagliate informazioni sulle opere d'arte descritte dai vari partner possano convergere e fornire un punto di accesso senza precedenti ai singoli cataloghi.

Nondimeno, le problematiche sollevate dai nuovi strumenti e dal nuovo paradigma sono diverse (Daquino e Tomasi 2016). Dalla necessaria ibridazione di descrizione archivistica, libraria e museale, alla definizione di modelli per la metadattazione e modelli formali per l'organizzazione della conoscenza (le ontologie), i problemi teorici e pratici da affrontare vanno moltiplicandosi (Daquino *et al.* 2017). Ogni istituzione ha infatti sviluppato nel corso delle proprie attività standard descrittivi adeguati alle proprie collezioni, utilizzando o ampliando standard nazionali o di comunità. Di conseguenza lo scambio di informazioni tra istituzioni affini richiede una prima fase di mappatura dispendiosa. In secondo luogo l'eredità di soluzioni tecnologiche – più o meno aperte e più o meno mantenibili – difficilmente permette di dialogare con fonti di dati esterne senza ulteriori costi e/o cambiamenti radicali. Non ultimo, il dialogo tra fonti richiede la revisione e l'aggiornamento delle basi di conoscenza, previa una sofferta fase di *data cleansing*. Il costo per aggiornare le descrizioni catalografiche è infatti dispendioso, in termini di risorse umane, *expertise* e tempo. Ciò comporta una

---

<sup>1</sup> "PHAROS: The International Consortium of Photo Archives," accessed December 18, 2018, <http://pharosartresearch.org>.

drastica selezione delle informazioni da condividere con enti esterni. La conseguenza immediata è un netto ridimensionamento delle domande di ricerca e delle aspettative a cui l'integrazione potrà dare risposta.

La situazione paradossale a cui ci si trova oggi di fronte è dettata dalla “chiamata alle armi” dei *Linked Open Data (LOD)* per il *Cultural Heritage* ed è conseguenza dell'*open data business model*, abbracciato dalle istituzioni culturali solo parzialmente. Le istituzioni culturali di tutto il mondo vengono spinte a disseminare i propri dati secondo il nuovo paradigma – dati strutturati, organizzati secondo modelli condivisi, ripuliti da incoerenze e variabili locali, riconciliati ad *authority* e arricchiti semanticamente –, il cui costo però è ancora alto in termini di expertise e di messa in produzione ed il cui ritorno in termini di visibilità, effettivo utilizzo e beneficio per gli utenti non è ancora né misurabile, né evidente. Ad oggi solo la Biblioteca Nazionale Svedese (KB) (<http://www.kb.se/libris/>) ha adottato il paradigma Linked Data appieno, abbandonando i precedenti sistemi gestionali e pubblicando la sua intera base di conoscenza utilizzando tecnologie per il Semantic Web. Pertanto, una valutazione puntuale su costi e benefici è ancora prematura. Il risultato evidente è l'accesso libero a informazioni – anche molto complesse e strutturate – che gli utenti (medi) non sono in grado di riutilizzare senza avere adeguate competenze informatiche e di cui non sono immediatamente evidenti i benefici.

L'aumento della complessità da affrontare per poter utilizzare appieno *cutting-edge technologies* riguarda tanto il fronte delle istituzioni culturali quanto quello degli interlocutori. La tecnologia impone una rivisitazione sostanziale delle pratiche di formazione del personale degli istituti culturali e di sviluppo della ricerca tradizionale. Per superare lo scoglio tecnologico – e massimalizzare il «potenziale epistemologico» dei dati, parafrasando il concetto espresso da Caraffa – è necessario un ulteriore contributo da parte delle istituzioni che si fanno promotrici di progetti legati all'*open data movement*.

Procedendo empiricamente verso progetti e soluzioni che muovono dall'utilizzo dei dati prodotti, musei, archivi e biblioteche promuovono precisi percorsi di ricerca. In altri termini, l'investimento delle istituzioni non si esaurisce nel liberare i dati, ma sconfinava nella ricerca stessa e nell'offerta di servizi che sappiano sfruttare tali dati per svolgere attività più complesse – le quali sarebbero escludenti e impensabili da sviluppare per utenti dalle minime competenze informatiche.

In tale contesto, gli archivi fotografici della storia dell'arte sollevano alcune ulteriori problematiche. La catalogazione qui investe potenzialmente oggetti molto eterogenei (documentazione fotografica, archivistica, oggetti museali ed eventualmente beni librari annessi) tali da richiedere sia una conoscenza complessiva del dominio dei beni culturali, sia nozioni specialistiche di catalogazione per la storia della fotografia e per la storia dell'arte. Servizi fondati su tali informazioni richiedono una visuale a 360° sul dominio dei beni culturali. L'utilizzo di standard, modelli e buone pratiche preesistenti in un caso d'uso così trasversale non è immediato, né è sempre sufficiente ad esprimere la complessità dello scenario. L'assenza di soluzioni – concettuali, terminologiche e tecnologiche – è rimarcata dalla non ancora effettiva integrazione MAB (Musei Archivi e Biblioteche), per cui la definizione di problematiche comuni e pratiche condivise non è ancora formalizzata e si procede empiricamente, caso per caso.

Inoltre, lo scenario interdisciplinare degli archivi per la storia dell'arte solleva problematiche che nelle singole discipline (archivistica, libraria e museale) non sempre vengono affrontate chiaramente. Ad esempio una chiara definizione dell'approccio ermeneutico adottato dai catalogatori nell'esegesi di

fonti (eterogenee appunto) che riportano informazioni contraddittorie non è formalizzata in alcun dominio. Per riprendere l'esempio dei connoisseur, gli archivisti devono districarsi tra differenti attribuzioni dell'opera d'arte rappresentata nella documentazione fotografica e valutare le relative fonti discordanti (siano queste annotazioni sul verso della foto, riferimenti bibliografici o perizie conservate in archivio), eventualmente fornendo una descrizione delle motivazioni e documentando la storia attribuzionistica alternativa. Come giudicare tali decisioni prese in sede di catalogazione è appannaggio dell'utente finale, a cui non sempre viene fornito il necessario contesto (le fonti utilizzate, il metodo e le motivazioni con cui si giudica una fonte, né le preferenze dell'istituzione di fronte a certe tipologie di fonti più o meno autorevoli).

In ultimo, la prospettiva di integrazione di dati affini effettuata secondo i limiti sopra evidenziati, rischia di risolversi in una destrutturazione e decontestualizzazione delle collezioni fotografiche, le quali perdono quel bagaglio informativo di cui lo storico poteva beneficiare in archivio. Si pensi ad esempio a collezioni fotografiche che sono il risultato di campagne fotografiche, documentate in cataloghi di fotografo. Tali campagne sono spesso il sintomo di un interesse generale (economico e culturale) nei confronti di un dato movimento artistico o genere, e possono evidenziare relazioni (latenti) tra opere le quali non sono state ancora investigate. Nel catalogo dell'archivio questa informazione viene preservata, esplicitando la provenienza della fotografia dell'opera dal catalogo in questione e descrivendo la relazione con altre opere. Facilmente queste informazioni verranno scartate nella selezione dei metadati da condividere con altre istituzioni, essendo spesso una informazione registrata in linguaggio naturale, difficile da strutturare automaticamente, che tratta di oggetti non sempre identificabili e riconciliabili con altri basi di dati.

In conclusione, finora gli enti sono stati promotori e primi attori nella definizione dei modelli per la rappresentazione del proprio patrimonio culturale, dando vita a molte risorse per supportare il passaggio al Semantic Web mantenendo salde le tradizioni descrittive delle discipline di partenza. Ad ogni modo, l'assenza di formalizzazione di percorsi trasversali tra domini affini, ovvero tra collezioni eterogenee (librarie, archivistiche e museali), e di definizione di domande di ricerca, richiede un ulteriore coinvolgimento attivo delle istituzioni. In particolare, fondamentale è il ruolo degli istituti nell'elaborazione di risorse, metodi e servizi per favorire il dialogo tra domini e, di conseguenza, la fase di *design* di nuove linee di ricerca.

## Risorse e metodi per la transizione al Semantic Web

Lungi dall'essere una panacea universale, le tecnologie per il Semantic Web sono strumenti efficaci per ideare e realizzare percorsi informativi, estendendo basi di conoscenza esistenti e supportando lo sviluppo di nuove applicazioni. In particolare, i LOD offrono soluzioni tecnologiche per rispondere a tre specifiche esigenze: identificazione, integrazione e scoperta dei dati.

L'identificazione di entità (persone, luoghi, soggetti, opere, ecc.) descritte in record catalogafici e la loro riconciliazione ad *authority file* è un metodo noto nel dominio dei beni culturali. L'utilizzo di thesauri, liste d'autorità e vocabolari controllati condivisi dalla comunità permette di uniformare la descrizione di entità, garantire la consistenza tra sistemi informativi e favorire la creazione di punti di accesso comuni alle collezioni. Tra le risorse più rilevanti per gli archivi fotografici di storia dell'arte si trovano il tesoro ICONCLASS per la soggettazione (Couprie 1983), i vocabolari del Getty (Baca

and Gill 2015) per l'identificazione di tecniche e tipologie di opera (AAT), artisti (ULAN) e luoghi storici (TGN).

Questo procedimento centralizzante in cui più *data provider* si rifanno alle specifiche di uno standard comune, è propedeutico alla realizzazione di Linked Data. A ciò si affianca una riconciliazione, per così dire, orizzontale. Il contributo degli authority file può essere infatti parziale, non sempre inclusivo di tutte le risorse necessarie a descrivere una fonte di dati specialistica. Un approccio decentralizzato alla riconciliazione, ovvero il *match* di informazioni equivalenti descritte in fonti di dati affini, egualmente autorevoli, permette di colmare alcune delle lacune in authority esistenti e confrontare direttamente informazioni concorrenti sugli stessi oggetti o entità. A sua volta però questo espediente solleva nuove problematiche, come la possibilità di dover confrontare informazioni contraddittorie. In questo caso, la descrizione della provenienza delle informazioni consente di preservare il contesto delle asserzioni stesse e supportare il giudizio finale dell'utente.

L'integrazione tra dati è favorita dall'utilizzo di un unico modello per la rappresentazione dei dati, i.e. RDF, e dal riuso di vocabolari e ontologie esistenti per descrivere scenari comuni. Ciò favorisce l'interrogazione – semplificando le richieste, riducendo complessità e rischio di errori – ed evidenzia immediatamente alcuni percorsi informativi (tramite l'aggregazione di risorse simili sotto le medesime classi) che altrimenti richiederebbero una conoscenza approfondita dei diversi modelli in uso e una mappatura tra essi. Tra le risorse ad oggi a disposizione per gli archivi fotografici della storia dell'arte ci sono CIDOC-CRM (Doerr 2003), il golden standard per la descrizione di oggetti museali, OAD e EAC-CPF (Ricci 2014) per la descrizione archivistica, le SPAR Ontologies (Peroni e Shotton 2018), Resource Description and Access (RDA)<sup>2</sup> e BIBFRAME (Kroeger 2013) per la descrizione delle risorse bibliografiche collegate. Inoltre, per preservare le informazioni sulla provenienza degli asserti potenzialmente contraddittori esistono PROV-O (Moreau *et al.* 2015), W3C standard per la rappresentazione della *provenance*, e HiCO (Daquino e Tomasi 2015), estensione di PROV-O per la descrizione del processo ermeneutico alla base di informazioni opinabili. Una vasta selezione di classi e proprietà provenienti dai modelli menzionati è inclusa nei modelli OA Entry e F Entry (Daquino *et al.* 2017), realizzati sulla base del mapping degli standard ICCD-OA e ICCD-F per la descrizione catalografica dell'opera d'arte e della fotografia.

L'integrazione può avvenire utilizzando metodi diversi, può essere caratterizzata da differenti livelli di complessità ed avere obiettivi differenti. L'identificazione di risorse equivalenti in due o più fonti permette di perfezionare descrizioni di risorse note, completando informazioni parziali o assenti in un dataset con informazioni presenti in un nuovo dataset. Ad esempio gli archivisti possono aggregare informazioni biografiche di artisti (descrizioni, date di nascita e morte, attività, relazioni con altri artisti, ecc.) da ULAN o DBpedia e arricchire così i propri authority locali. In questo caso il fine è l'arricchimento della base di conoscenza iniziale, sfruttando contenuti di qualità già esistenti e risparmiando il tempo delle attività di ricerca dei catalogatori. Inoltre, il confronto tra fonti che riportano informazioni contraddittorie può portare ad una modifica della base di conoscenza iniziale. Confrontando i dataset di archivi fotografici per la storia dell'arte, musei e biblioteche ad esempio è possibile ricostruire (tra gli altri aspetti) la storia attribuzionistica, valutare le informazioni più aggiornate e documentate e modificare le informazioni incluse nei cataloghi – che possono essere

---

<sup>2</sup> "RDA Registry," accessed December 18, 2018, <https://www.rdaregistry.info/>.

poco aggiornate, parziali, incomplete o scorrette. In questo caso l'obiettivo è selezionare informazioni già vagliate da fonti autorevoli e risparmiare tempo nella fase di ricognizione. In ultimo, l'integrazione può avvenire tramite aggregazione e può essere portata avanti senza finalità definite a priori. È questo il caso degli aggregatori come Europeana (Isaac e Haslhofer 2013) o PHAROS, dove l'esplorazione e la definizione delle linee di ricerca è demandata all'utente.

L'ultimo elemento chiave caratterizzante i Linked Data è la scoperta di nuovi percorsi informativi. Questa può avvenire tramite semplici meccanismi di *serendipity*, ovvero la scoperta casuale di informazioni messe in risalto dalla riorganizzazione della conoscenza, o *follow your nose*, percorrendo i nodi intermedi che collegano entità. Altre strategie per la scoperta di nuove informazioni si possono ottenere dall'integrazione di fonti eterogenee, ovvero *mashup*. Applicazioni basate su *mashup* consentono di generare nuove tipologie di servizi per gli utenti, evidenziando percorsi impossibili da ottenere basandosi sulla sola conoscenza disponibile in una fonte di partenza. Ad esempio è possibile aggregare le storie attribuzionistiche registrate in differenti istituzioni, includendo informazioni bibliografiche e risorse archivistiche collegate, e fornire così un servizio specialistico per storici dell'arte. In ultimo, la possibilità di generare inferenze consente di scoprire conoscenza latente, non evidente nei percorsi possibili nei cataloghi di partenza. Sempre a titolo di esempio, è possibile utilizzare informazioni biografiche sui fotografi per inferire la datazione di nuove fotografie da catalogare.

La ricerca di euristiche per *knowledge discovery* (automatica o semi-automatica) nel campo dei beni culturali dipende ancora molto dal contributo degli esperti di dominio, i quali forniscono linee guida e spiegazioni sulle dinamiche sottintese ai dati. Fondamentale è il loro contributo per formalizzare numerosi aspetti che sono latenti nelle descrizioni archivistiche e catalografiche, come ad esempio il già menzionato approccio ermeneutico dei catalogatori. La definizione formale di quali fonti siano autorevoli, e come queste vengano giudicate tali, permetterebbe infatti di comprendere i meccanismi profondi che sottostanno alla creazione della conoscenza negli istituti culturali così da riprodurli e supportare gli utenti nel giudizio di fronte a informazioni contraddittorie.

## **Linee di ricerca per gli archivi fotografici per la storia dell'arte in *Digital Art History***

Assumendo che il compito principale delle istituzioni culturali sia preservare, valorizzare il patrimonio culturale di cui dispongono e supportare gli utenti nell'esplorazione, si può intuire che le linee di ricerca richiedenti un coinvolgimento diretto degli enti si plasmino proprio attorno ai percorsi informativi degli utenti. Le necessità di un fruitore di un archivio fotografico per la storia dell'arte possono essere schematizzate nei seguenti tre punti: (i) essere accompagnato nella ricognizione della documentazione fotografica, avendo chiaro il potenziale informativo delle risorse e la visione d'insieme sulla collezione e i collegamenti interni; (ii) essere supportato nella ricognizione di fonti (eterogenee) collegate all'oggetto di studio, cogliendo così i legami tra la collezione fotografica e la documentazione d'archivio o libraria; (iii) evidenziare le fonti rilevanti, aggiornate e autorevoli sul tema in analisi.

Quando la ricerca avviene tramite i cataloghi online, l'utente deve innanzitutto intuire quali informazioni sono effettivamente presenti e dove reperirle. Successivamente utilizzerà alcuni

strumenti contemporaneamente (p.e. OPAC, altri cataloghi e siti web) per raccogliere tutte le fonti necessarie a inquadrare il tema di ricerca. In ultimo, avvierà la fase di valutazione delle fonti. Le operazioni di ricognizione possono essere lunghe e portare a risultati incompleti, dovuti alle difficoltà derivanti dall'utilizzo dei servizi online, errori dell'utente o l'assenza della risorsa digitale ricercata. Inoltre, mancando il supporto dell'esperto di dominio, la fase di valutazione delle fonti può richiedere più tempo e può risentire della conoscenza parziale sull'argomento. Il risultato è uno spaesamento dell'utente, che non è certo di aver reperito tutte le informazioni disponibili, sia perché queste non sono effettivamente disponibili online o perché non è stato in grado di reperirle correttamente.

In tale contesto il ruolo degli archivi non può esaurirsi nel disseminare e aggregare open data e demandare all'utente la comprensione dei percorsi conoscitivi. I servizi *user-centered* si basano sulla convivenza di informazioni e tool a supporto della ricerca, dove percorsi flessibili alle caratteristiche dell'utente sono identificati innanzitutto sulla base delle specificità del fornitore di dati (ovvero la natura delle ricerche effettivamente possibili in un dato archivio) e formalizzati per rispondere a determinate domande di ricerca.

Una chiara linea di ricerca per gli archivi fotografici per la storia dell'arte (ma ciò si estende ad altre tipologie di istituti) è allora riprodurre quella metodologia che gli archivisti hanno sviluppato nella catalogazione delle collezioni e metterla al servizio degli utenti. Così facendo è possibile condividere non solo dati, ma il contesto dei dati, ovvero l'elemento che consente di sviscerare il "potenziale epistemologico" delle collezioni. Ciò comporta (i) analizzare la metodologia adottata dagli archivisti nel reperimento delle informazioni, (ii) formalizzare gli aspetti del loro approccio ermeneutico e (iii) applicare metodi di ranking per la valutazione dell'autorevolezza dei dati in un certo dominio. Il risultato è l'offerta di strumenti sofisticati per il supporto alla ricerca, capaci di replicare la metodologia di esperti di dominio e supportare (senza ovviamente sostituire) le decisioni di un utente. Il progetto *mAuth - mining Authoritativeness in Art History* (<http://purl.org/emmedi/mauth>) va in questa direzione. Metodi quantitativi di *data analysis* vengono utilizzati per estrapolare informazioni sul metodo di lavoro degli archivisti a partire da un corpus di record catalografici che includono dettagliate informazioni sulla storia attribuzionistica. Il corpus comprende record di tre archivi fotografici per la storia dell'arte, appartenenti alla Fondazione Federico Zeri (Università di Bologna), a Villa I Tatti - Berenson Library (Harvard University) e alla Frick Art Reference Library di New York.

Il risultato della prima fase di analisi è un vocabolario controllato di termini inerenti le motivazioni a supporto delle attribuzioni (bibliografia, attribuzione di storico, attribuzione d'asta, ecc.). Per stabilire come gli archivisti scelgano il criterio più affidabile per giudicare la bontà di un'attribuzione, è stato chiesto a esperti di dominio delle istituzioni in questione di rivedere la lista di termini estratta e fornire un primo *rating* dei criteri. Ad esempio, l'attribuzione fornita da uno storico dell'arte o riportata su una pubblicazione scientifica è tendenzialmente considerata più affidabile di un'attribuzione fornita da una casa d'asta (il cui giudizio, se contraddetto da altre fonti, può essere reputato influenzato da interessi economici).

Le ipotesi formulate da esperti di dominio sono state poi validate da una ulteriore analisi dei dati. In particolare, i criteri a sostegno delle attribuzioni accettate dalle tre istituzioni sono stati confrontati, uno ad uno, con i criteri a sostegno di attribuzioni scartate (comunque registrate nei record



catalografici). Il risultato è un rating finale di criteri con cui effettuare un primo ranking delle fonti di attribuzioni.

In seguito a questa analisi comparativa, che investe esplicitamente gli aspetti documentali delle attribuzioni (e non l'analisi della fonte primaria, *i.e.* l'opera stessa), altri fenomeni sono risultati essere influenti nella decisione finale degli archivisti. Innanzitutto, la data dell'attribuzione è considerata significativa. Le attribuzioni più recenti, supportate da criteri autorevoli, sono comunemente repute dagli esperti di dominio come più plausibilmente veritiere delle precedenti. Allo stesso modo, l'accettazione dell'attribuzione da parte di più istituzioni è considerato un valore aggiunto alla credibilità dell'attribuzione stessa. Non ultimo, l'autorevolezza di chi ha per primo effettuato l'attribuzione (sia esso uno storico dell'arte, un museo o altra istituzione) e di chi pubblica l'informazione in una fonte secondaria sul web (in questo caso l'archivio fotografico che pubblica il catalogo online) sono aspetti necessariamente valutati.

A differenza dei precedenti criteri (rating delle motivazioni, datazione, numero di fonti in accordo), l'autorevolezza dello storico dell'arte è difficilmente quantificabile. Ad oggi le metriche per valutare accademici si basano su indici citazionali (Mitchell *et al.* 2011), basati sul conteggio di citazioni di pubblicazioni accademiche. Sfortunatamente i dati citazionali sono carenti o completamente assenti per quegli storici dell'arte che hanno operato tra la fine del diciannovesimo secolo e l'inizio del ventesimo secolo. Tale limite non consente di avere indici affidabili per misurare l'impatto di uno storico dell'arte all'interno del dibattito attribuzionistico. Inoltre, la peculiarità degli archivi fotografici risiede in differenti forme di accreditamento delle informazioni – e delle attribuzioni in particolare. Non solo pubblicazioni scientifiche, ma annotazioni sul verso di fotografie, perizie documentate e comunicazioni verbali vengono citate in record catalografici a supporto di un'attribuzione, dando forma a differenti forme citazionali.

Pertanto, diverse formalizzazioni di un atto citazionale meritano di essere incluse nel calcolo delle metriche, qualsiasi razionale esse seguano, così da identificare compiutamente l'impatto di uno storico nel dibattito storico-artistico. È questa un'altra interessante linea di ricerca per gli archivi fotografici per la storia dell'arte. Uscendo dai limiti del solo ambito disciplinare (storico-artistico) e dalle convenzioni archivistiche, si promuoverebbe una sapiente integrazione dei domini MAB. L'intersezione tra collezioni archivistiche, museali e librerie di enti che condividono un dato tema di ricerca è un obiettivo di massima condiviso, che passa per l'identificazione delle risorse in comune, l'integrazione dei dati e la scoperta dei percorsi informativi creati non tramite ecumenici contenitori di (pochi) dati eterogenei, ma utilizzando la lente delle domande di ricerca come guida nella creazione, disseminazione e valorizzazione di dati dei beni culturali.

Tali percorsi profilati, ovvero dipendenti da uno o più specifici punti di vista (del catalogatore e dello storico) sono una delle promesse fatte dalle tecnologie del Semantic Web. In questo senso, il progetto mAuth intende utilizzare il patrimonio linked data creato dalle istituzioni culturali per definire una banca dati di citazioni per la storia dell'arte e portare avanti un'ulteriore fase di analisi dei dati citazionali. L'obiettivo è definire percorsi storiografici, identificare gli attori che hanno influenzato il dibattito storico-artistico su determinati temi e riprodurre le trame della storia dell'arte a partire dagli archivi in cui originariamente trovavano un luogo proficuo.

## Conclusioni

Gli scenari sopra elencati aiutano a comprendere quali siano i benefici che metodi quantitativi e tecnologie per il Semantic Web possano portare agli archivi fotografici per la storia dell'arte. È stato rimarcato il ruolo attivo degli archivi e dei catalogatori nel processo di presa di coscienza delle possibilità offerte dalle tecnologie, così come nella definizione di linee di ricerca declinate sugli interessi di istituzioni e utenti. Infine sono state descritte alcune direzioni di ricerca in cui gli archivi naturalmente trovano un campo di applicazione. Ciò che emerge è che l'applicazione di tecnologie web per gli istituti di storia dell'arte non deve limitarsi alla creazione di raccolte di dati eterogenei, da cui aspettarsi serendipitose scoperte, bensì, deve sapersi misurare con specifici problemi di ricerca, da questi generalizzare – adottando un processo induttivo – e delineare linee di ricerca interdisciplinari e trasversali al dominio dei beni culturali, capaci di rispondere alle esigenze informative degli utenti.

## Riferimenti bibliografici

- Baca, Murtha, and Melissa Gill. 2015. "Encoding multilingual knowledge systems in the digital age: the getty vocabularies." *Knowledge Organization* 42, 4:232–243.
- Belovari, Susanne. 2013. "Professional minutia and their consequences: provenance, context, original identification, and anthropology at the Field Museum of Natural History, Chicago, Illinois." *Archival Science* 13, 2-3:143–193.
- Caraffa, Costanza (Ed.). 2011. *Photo archives and the photographic memory of art history*. Berlin: Deutscher Kunstverlag.
- Couprie, Leendert D. 1983. "Iconclass: an iconographic classification system." *Art Libraries Journal* 8, 2:32–49.
- Daquino, Marilena, *et al.* 2017. "Enhancing semantic expressivity in the cultural heritage domain: exposing the Zeri Photo Archive as Linked Open Data." *Journal on Computing and Cultural Heritage* 10, 4:21.
- Daquino, Marilena, and Francesca Tomasi. 2016. "Digital Humanities e Library and Information Science. Through the lens of knowledge organization." *Bibliothecae.it* 5, 1:130–150. Accessed December 18, doi: [10.6092/issn.2283-9364/6109](https://doi.org/10.6092/issn.2283-9364/6109).
- Daquino, Marilena e Francesca Tomasi. 2015. "Historical Context Ontology (HiCO): a conceptual model for describing context information of cultural heritage objects." In *Research Conference on Metadata and Semantics Research*, 424–36. Springer, Cham.
- Doerr, Martin. 2003. "The CIDOC conceptual reference module: an ontological approach to semantic interoperability of metadata." *AI magazine* 24, 3:75.
- Edwards, Elizabeth, and Janice Hart (Eds.). 2004. *Photographs objects histories: on the materiality of images*. Routledge.

- Franceschini, Chiara and Katia Mazzucco (Eds). 2014. "Classifying content. Photographic collections and theories of thematic ordering". In *Visual Resources. An international journal of documentation* 30, 2.
- Gosden, Chris, Frances Larson, and Alison Petch. 2007. *Knowing things: exploring the collections at the Pitt Rivers Museum, 1884-1945*. Oxford: Oxford University Press.
- Isaac, Antoine and Bernhard Haslhofer. 2013. "Europeana linked open data–data. europeana. eu." *Semantic Web* 4, 3:291–297.
- Kroeger, Angela. 2013. "The road to BIBFRAME: the evolution of the idea of bibliographic transition into a post-MARC Future." *Cataloging & classification quarterly* 51, 8:873–890.
- Lien, Sigrid, and Elizabeth Edwards (Eds.). 2014. *Uncertain images: Museums and the work of photographs*. Ashgate Publishing, Ltd.
- Mambelli, Francesca. 2018. "Bene, documento, fonte. La fotografia negli archivi fotografici degli storici dell'arte." *INTRECCI d'arte* 7:91–101. Accessed December 18, doi: [10.6092/issn.2240-7251/8609](https://doi.org/10.6092/issn.2240-7251/8609).
- Marien, Mary Warner. 2006. *Photography: A cultural history*. Laurence King Publishing.
- Mitchell, Gordon R., *et al.* 2011. "Measuring scholarly metrics."
- Moreau, Luc, *et al.* 2015. "The rationale of PROV." *Web Semantics: Science, Services and Agents on the World Wide Web* 35:235–57.
- Peroni, Silvio and David Shotton. 2018. "The SPAR ontologies." In *International Semantic Web Conference*, 119–136. Springer, Cham.
- Ricci, Francesca. 2014. "Il progetto italiano ReLOAD al LODLAM Summit 2013 Linked open data in libraries archives and museums." *DigItalia* 2:173–181.
- Schlak, Tim. 2008. "Framing photographs, denying archives: the difficulty of focusing on archival photographs." *Archival Science* 8, 2:85–101.
- Schwartz, Joan M. 1995. "We make our tools and our tools make us: Lessons from Photographs for the Practice, Politics, and Poetics of Diplomats." *Archivaria* 40.