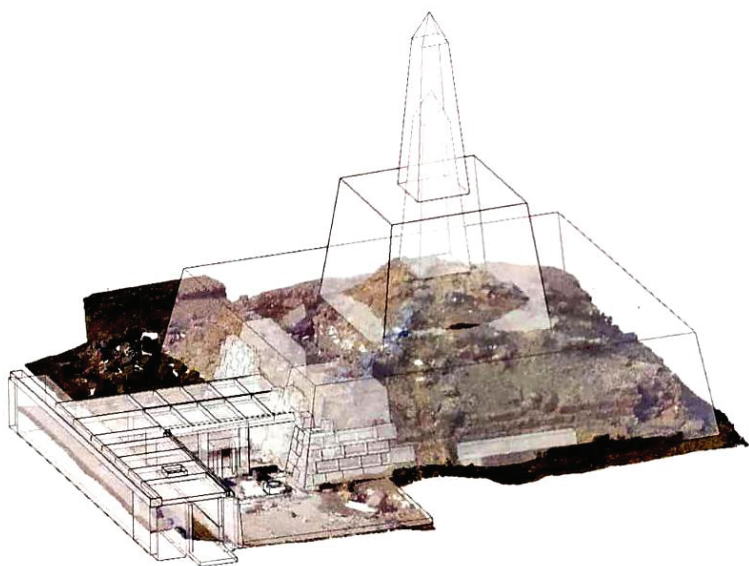


Andrea D'Andrea

# I dati archeologici nella società dell'informazione



UniorPress







UNIVERSITÀ DI NAPOLI L'ORIENTALE  
DIPARTIMENTO ASIA AFRICA E MEDITERRANEO

Andrea D'Andrea

I DATI ARCHEOLOGICI  
NELLA SOCIETÀ DELL'INFORMAZIONE



UniorPress  
Napoli 2023

*In copertina:* ricostruzione del tempio solare di Niuserra  
(Abu Gurab, Egitto; elaborazione di Patrizia Zanfagna, da Bosco *et al.* 2018)

Edizione digitale con licenza  
Creative Commons Attribution 4.0 International



UniorPress  
Via Nuova Marina 59, 80133 - Napoli

ISBN 978-88-6719-272-4

# INDICE

PREMESSA	9
Capitolo 1 - L'ARCHEOLOGIA DIGITALE	15
1.1. Introduzione	15
1.2. L'Archeologia Digitale come disciplina storica	17
1.3. Dalla <i>digitization</i> alla <i>digital transformation</i>	25
1.4. I legacy data	32
1.5. <i>Open Knowledge</i>	41
1.6. Riflessioni di metodo	45
Capitolo 2 - L'INFORMAZIONE TESTUALE	53
2.1. La nascita dell'Archeologia Digitale	53
2.2. I dati testuali	57
2.3. La formalizzazione	60
2.4. Dal Semantic Web alla Scienza Aperta	62
2.4.1. L'architettura del Semantic Web	68
2.4.2. Open Data e <i>Linked Open Data</i>	70
2.4.3. Thesauri e vocabolari	74
2.4.4. FAIR Data	80
Capitolo 3 - OPEN ARCHAEOLOGY	87
3.1. Introduzione	87
3.2. Thesauri e strumenti terminologici	93
3.3. I dati FAIR in archeologia	103
3.4. Metadati	107
3.5. <i>Paradata</i>	111
3.6. Ontologie di dominio e <i>task ontology</i>	115
3.6.1. Il CIDOC-CRM	121
3.7. Repository e infrastrutture	124
3.8. Aggregatori	130

Capitolo 4 - L'INFORMAZIONE SPAZIALE	135
4.1. I dati spaziali	135
4.2. Standard per l'informazione geografica	139
4.3. Il rilievo digitale in archeologia	146
4.4. Oggetti archeologici 3D	154
4.4.1. Metadati per il 3D	157
4.4.2. IPR e 3D	161
4.4.3. Principi di base per il 3D	163
4.5. La visualizzazione del dato geometrico 3D	164
CONCLUSIONI	171
5.1. Organismi internazionali per la documentazione archeologica	171
5.2. L'archeografia digitale	174
5.3. I dati territoriali	180
5.4. I dati 3D	181
5.5. Archeologia e <i>data-science</i>	186
ABBREVIAZIONI	191
BIBLIOGRAFIA	195
SITOGRAFIA	217

A Dora e Mimmo,  
*Ai miei genitori che mi hanno trasmesso  
l'arte della comprensione umana*





## PREMESSA

«... Two roads diverged in a wood, and I –  
I took the one less travelled by,  
And that has made all the difference».  
(ROBERT FROST, *The Road Not Taken*, 1915)

Pubblicando nel 2006 il volume “Documentazione Archeologica, Standard e Trattamento Informatico” avevo riservato la seconda parte del lavoro alla illustrazione di alcuni progetti GIS che avevo coordinato tra gli anni '90 ed il 2000. La parte introduttiva di quella sezione si soffermava sulla ricchezza del dibattito sui GIS in archeologia, un tema circoscritto, almeno in quegli anni, all'analisi dei potenziali avanzamenti nella metodologia e nella ricerca archeologica e all'approfondimento della tecnologia a quel tempo disponibile.

Se l'applicazione dei GIS per la ricostruzione del territorio antico trova oggi poco spazio nei convegni scientifici, ciò non si deve tanto ad una rimozione disciplinare o ad un ripensamento metodologico, quanto piuttosto alla banale considerazione che, ad oltre 35 anni dalla comparsa dei primi contributi sui sistemi digitali per l'analisi spaziale, oramai il settore delle applicazioni GIS nel campo delle Humanities non suscita più un particolare interesse nei ricercatori. Questa situazione contrasta con una diffusa richiesta di sviluppo di sistemi informativi che non sempre appare soddisfatta da professionisti in grado di progettare applicazioni che resistano all'avanzamento della tecnologia e, soprattutto, alla grande quantità di dati di qualsiasi natura disponibili liberamente sulla rete.

Poiché le ragioni alla base di quel volume restano, a mio avviso, ancora valide e forse ancora di più, considerata la semplificazione nell'accesso all'informazione ed al suo trattamento, ho ritenuto importante riprendere il lavoro già avviato anni fa, concentrando la mia attenzione sul tema più complessivo dei dati e della loro analisi in archeologia, un ambito di ricerca che ha conosciuto, grazie alla dimensione digitale, un significativo incremento e che richiede opportune riflessioni di metodo.

Mentre cresce la richiesta di ideazione e costruzione di sistemi informativi, anche come effetto positivo della diffusione di banche dati online facilmente scaricabili e riutilizzabili, dobbiamo registrare, almeno nel nostro paese, un ritardo istituzionale e formativo, oramai quasi strutturale, che ha profonde radici culturali. Una parte non minoritaria di studiosi considera il digitale in termini di adeguamento tecnico all'evoluzione degli strumenti non avendo alcuna influenza sui più noti e sperimentati processi di conoscenza. Poiché l'analisi computazionale viene derubricata a semplice cambiamento tecnico, la progettazione e la realizzazione di un sistema informativo può essere appaltata, economicamente e concettualmente, agli esperti (ingegneri, informatici, geografi e architetti).

L'opposizione, in primo luogo accademica, nei confronti dell'avanzamento di una Archeologia Digitale come disciplina autonoma, ha prodotto numerosi e visibili risultati distorsivi, principalmente al livello della gestione ordinaria del patrimonio archeologico.

In questo periodo, caratterizzato dalle preoccupazioni economiche e sociali conseguenti al pericolo di nuove crisi pandemiche, la discussione nel nostro paese è focalizzata sull'impiego dei cospicui fondi che l'EU ha assegnato all'Italia. Una quota significativa delle risorse indicate nel piano Next Generation EU è, ancora una volta, destinata alla digitalizzazione dei beni culturali, una strada oramai percorsa da moltissimi anni con esiti in larga misura disastrosi. Questo piano si configura come l'ennesima riproposizione di un approccio inaugurato con la stagione dei Giacimenti Culturali, termine coniato dall'allora Ministro De Michelis: l'art. 15 della Legge 41/1986 stanziò 600 Mld di lire (poco più di 3 Mln di euro) per finanziare progetti presentati da società informatiche per la catalogazione del patrimonio culturale italiano attraverso l'utilizzazione delle tecnologie più avanzate.

Per dare un po' di ossigeno all'economia, la digitalizzazione ha rappresentato storicamente - e rappresenta tutt'oggi - un vero e proprio asset strategico in grado di ricevere l'approvazione entusiastica ed immediata dell'opinione pubblica, di larga parte degli intellettuali e della stampa, senza, però, riscuotere analoga attenzione da parte degli addetti ai lavori alle prese con problemi, di frequente, più complessi e non legati alla semplice trasformazione digitale delle loro attività ordinarie. Con ciò non si vuole assolutamente esprimere una posizione preconcepita contraria alla

digitalizzazione. Al contrario, si vuole mettere l'accento sulle contraddizioni di una progettualità di corto respiro che, sebbene condivisibile nelle intenzioni, non si pone l'obiettivo di promuovere una diversa modalità di gestione, tutela e valorizzazione di un ingente patrimonio culturale composto da aree e siti archeologici, musei grandi e piccoli, collezioni nazionali e civiche, frutto di acquisizioni realizzatesi nel corso degli anni.

Per superare le incoerenze dei piani strategici, periodicamente approvati, si deve incidere parallelamente su due piani: incrementare le risorse ordinarie assegnate alle istituzioni che si occupano della tutela dei beni archeologici e culturali e, nello stesso, formare una classe di archeologi che abbiano specifiche competenze nell'ambito della progettazione e nell'uso di sistemi digitali.

Il mio nuovo lavoro sui dati digitali in archeologia non avrà, certamente, il potere di condizionare chi governa ad accrescere le dotazioni finanziarie delle Soprintendenze, dei Musei o dei Parchi, magari con programmi pluriennali di investimenti anche in assunzioni di personale scientifico. Più sommessamente con questo nuovo scritto, mi propongo di contribuire alla preparazione dei futuri professionisti responsabili della salvaguardia e della valorizzazione del nostro patrimonio archeologico. Non è certamente mia intenzione formare il lettore in una specifica competenza informatica; il mio obiettivo, meno ambizioso, punta a proporre una panoramica ragionata sulle soluzioni informatiche consolidate e quelle ancora in fase di sperimentazione per affrontare efficacemente una serie variegata ed articolata di compiti.

Per queste ragioni, un nuovo libro sulle metodologie digitali in archeologia nasce per fornire risposte aggiornate alla crescente domanda di comprensione su ciò che l'informatica mette oggi a disposizione degli archeologi e, soprattutto, su come sfruttare al meglio le opportunità che la scienza digitale propone in generale per il mondo delle Humanities. L'intento è ancora quello di porre al centro dell'attenzione degli archeologi le questioni del metodo, proponendo una Archeologia Digitale come disciplina contraddistinta da un processo non lineare, nel quale avanzamenti tecnologici si mescolano a pause riflessive e ripensamenti teorici e metodologici. L'obiettivo, più modesto, è quello di stimolare un ragionamento sull'evoluzione dei metodi in archeologia, evoluzione che si sviluppa anche grazie al contributo delle tecnologie digitali. La speranza è

che gli archeologi comprendano in che modo adattare gli avanzamenti della conoscenza, non solo in termini di nuove strumentazioni, ma soprattutto di obiettivi e programmi da realizzare nell'ambito di progetti realmente interdisciplinari che evitino forme di colonizzazione culturale.

Il primo capitolo è dedicato ad una descrizione dell'Archeologia Digitale, una nuova disciplina che lentamente, pur tra mille resistenze, emerge con l'obiettivo di approfondire il legame tra tecnologie informatiche e archeologia. Questo campo di studi, che rientra tra le metodologie di ricerca archeologica, si propone di analizzare il ruolo che gli approcci computazionali hanno nella costruzione del dato e dell'informazione archeologica, evidenziando la progressiva importanza assunta dalla conoscenza negli scenari dell'OS.

Il secondo capitolo introduce il lettore al tema del trattamento dell'informazione testuale in riferimento, soprattutto, allo sviluppo delle tecnologie semantiche e ai più recenti progressi finalizzati alla condivisione e accessibilità delle risorse digitali. Mentre nuovi strumenti consentono di pubblicare velocemente e in modo semplificato i dati in rete raggiungendo qualsiasi tipo di utenza, numerose raccomandazioni e linee guida sono definite per promuovere l'accessibilità, il riuso e la long-term preservation degli archivi. Il tema era stato già esaminato nel volume del 2006 e molte delle traiettorie delineate allora hanno conosciuto, di recente, una rapida esplosione sebbene la tesi di fondo, e cioè la condivisione, resti l'obiettivo finale dei nuovi paradigmi semantici.

Nel terzo e quarto capitolo viene illustrato in che modo le acquisizioni del Web Semantico siano state recepite dal mondo della ricerca archeologica e come la stessa disciplina si stia avvantaggiando di una tale trasformazione tecnica e, soprattutto, metodologica. Nuove modalità di registrazione, standardizzazione e trattamento dell'informazione archeologica di tipo testuale sono messe a punto per rendere concreti gli scenari innovativi delineati dall'AI e dai BD. L'uso di formati interoperabili e la disponibilità di piattaforme e infrastrutture di ricerca online, consentono di integrare archivi testuali e spaziali conformi alle linee guida del Semantic Web moltiplicando potenzialmente le possibilità di esaminare e comparare differenti archivi archeologici alla scoperta di nuove aggregazioni e significati.

L'ultimo capitolo delinea il ruolo futuro della Archeologia Digitale verso una nuova forma di ricerca archeologica nella quale archeografia

e archeologia si ricongiungono e nella quale scoperta del dato, scrittura del record e interpretazione dell'informazione sono parti essenziali di unico processo conoscitivo. Il fenomeno degli Open Data contribuirà ad aumentare la consapevolezza sui benefici di una informazione condivisa, mentre la disponibilità di dati, scaricabili ed agevolmente riutilizzabili, promuoverà forme innovative di co-working e lavoro a distanza.

Rispetto alle riflessioni pubblicate nel 2006, il quadro attuale sembra caratterizzato da una maggiore coscienza culturale e scientifica da parte degli attori in campo che riconoscono il rilievo che nel ragionamento archeologico hanno i dati e le tecnologie digitali.

A tutti i colleghi, amici e frequentatori – anche occasionali - del Centro Interdipartimentale di Servizi di Archeologia (CISA) dell'Università di Napoli L'Orientale va il mio speciale ringraziamento. Ho sempre concepito la ricerca storico-archeologica come l'esito di una esperienza nata da un confronto collettivo, disciplinare ed interdisciplinare, e non come il risultato originale dell'intuito del singolo. Per questi motivi il debito di riconoscenza per quanti, professori, ricercatori e studenti, hanno accompagnato il mio percorso, soprattutto negli ultimi anni, è particolarmente significativo. Sento anche di dover esprimere un sentimento di gratitudine per i Presidenti del CISA che si sono succeduti negli anni dopo la comparsa del mio primo volume, dal compianto Rodolfo Fattovich, a Fabrizio Pesando, Bruno Genito e, infine, Roberta Giunta e ai Direttori del Dipartimento Asia, Africa e Mediterraneo, Roberto Tottoli, Michele Bernardini e Andrea Manzo, per il loro costante sostegno, supporto ed incoraggiamento anche nei momenti meno esaltanti della mia vita professionale e di ricerca.

Nel mio film preferito, che racconta la storia della missione dell'Apollo 13 lanciato nello spazio nel 1970, i piloti dell'astronave e i tecnici della base di Houston sono impegnati nel programmare il rientro della navicella sperimentando una procedura mai presa in considerazione in precedenza. Alcuni scienziati sono concentrati nel disegnare una rotta sfruttando le orbite e la gravità terrestre e lunare; nello stesso tempo, esperti studiano come adattare i filtri di stoccaggio dell'anidride carbonica; infine, tecnici mettono a punto le procedure per assicurare il corretto angolo di incidenza necessario all'ingresso nell'atmosfera terrestre ed evitare il surriscaldamento dell'abitacolo. In quei frangenti, in cui si

manifestano situazioni imprevedibili e apparentemente irrisolvibili, entra in azione la creatività della scienza, una forma straordinaria e geniale di adattamento agli imprevisti che consente di affrontare problemi nuovi con soluzioni mai prima immaginate. Nel momento più drammatico, in cui il sapere degli scienziati non sembra più in grado di dominare le leggi della natura e gli eventi sembrano condannare gli astronauti ad un triste destino, emerge una insolita intelligenza collettiva che salva l'equipaggio. Questa attitudine alla ricerca interdisciplinare, talvolta portata avanti anche in forma eretica, ha rappresentato il riferimento del mio modo di osservare la realtà e ricavarne esperienze.

A Rosario Valentini, Angela Bosco, Gilda Ferrandino, Laura Carpentiero, Francesca Forte, Eleonora Minucci va il mio più sincero ringraziamento per le costanti discussioni e riflessioni sui temi della metodologia informatica in archeologia e, più in generale, della formazione universitaria. Infine, un grazie speciale ad Alexia Pavan per aver guardato e riguardato il testo ed avermi aiutato nel costruire al meglio l'impostazione del volume. Senza il suo prezioso contributo il mio lavoro sarebbe stato certamente meno incisivo.

# Capitolo 1

## L'ARCHEOLOGIA DIGITALE

### 1.1. Introduzione

Nell'illustrare le ragioni che mi avevano spinto nel 2006 a pubblicare un libro sui metodi informatici in archeologia (D'Andrea 2006), avevo sottolineato quanto arduo fosse questo compito sia per la complessità del tema proposto che per le profonde implicazioni sul versante dell'innovazione metodologica e delle interazioni interdisciplinari. Per sostenere i motivi di quella scelta, alquanto singolare nel panorama italiano, mi ero soprattutto concentrato sul tema della formalizzazione e della standardizzazione, due passaggi particolarmente complessi nel processo di elaborazione computazionale. Il volume offriva un panorama piuttosto variegato di tecniche e metodi per migliorare la descrizione, l'organizzazione e l'elaborazione dei dati.

Nel dedicare, dopo molti anni, un nuovo contributo ai metodi digitali in archeologia, ho avvertito l'esigenza di aggiornare quanto a suo tempo pubblicato, consapevole che offrire al lettore un quadro esaustivo è impresa quasi impossibile considerando la rapida evoluzione dei metodi computazionali in archeologia. Per questo motivo un nuovo libro sui dati in archeologia ha l'obiettivo di offrire un approfondimento su questioni di metodo e su ambiti applicativi che oscillano tra la tradizionale pratica archeologica e orizzonti sperimentali collegati al nuovo paradigma dei BD e delle tecniche di *machine o deep-learning*.

Agli inizi del 2000 temi come thesauri, standard, metadati, ontologie apparivano estranei ad un dibattito concentrato sull'applicazione di strumenti digitali a supporto della ricerca e della comunicazione. Con l'affermazione dello scenario dell'OS anche gli archeologi sono stati obbligati a fare i conti con una tecnologia che impone una riflessione disciplinare che ha origine, in primo luogo, in un diverso atteggiamento culturale verso la conoscenza condivisa e i risultati della ricerca.

Sebbene un tale processo di revisione sia stato da tempo avviato, la strada da percorrere è ancora lunga. L'assenza di organismi sovra-inter-



nazionali non semplifica l'adozione di modelli di documentazione comuni e, soprattutto, integrabili (*infra* 171). La Commissione Europea nel novembre del 2021 ha adottato una delibera (Raccomandazione 1970) finalizzata alla creazione di uno spazio comune europeo di dati per il patrimonio culturale; alcune delle sollecitazioni contenute nel documento evidenziano come la mancanza di specifiche competenze digitali stia rallentando, non solo l'avvio di un più duraturo processo di digitalizzazione dei beni culturali, ma anche lo sviluppo di politiche indirizzate alla conservazione digitale, cioè alla cura dei documenti digitali prodotti nel passato che devono essere utilizzabili e comprensibili nel futuro. La Commissione, nello stesso tempo, ha sottolineato l'importanza della tutela del diritto d'autore anche nel settore dei beni digitali e il ruolo che gli Stati membri devono avere nel garantire la piena attuazione delle norme in materia di riconoscimento del diritto di paternità della scoperta.

Le linee guida suggerite dal documento non vanno oltre le esortazioni per la predisposizione e l'aggiornamento di una strategia digitale globale che dovrebbe fissare obiettivi di digitalizzazione e conservazione digitale. Per assicurare la condivisione e l'interoperabilità delle risorse online, la Commissione si limita ad invitare gli istituti di tutela del patrimonio culturale ad aderire agli standard di contenuto e di metadati della biblioteca digitale EUROPEANA (*infra* 130).

Nonostante i recenti orientamenti sostenuti anche dalla Commissione Europea, il corpus della conoscenza archeologica digitale resta suddiviso tra tante differenti istituzioni, enti e università, e appare ancora frammentario a causa della mancanza di soluzioni in grado di assicurare il corretto scambio di dati, mentre le pratiche archeologiche, orientate alla acquisizione ed interpretazione del dato *digital-born*, risultano, in gran parte, ancora scarsamente codificabili.

L'evoluzione delle tecnologie 3D, lo sviluppo di un settore di ricerca destinato all'impiego di algoritmi di AI per l'analisi dei dati culturali, il *cloud computing*, le tecnologie semantiche, la realtà virtuale e la realtà aumentata, creano nuove opportunità per la digitalizzazione, l'accesso online e la long-term preservation. L'uso di strumenti digitali nelle attività ordinarie contribuisce alla modernizzazione dei processi che diventano così più efficienti e integrati, ma la strada verso una archeologia 3.0 è tutta in salita e piena di contraddizioni. Malgrado ciò, il cambiamento

è divenuto irreversibile e i costanti avanzamenti tecnologici rendono lo sfondo dell'OS un obiettivo maturo anche per la ricerca archeologica.

### 1.2. L'Archeologia Digitale come disciplina storica

L'affermazione di un modello scientifico basato sui dati (D'Andrea 2016; Huggett 2022) piuttosto che sui modelli, ha contribuito a porre le basi per un cambiamento di paradigma che deve essere correttamente esaminato anche per gli esiti che esso ha sulla ricerca archeologica. La complessità raggiunta dalle tecniche informatiche richiede una specifica competenza in ciascun dominio affinché siano corrette le procedure per la registrazione, il trattamento e la pubblicazione dei dati. In campo archeologico è l'Archeologia Digitale ad occuparsi di questo insieme di tematiche, tecnologiche e di metodo, che investono i vari momenti e i differenti aspetti della ricerca storica nel terzo millennio. Specifici settori dell'Archeologia Digitale sono la fotografia e il disegno digitale, il rilievo e il relativo trattamento dell'informazione geometrica, i GIS, i BIM, i sistemi CAD fino a comprendere l'archeologia virtuale con il suo peculiare ambito applicativo rappresentato dallo sfruttamento del dato archeologico a fini prevalentemente divulgativi.

In questo nuovo dominio sembrano affermarsi due differenti prospettive o punti di vista strategici in parte inconciliabili: quello che considera il digitale come una semplice tecnica, equiparabile ad altri procedimenti, tra i quali i metodi al radiocarbonio o le prospezioni geofisiche, che servono per datare un oggetto o proporre una ricostruzione ambientale. Questi strumenti sarebbero a-teoretici e potrebbero essere utilizzati da qualsiasi archeologo indipendentemente dalla sua formazione. Una tale considerazione spingerebbe a concludere che gli sviluppi in campo digitale siano paragonabili a quelli di una altra qualsiasi tecnica senza alcun rapporto con le trasformazioni del metodo. L'altro punto di vista, invece, sostiene che l'evoluzione del digitale ha importanti conseguenze sul metodo enfatizzando, in questo modo, la stretta connessione tra strumenti informatici e obiettivi di ricerca (Orlandi 1997; Lock 2003; Zubrow 2006).

A queste due contrapposte visioni se ne può aggiungere una terza che parte dalla semplice constatazione che il digitale rappresenta per tutti gli studiosi una opportunità concreta di trasformazione del modo di fare ricerca e che gli strumenti digitali sono divenuti parte integrante del casset-

to degli attrezzi dell'archeologo. Una tale posizione, meno estrema e più permeabile alle varie sollecitazioni che provengono da domini tecnici e/o filosofici, cerca invece di esaminare in quali forme l'impiego di sistemi informatici e di dispositivi digitali stia modificando l'agire dello studioso e di come la stessa evoluzione tecnologica imponga una costante revisione delle pratiche tradizionali di ricerca. Non è estranea a questo processo di revisione una più ampia riorganizzazione dei saperi all'interno di una società in rapida trasformazione che modifica e adegua la propria organizzazione delle conoscenze e della sua azione sociale e del suo tempo di ricerca alle nuove sfide create dalla diffusione delle tecnologie digitali.

L'Archeologia Digitale si propone come una disciplina, o meglio un campo di interesse, in movimento, in cui convivono varie sfaccettature e punti di vista, talvolta alternativi, e limiti settoriali e di pensiero non ancora ben definiti. Questo sapere oscilla tra l'analisi degli strumenti adottati, gli esiti della ricerca e, infine, lo studio dell'interazione tra archeologia e società contemporanea. In particolare, l'Archeologia Digitale interviene nella ricerca evidenziando l'apporto del digitale nel processo di conoscenza e segnalando errori o inconsistenze nelle fasi di identificazione e digitalizzazione del record archeologico. Il tema non può essere disgiunto da quello, oramai centrale, della cura, della manutenzione e della preservazione degli archivi. Da anni si assiste ad una pubblicazione online dell'informazione, differenziata in base ai tipi di media scelti; la separazione degli archivi di testi da quelli fotografici costituisce un esempio della specializzazione delle fonti digitali che, in questo modo, perdono i nessi con le risorse in rete potenzialmente associabili.

La maggiore considerazione per i dati grezzi sta spostando l'interesse degli studiosi dalla pubblicazione dei media (foto, disegni, ricostruzioni, etc.), necessari alla illustrazione del contenuto scientifico edito, alla creazione di strumenti (repository, piattaforme, infrastrutture) dove depositare, condividere e mantenere gli archivi digitali. Gli standard, le ontologie, i metadati e, in generale, le tecnologie e i linguaggi semantici rendono concreto uno scenario che tragherà nei prossimi anni l'archeologia da scienza circoscritta ad un ambito ristretto di esperti ad un sapere ampio e partecipato.

Se per alcuni anni l'espressione per definire questa area di interesse è oscillata tra archeologia computazionale, informatica archeologica, me-

todi informatici per l'archeologia, sembra, oggi, che il termine Archeologia Digitale sia entrato nel lessico comune e anche negli insegnamenti dei corsi universitari, riempiendo in questo modo una grave lacuna del nostro sistema formativo; non si tratta di una disciplina orientata a fornire una specifica competenza tecnica nell'uso di un determinato software, quanto di un'area di studi metodologici finalizzata a testare l'uso delle tecniche digitali nella ricerca archeologica e a fornire strumenti per la loro validazione allo scopo di contribuire ad una proiezione del ruolo dell'archeologia del prossimo futuro.

Sulla definizione della disciplina è intervenuta, di recente, P. Moscati (2019) che ha sottolineato come il termine Archeologia Digitale rifletta una connotazione tecnologica degli studi archeologici nei quali l'attenzione alla metodologia di analisi è sostituita dall'approfondimento delle tecniche e degli strumenti computazionali. Come vedremo più avanti parlando della *digital transformation* e del *digital twin*, una corretta valutazione degli scenari inaugurati dall'Archeologia Digitale può essere tracciata partendo dalla constatazione che la ricerca si trasforma parallelamente all'evolversi delle scoperte e delle nuove tecnologie e che anche l'archeologia subisce le influenze del mondo circostante.

Mentre gli sviluppi della disciplina spingono a ritenere che tutti siano oramai archeologi digitali da almeno un ventennio (Evans, Daly 2006), la pratica quotidiana e soprattutto l'evoluzione della tecnologia richiede una riflessione su quanto la dimensione digitale dell'agire dell'archeologo influenzi il record, la ricostruzione e, infine, l'interpretazione storica. Inoltre, è allo stesso modo importante sviluppare un ragionamento che includa una previsione sulla conservazione di un archivio codificato in base a formati interoperabili e a strutture formali facilmente migrabili ed aggiornabili in sistemi più moderni. L'Archeologia Digitale non si occupa soltanto del dato digitale e del suo migliore impiego sul terreno della ricerca, ma anche della sua salvaguardia e manutenzione fisica e, soprattutto, della sua accessibilità e riuso nell'ottica di un sapere sempre più condiviso e che travalica i confini della ricerca storica per estendere il suo impiego anche in ambiti differenti.

Il rapporto con la società e le sue rapide modifiche e trasformazioni è un'altra delle aree di studio dell'Archeologia Digitale che, tra i suoi compiti, pone anche quelli relativi ad una maggiore integrazione della

disciplina con le esigenze di una società nuova. L'archeologo non può più vivere confinato nella sua ricerca e nella sua torre digitale; i prodotti del suo lavoro devono essere sottoposti ad un giudizio, prima di tutto scientifico, e poi collettivo, considerati i costi spesso elevati dei progetti di archeologia sul campo.

Definire in dettaglio i campi di interesse dell'Archeologia Digitale è un compito arduo, a causa della impossibilità di circoscrivere con precisione i confini del digitale. Tuttavia, per tentare di delineare delle aree specifiche di intervento possiamo riprendere i temi trattati nel simposio dal titolo *Digital Archaeological Heritage* organizzato a Brighton nel 2016 dall'EAC (May 2017). In particolare, le sessioni si sono focalizzate su tre tematiche specifiche:

- Misurare e percepire: orientata ad esplorare come percepiamo, identifichiamo e caratterizziamo il paesaggio archeologico, i siti e i reperti;
- Dai dati alla conoscenza: destinata a illustrare come sia possibile estrarre nuova conoscenza dai dati grezzi;
- Visualizzare il passato: diretta a mostrare come il mondo antico sia oggi presentato attraverso le tecnologie digitali.

Mentre la prima sessione è stata dedicata ad evidenziare il potenziale di uso dei GIS nella comprensione del paesaggio antico, il secondo panel ha mostrato come l'implementazione di infrastrutture e reti in grado di assemblare dataset eterogenei possa contribuire a produrre nuova conoscenza usando algoritmi di *machine-learning*; infine, nella terza sessione, la discussione si è concentrata sulle innumerevoli modalità di visualizzazione del paesaggio antico tramite tecniche di realtà virtuale, aumentata o mista, e sugli strumenti che consentono di offrire ad un pubblico inesperto la possibilità di interagire con mondi antichi ricostruiti digitalmente; questa parte dell'incontro ha affrontato il tema, altrettanto delicato, del coinvolgimento del pubblico nella fruizione del passato e sulla dimensione social dell'archeologia pubblica (May 2017).

La tecnologia digitale ha avuto un impatto significativo su alcune aree della ricerca archeologica consentendo, soprattutto, di costruire una rappresentazione completa del mondo reale antico e moderno. Nel-

lo stesso tempo, il digitale può manipolare e aggregare dati per simulare processi di interazione uomo-ambiente e di trasformazione.

Le sfide poste dall'Archeologia Digitale non riguardano soltanto una riflessione metodologica sull'impiego di strumenti informatici in archeologia, ma comprendono anche lo studio di forme innovative di registrazione, acquisizione e processamento dei dati che includano aspetti relativi all'uso di appropriate metodologie di raccolta e controllo delle informazioni che oltrepassano i tradizionali confini delle competenze professionali archeologiche (Zubrow 2006). Non si tratta di trasformare una pratica normale in un sistema complesso di procedure digitali, sebbene innovative, quanto di comprendere come quella tecnologia e soprattutto il suo impiego possano modificare l'agire del ricercatore sul campo o in laboratorio e come questo cambiamento sia riflesso nei dati acquisiti.

Se il rischio è di re-ingegnerizzare la metodologia archeologica rendendola, quasi, subalterna al codice informatico, il pericolo maggiore, che si annida in un tale processo di ammodernamento delle procedure di raccolta dati, è l'uso acritico di uno strumento adoperato senza conoscere le potenzialità e soprattutto limiti funzionali e/o tecnici. Un tale atteggiamento può produrre effetti dannosi soprattutto se si pensa alla dimensione pubblica e aperta delle informazioni che circolano in rete, talvolta prive di ogni necessaria descrizione relativa ai processi di digitalizzazione e trattamento del record.

L'uso di strumenti innovativi modifica il rapporto con l'oggetto delle nostre indagini (Huggett 2017) e, in modo particolare, la relazione cognitiva tra soggetto e dato osservato; anche il modo di analizzare e ricostruire il passato muta profondamente (Huggett 2004; 2012). Capire come le tecnologie operano su di noi e per noi conduce inevitabilmente a comprendere meglio il funzionamento di quello strumento e di quello specifico algoritmo e, soprattutto, consentono di rimuovere l'idea che, in fin dei conti, l'informatica sia un elettrodomestico che estende le nostre facoltà corporee.

Per Hodder (2012) uno strumento digitale contiene in sé una conoscenza scientifica che in un qualche modo si riflette nell'oggetto acquisito finendo con il condizionare ogni ricostruzione archeologica; i dispositivi di misura (laser scanner, stazione totale, GPS, GNSS) incorporano un concetto di misura legato alle proprietà fisiche dell'oggetto e algorit-

mi di calcolo che identificano e conservano la geometria dell'informazione acquisita. La Post-Processual Archaeology ha contrastato fortemente l'introduzione del digitale nella ricerca motivando il rifiuto con l'idea erronea che il digitale in fin dei conti portasse nell'analisi dei dati una descrizione in termini esclusivamente numerici della realtà; se la Post-Processual Archaeology è riflessiva, decostruttiva e interpretativa, allora il digitale è analitico e deterministico.

Gli aspetti che legano pratiche e sistemi cognitivi (Heersmink 2012) trasformano le capacità di fare ricerca e orientano verso specifici ambiti di conoscenza scientifica. Pertanto, l'Archeologia Digitale deve necessariamente indagare i dispositivi, il funzionamento, l'impiego e i risultati ottenuti per tentare di delineare in che modo le procedure digitali cambiano, o alterano, le forme di scoperta ed esplorazione del passato. Poiché l'attività dell'archeologo ha prodotto nel tempo dati, metodi, teorie e ricostruzioni, l'Archeologia Digitale deve anche interrogarsi su come gli archivi prodotti nel passato possano integrarsi senza sovrapporre in modo distorto la conoscenza implicita presente in un dataset strutturato in un certo periodo con quella acquisita in precedenza o dopo. Il caso più evidente è quello della documentazione di scavo registrata in una fase successiva all'introduzione del metodo stratigrafico che non può automaticamente essere allineata agli archivi registrati con tecniche di scavo differenti.

Come per l'indagine stratigrafica, che è guidata dall'esperienza del ricercatore che sa come identificare e rimuovere gli strati, anche gli strumenti digitali sono adoperati seguendo un workflow definito dall'utilizzatore; non c'è nulla di automatico nell'impiego di una stazione totale, di un laser scanner o di un drone. Le attrezzature per il rilievo digitale registrano i dati cognitivi in base ad una conoscenza interna alla tecnologia impiegata, ma il setup degli strumenti e la progettazione dell'intervento restano nelle mani del topografo o del disegnatore. Se le tecnologie spaziali producono una misura oggettiva entro un range noto di errore strumentale, l'accuratezza finale del rilievo sarà in funzione dell'uso appropriato dei dispositivi da parte dell'utente. Nella produzione del dato e della relativa conoscenza, la componente umana gioca, dunque, un ruolo cognitivo centrale che si trasferisce sull'oggetto osservato.

Analogamente agli strumenti per il rilievo, anche il computer possiede una specifica visione del mondo costruita su algoritmi che gestiscono

e aggregano oggetti digitali e che si proietta sui dati raccolti e memorizzati. L'esempio più lampante di un approccio *digital-driven* è, forse, fornito dal GIS che impone allo studioso una lettura deterministico-ambientale, basata su una impostazione prevalentemente geometrica del paesaggio e delle azioni umane (*infra* 135).

L'Archeologia Digitale si muove con l'obiettivo di identificare le aree di criticità nell'applicazione di strumenti digitali alla ricerca storica, non limitando il suo raggio di azione ad uno specifico settore di intervento (database, semantica, statistica, rilievo, GIS, realtà virtuale), ma anzi estendendolo a pratiche e metodologie interdisciplinari che comprendono la teoria archeologica e le tecniche informatiche.

La centralità del dato nella ricerca deve imporre ai ricercatori una maggiore standardizzazione delle risorse digitali e una loro semplice e immediata pubblicazione online; la creazione di infrastrutture o repository può accelerare questo processo di condivisione e circolazione dei dati innescando pratiche virtuose e, soprattutto, incentivando forme di co-working. Senza una convinta adesione a questi principi il rischio è di frammentare ulteriormente la conoscenza o parti di essa attraverso l'adozione di sistemi di registrazione e disseminazione che seguono canali individuali che finiscono per sottrarre informazioni all'intera comunità scientifica.

L'informatica come scienza del trattamento automatico dell'informazione può fornire un contributo determinante nella cura degli archivi o nel recupero di essi, ma per fare ciò, secondo, K. May (2017) l'archeologo deve:

- identificare i target del pubblico scientifico;
- comprendere le modalità di uso e riuso dell'informazione;
- ridefinire l'uso della letteratura grigia pubblicandola in rete con appropriati DOI;
- rendere i dati facilmente leggibili;
- promuovere la condivisione dei dati per rendere più comprensibili eventuali sintesi dei risultati ottenuti nelle ricerche;
- esplorare nuovi metodi per fornire esempi di archivi digitali online.

In parte, le suggestioni formulate da May sono contenute all'interno dei FAIR Principles (*infra* 80). Il tema di una maggiore attenzione alla strut-



turazione e al destino dei dati digitali si collega anche al riuso dei *legacy data* (*infra* 32) e alla messa a punto di strategie per la conversione e il trattamento di questo ricco patrimonio di informazioni e dati ereditati che spesso è inutilizzato a causa della obsolescenza della tecnologia con la quale è stato codificato.

Senza voler limitare il campo di riflessione e sperimentazione, questo nuovo settore di studi che definiamo Archeologia Digitale:

- promuove linee guida per la corretta digitalizzazione dei dati testuali, spaziali, geografici e rilevati;
- esamina le caratteristiche di nuove strumentazioni e software e ne testa le potenzialità di uso in campo archeologico;
- sostiene e migliora la comunicazione dei risultati della ricerca tra gli studiosi e il grande pubblico favorendo e supportando approcci aperti;
- identifica le pratiche per standardizzare archivi e dati attraverso l'uso di tecnologie semantiche e di portali, infrastrutture o repository.

Se i confini dell'Archeologia Digitale non sono ancora ben delineati, il campo di indagine e le tecniche di ricerca restano ancorati alla capacità dei ricercatori di sviluppare una riflessione sulla metodologia archeologica e sull'influenza che i metodi informatici hanno sugli esiti della ricerca storica. L'incertezza lascia il campo aperto ad incursioni provenienti da esperti che appartengono ad altri domini. Ingegneri, architetti, informatici e geologi trasferiscono nella ricerca archeologica competenze di tipo tecnico consolidando l'idea che, in fin dei conti, l'evoluzione della tecnologia non abbia conseguenze, se non limitate, sul modo di operare dell'archeologo. Questo atteggiamento ha esiti negativi non soltanto sul fronte della formazione dei giovani archeologi, ma anche sulla sostenibilità degli interventi, perché, trascorso il periodo legato alla affermazione di una nuova tecnica, mancano professionisti in grado di garantire la sopravvivenza a regime di quel nuovo modo di fare ricerca e produrre dati scientifici.

A questo particolare obiettivo è dedicata l'Archeologia Digitale, il cui contenuto disciplinare rientra nel più ampio settore delle metodologie di ricerca in archeologia, un ambito nel quale confluiscono pratiche e metodi eterogenei che vanno dallo scavo alla classificazione, fino ai più

recenti studi sull'archeologia pubblica. L'Archeologia Digitale si propone di approfondire tematiche come la gestione/descrizione di una collezione di oggetti digitali, lo studio delle relazioni inter e intra-site eseguite mediante tecnologie spaziali, e, infine, la costruzione di scenari per l'archeologia virtuale e/o immersiva.

### 1.3. Dalla *digitization* alla *digital transformation*

Per comprendere al meglio il ruolo e la funzione di questa nuova disciplina, o campo di interessi, dobbiamo allargare la prospettiva oltre lo spazio ristretto delle logiche dell'insegnamento universitario, e considerare le trasformazioni sociali e culturali che attraversano il mondo dei beni archeologici e culturali e, più in generale, la sfera delle scienze. Possiamo così intuire che la ricerca, indipendentemente dallo specifico ambito di studi, mette al centro della propria riflessione disciplinare il tema della *digital transformation*, cioè dello sviluppo di strumenti e tecnologie digitali per lo svolgimento delle ordinarie attività lavorative e di studio. Questo ripensamento epistemologico, che sta investendo i settori comunemente appartenenti alle Humanities, non è soltanto il risultato di una progressiva digitalizzazione delle fonti a disposizione del ricercatore che abbandona la tradizionale biblioteca a vantaggio del Web, ma anche l'effetto dell'impiego di algoritmi che aggregano le informazioni e che modificano, nello stesso tempo, metodi di analisi e attitudini degli studiosi.

Che il cambiamento fosse epocale e irreversibile era già prevedibile nella fase iniziale della digitalizzazione che ha prodotto la conversione numerica di consistenti archivi destinati all'oscurità o alla distruzione. Un notevole incremento dei dati digitali a disposizione del ricercatore si è accompagnato allo sviluppo di specifiche applicazioni computazionali, originate da quel fenomeno descritto da F. Djindjian (1996) come "l'informatica senza informatici", una fase storica nella quale emergeva da parte dei ricercatori un atteggiamento pragmatico che poneva l'accento sul calcolatore, considerato un nuovo e più potente elettrodomestico, piuttosto che uno strumento che modificava la stessa metodologia archeologica.

Oggi, dopo la fase della digitalizzazione, si riflette su come lo sviluppo tecnologico abbia cambiato il modo di fare ricerca e cultura e, parallelamente, si cerca di circoscrivere al meglio gli obiettivi dei distinti aspetti della digitalizzazione e di definire in modo non più ambiguo la terminologia.

Se l'espressione digitizzazione identifica il passaggio analogico-digitale di una qualsiasi fonte audio, video, immagine e testo, è più difficile chiarire il significato di digitalizzazione e transizione digitale. I concetti sono profondamente differenti e non devono essere confusi: mentre la digitizzazione si riferisce ai dati e alle informazioni, con il termine digitalizzazione si indicano i processi, e, infine, con trasformazione digitale si definisce un cambiamento sistemico dei modelli industriali, di business, ma anche formativi e culturali.

L'introduzione di nuovi formati per sostituire gli archivi cartacei con documenti digitali ha modificato la modalità di memorizzazione, accesso ed elaborazione dei dati, rappresentando una rivoluzione nella costruzione e gestione dell'informazione e mettendo per la prima volta in connessione il mondo fisico con il computer. La conversione analogico-digitale, nata negli anni '60 e, talvolta, eseguita in forma automatica o semi-automatica, ha caratterizzato la digitizzazione che ha consentito di creare numerosi archivi digitali, trattati, però, con i tradizionali processi preesistenti. In campo archeologico si è assistito alla conversione della documentazione (foto, diari, disegni), mentre soltanto a partire dagli anni '90 si è realizzata la sperimentazione di una acquisizione direttamente sullo scavo (principalmente foto e report) per un successivo trattamento in laboratorio.

L'esigenza di rendere accessibile ed utilizzabile questa massa di dati digitali è alla base dello sviluppo della fase della digitalizzazione contraddistinta dalla identificazione di nuove forme di gestione, comunicazione e sfruttamento dei dati grazie, in particolare, alla diffusione delle connessioni digitali. Le tecnologie, attraverso applicazioni e/o piattaforme, garantivano una nuova interazione tra l'utente e i dati digitali migliorando l'efficienza dei processi e stimolando in tal modo nuove opportunità. La realtà aumentata, ad esempio, utilizzando rilievi o ricostruzioni digitali realizzate per obiettivi di ricerca, produceva un nuovo oggetto digitale anche per la fruizione turistica. I GIS hanno modificato l'accesso e il trattamento dell'informazione geografica consentendo la creazione e l'esplorazione di mondi digitali, del tutto virtuali, nei quali il ricercatore poteva simulare il movimento di agenti o le trasformazioni del paesaggio antico.

La digitalizzazione ha incrementato il livello di automazione dei procedimenti attraverso un uso massiccio delle tecnologie digitali, senza

modificare, tuttavia, i processi di analisi; l'integrazione delle risorse digitali è uno dei più evidenti e utili passaggi verso un modo nuovo di utilizzare gli archivi. La digitalizzazione ha imposto un modo nuovo di raccogliere, acquisire e trattare i dati digitali producendo un ripensamento nelle strategie di ricerca e diffusione delle tecnologie informatiche per la realizzazione delle diverse fasi dell'attività scientifica. La trasformazione dei processi di ricerca ha richiesto un consistente investimento tecnologico che ha coinvolto tutti gli stadi dell'indagine archeologica sul campo e in laboratorio; nello stesso tempo, l'esigenza di preservare e condividere i dati ha incentivato l'uso di formati aperti e forme di standardizzazione della conoscenza e dell'informazione. L'implementazione di infrastrutture ha reso più democratico e semplice l'accesso ai dati e alla loro elaborazione consentendo di superare i vincoli dei confini geografici e, soprattutto, le barriere delle tradizionali competenze scientifiche.

I cambiamenti introdotti dalla digitalizzazione, dal decentramento dell'informazione e dal conseguente aumento dell'interazione digitale a tutti i livelli, non solo scientifici, ma anche culturali e soprattutto sociali, hanno favorito l'emergere delle tecnologie dell'IoT. Oggi siamo di fronte ad una nuova rivoluzione digitale nella quale un ruolo centrale hanno quelle tecnologie che supportano l'interconnessione automatica in rete di oggetti digitali identificabili ed intelligenti. Questo mutamento sta producendo un cambiamento nelle strategie di costruzione, uso e condivisione del dato digitale, mentre una nuova cultura organizzativa si è diffusa per una gestione affidabile, efficiente e duratura degli archivi. Il risultato della trasformazione radicale nei modi di pensare e organizzare il lavoro, definito più semplicemente con l'espressione *digital transformation*, rappresenta un cambiamento che spinge verso un riorientamento del modo di creare e utilizzare le risorse digitali. Lo sviluppo della sensoristica, come dell'IoT, sposta il controllo remoto dei processi dall'uomo direttamente alle macchine, mentre gli algoritmi basati sull'AI possono modificare gli stessi processi organizzativi e produttivi.

Uno dei fattori che guida le innovazioni introdotte dalla *digital transformation* è il gemello digitale, un oggetto virtuale che riproduce la realtà e la rende interrogabile e i cui comportamenti agli eventi sono simulabili. Oggetto, processo, relazioni, luoghi e comportamenti sono parti di un ecosistema digitale nel quale agiscono automi in grado di

fornire analisi, risposte e modelli predittivi di fronte a stimoli provenienti da differenti circostanze. Il gemello digitale non è la semplice digitalizzazione della realtà fisica; esso ingloba un flusso di lavoro e di procedure che contestualizzano elementi del mondo reale integrando sistemi, funzioni e dati diversi in un ciclo di vita, dalla creazione alla fine di utilizzo dell'oggetto. L'ecosistema digitale nel quale si diffondono i gemelli digitali è un ambiente nel quale si sperimentano nuovi approcci e algoritmi, come l'AI, i BD e l'IoT, ma è, nello stesso tempo, uno spazio virtuale dove interagiscono le nostre attitudini sociali, culturali, tecniche e scientifiche.

Poiché la simulazione e l'attività predittiva sono due componenti essenziali della ricerca, i gemelli digitali rappresentano la nuova frontiera della digitalizzazione, grazie alla capacità dei sistemi e delle infrastrutture di gestire in un ambiente condiviso programmi, dati, competenze ed esperienze.

L'esigenza di testare in un mondo virtuale oggetti fisici tratti dal mondo reale ha incoraggiato gli studiosi a sviluppare modalità affidabili per la creazione e il rendering grafico di accurati oggetti e spazi tridimensionali; queste repliche digitali, fedeli nella riproduzione geometrica e meccanica, sono sottoposte a esperimenti che consentono di analizzarne il comportamento in un dato contesto e a determinate condizioni. Le connessioni, che nel mondo digitale possono essere costruite dall'IoT e dagli algoritmi dell'AI, ampliano le possibilità di testare il modello e attivare una rete pressoché infinita di pratiche ed esperienze di ricerca.

La lenta e progressiva crescita delle applicazioni informatiche, e, soprattutto, della loro peculiare utilizzazione per lo svolgimento di numerosi compiti ordinari della nostra vita sociale e professionale, ha contribuito alla definizione di un nuovo modello organizzativo nel quale si impiega l'informatica per automatizzare le normali attività e si intensificano i processi lavorativi creando nuovi prodotti e ridisegnando il workflow dei servizi. Lavorare in cloud, condividere software o strumenti digitali, interagire con dati presenti in rete, sono aspetti che spingono nella direzione di trasformare una semplice navigazione in Internet in una attività scientifica ad alta intensità.

La transizione digitale richiede un radicale cambio di mentalità, soprattutto funzionale allo sviluppo di soluzioni che si basano sulla collabo-

razione. Grazie alla rete, l'annullamento delle distanze fisiche crea opportunità di lavoro e cooperazione tra scienziati, costringendo anche i nativi analogici a profondi cambiamenti di visione. La virtualizzazione del dato o la sua dematerializzazione sollecitano un atteggiamento innovativo per la registrazione, la manutenzione e il trattamento dell'informazione. Non è necessario possedere specifiche competenze informatiche, la transizione digitale mette a disposizione dell'utente straordinarie soluzioni basate sull'IoT, sull'AI e sulle tecnologie dei BD. Occorre, però, essere consapevoli che il modello di scienza ha la sua centralità sui dati e che, di conseguenza, molti dei nostri sforzi devono andare nella direzione di rendere i nostri archivi standardizzati e condivisi per accrescere la connettività non soltanto in termini di ricerca, ma anche di comunicazione e in modo particolare per rafforzare le collaborazioni a livello interdisciplinare.

Parallelamente alle trasformazioni tecnologiche, anche il mondo dell'archeologia ha conosciuto cambiamenti senza precedenti. Dai block notes cartacei siamo passati ai database, ai disegni CAD, all'implementazione di GIS, alla modellazione, alla creazione di accurate repliche 3D fino ad una serie di innovazioni, tra cui la simulazione basata su agenti. Rileggendo la storia della disciplina e dei cambiamenti di metodo a partire dagli anni '70 si ha la percezione netta di una lenta, ma costante trasformazione digitale che ha interessato tutti i settori della ricerca archeologica, sul campo e in laboratorio.

Dopo una fase in cui le innovazioni sono state ad esclusivo appannaggio di pochi istituti di ricerca e di università in grado di acquistare ed utilizzare calcolatori costosissimi, oggi la continua diminuzione dei prezzi delle strumentazioni rende la tecnologia accessibile a tutti gli studiosi. La possibilità di poter disporre di un computer, di applicativi specifici per la ricerca e di dispositivi digitali per il disegno e la fotografia ha ampliato le competenze dei singoli ricercatori e le potenzialità di indagine. La democratizzazione della tecnologia, però, non ha sempre prodotto un vantaggio immediato se si pensa alle migliaia di piccoli e grandi scavi che restano inediti o pubblicati attraverso brevi report. Poiché la crescita esponenziale dei dati (testuali, tabellari, fotografici e grafici) attende ancora una sistematizzazione e ottimizzazione, è lecito, in taluni casi, interrogarsi sui reali benefici che la tecnologia ha apportato a quella specifica ricerca.

La grande disponibilità teorica di dati rende più complessa la realizzazione di un qualsiasi progetto di studio, richiedendo una precisa strategia che consideri anche come accedere alle risorse, incluse quelle ereditate (*infra* 32). La difficoltà di districarsi in un mare infinito di dati, spesso inaccessibili e non pubblicati, è testimoniata dalla attività ultradecennale dell'ADS (Richards 2017) che ha pubblicato online e reso accessibili gratuitamente circa 64.000 report relativi a scavi eseguiti nel Regno Unito. Nella stessa direzione si è mosso l'olandese DANS che ha implementato un portale con oltre 65.000 documenti indicizzati e interrogabili attraverso il motore AGNES (Brandsen, Lippok 2021; Holander 2021). L'enfasi riservata al miglioramento delle tecniche di pubblicazione e accesso ai dati contenuti nella letteratura grigia è confermata dagli obiettivi del progetto europeo ARIADNE (Richards, Niccolucci 2019) che ha integrato i portali dell'ADS e del DANS e di altre risorse online mettendo a disposizione degli utenti circa 3.500.000 dati accessibili da un unico motore. L'incremento dei dati in rete si accompagna ad una maggiore qualità dell'informazione e degli strumenti per il recupero e l'aggregazione delle risorse digitali.

L'archivio dell'ADS conserva oltre 3,6 milioni file in 308 formati e circa 1,3 milioni di record di metadati; il sito ospita 4087 collezioni, tra cui database di manufatti, sepolture, monete, rilievi edilizi, archivi e rapporti di lavoro sul campo, set di dati scientifici e registrazioni di siti e monumenti. L'*Historic Environment Scotland's Canmore* ha superato il milione di oggetti digitali nel 2020 (McKeague 2021), mentre il *PAS for England and Wales* ha poco più di un milione di record relativi a oltre 1,6 milioni di oggetti.

Una analisi sulla evoluzione delle tecnologie e dei processi di digitalizzazione che hanno interessato la società e le scienze nel loro complesso, porta a concludere che esista una forte correlazione tra innovazioni, sperimentazione dei metodi digitali in archeologia, e, infine, adattamenti della teoria ai nuovi risultati. Come l'introduzione dei GIS ha contribuito a modificare la concezione dello spazio in archeologia e la stessa capacità di interrogare variabili geografiche, analogamente la crescente disponibilità di dati in rete incoraggia gli studiosi a creare archivi aperti e accessibili e, soprattutto, a mettere in atto strategie innovative per la registrazione dei dati che aumentano quantità e qualità dell'informazione (Opgenhaffen 2022).

Uno degli aspetti ancora poco delineati, ma, nello stesso tempo, di straordinario interesse per le sue implicazioni sotto il profilo della ricerca e della conservazione, è il ruolo del gemello digitale in archeologia. Come abbiamo visto in precedenza la replica digitale viene considerata uno dei pilastri delle innovazioni introdotte dalla *digital transformation*, per la capacità offerta dal documento digitale di riprodurre una realtà interrogabile e simulabile quando esposta a particolari condizioni. Se per alcuni ricercatori la semplice digitalizzazione accurata della realtà fisica produce automaticamente *digital twins*, va precisato che la copia di un oggetto reale diviene un gemello digitale se la risorsa include, accanto alle proprietà geometriche, anche relazioni, luoghi e comportamenti che agiscono autonomamente di fronte a stimoli esterni (Hermon *et al.* 2022).

Nell'ecosistema virtuale il ricercatore può sperimentare nuovi approcci digitali tratti dall'AI, dai BD e dall'IoT che interagiscono con la replica digitale. Il sistema che maggiormente si avvicina a questo scenario è quello dell'HBIM (Bosco *et al.* 2021), la cui tecnologia combina forme testuali, relative alla descrizione dell'oggetto, grafiche, per la identificazione delle forme geometriche e materiali, e, infine, comportamentali per la capacità del sistema di eseguire operazioni su tutti gli elementi archiviati. Una componente essenziale del processo è la creazione di un oggetto 3D esplorabile, realizzato utilizzando tecniche di modellazione oppure una delle 3D *reality-based* tecnologie disponibili (laser scanner, fotogrammetria aerea e/o terrestre, etc.). Il BIM associa alla risorsa 3D qualsiasi altro tipo di informazione testuale, tabellare, grafica, fotografica, contribuendo alla definizione di una rappresentazione digitale complessiva del record che può essere utilizzata per motivi di ricerca, conservazione, valorizzazione, gestione o fruizione. Il ricercatore può eseguire interrogazioni sia sull'informazione geometrica che su uno specifico dato cronologico e/o funzionale oppure relativo alla temperatura, all'umidità o alla muratura. A differenza del più tradizionale GIS, che opera a livello bidimensionale e territoriale, o dei vari esempi di GIS 3D (Dell'Unto, Landeschi 2022), il BIM tratta il dato geometrico/spaziale alla scala dell'edificio con un dettaglio dell'informazione che può arrivare anche alla descrizione del singolo blocchetto nel caso, ad esempio, di progetti di restauro. Poiché GIS e BIM elaborano informazioni spaziali collegate ad archivi descrittivi è stata da tempo progettata



una maggiore integrazione tra i due sistemi (Dore, Murphy 2012) e le principali industrie del software GIS sono già in grado di operare con file di interscambio IFC generati dai software BIM riducendo la distanza concettuale e applicativa tra i due sistemi.

La tecnologia *digital twin* è ancora poco utilizzata nel mondo dei beni culturali sebbene le potenzialità di impiego al costruito storico risultino evidenti, in particolare, per la capacità del sistema di conservare e reimpiegare qualsiasi altro tipo di fonte informativa coinvolgendo una pluralità di specialisti in un unico progetto. Il settore che può trarre il più ampio vantaggio dall'impiego del BIM nella sua versione HBIM o ABIM o Archeo-BIM (Bosco *et al.* 2018a; Garagnani *et al.* 2021) è certamente quello della progettazione ed esecuzione del restauro e della manutenzione, ma non mancano significativi esempi di soluzioni, basate sull'impiego di ontologie di dominio, che mostrano come collegare in modo interattivo i mondi virtuali e fisico del gemello digitale (Niccolucci *et al.* 2022; Tan *et al.* 2022).

La trasformazione digitale contribuisce a modificare l'approccio ai beni archeologici obbligando, nello stesso tempo, i ricercatori ad una rivisitazione delle metodologie archeologiche soprattutto alla luce dell'introduzione, oramai in forma quasi stabile, dei metodi digitali. Per sfruttare appieno le potenzialità di cambiamento implicite nella *digital transformation* è, però, necessario cambiare quell'atteggiamento di rifiuto aprioristico nei confronti delle innovazioni o, peggio, ogni forma di adozione acritica dell'ultima tecnologia al momento disponibile. La trasformazione digitale richiede:

- una adeguata alfabetizzazione informatica destinata alla produzione di dati affidabili e accessibili;
- una buona educazione digitale che consenta di utilizzare in modo consapevole gli strumenti messi a disposizione, spesso gratuitamente
- e, infine, un comportamento civico che sia in grado di rispettare la privacy e, soprattutto, di riconoscere i diritti d'autore dei produttori di informazioni presenti in rete.

#### 1.4. I *legacy data*

Numerosi interventi sono stati realizzati per costruire sistemi in grado di rendere accessibili i dati online. Nonostante siano stati prodotti mol-

teplici sforzi in questa direzione, un'area che resta ancora poco trattata nella condivisione è quella dei record predigitali o acquisiti con i primi dispositivi e memorizzati in forme embrionali di database. L'esigenza di gestire al meglio questa ricca documentazione, spesso inaccessibile e non utilizzabile in più moderne applicazioni, ha spinto gli studiosi a sviluppare sistemi per recuperare archivi di dati di vecchie indagini o scavi, talvolta dimenticati o divenuti nel frattempo obsoleti.

I dati, definiti *legacy* ovvero ereditati, sono in gran parte un prodotto dell'Archeologia Digitale poiché la velocità dei processi di digitalizzazione rendono in poco tempo già vecchi dati registrati pochi anni fa; una tale rapida obsolescenza delle informazioni registrate anche in periodi predigitali non deriva esclusivamente dall'evoluzione tecnologica, ma talvolta da approcci di ricerca abbandonati oppure dall'emergere di nuovi filoni a cui i nostri dati non sembrano adeguarsi senza consistenti interventi di ristrutturazione delle forme di rappresentazione dell'informazione (Börjesson *et al.* 2022).

Anche a causa delle continue innovazioni strumentali e tecnologiche introdotte nel mercato, gli archivi possono restare inutilizzati o sottoutilizzati senza interventi che potremmo definire di archeologia di emergenza. A questa specifica caratteristica dei vecchi archivi era dedicato un contributo pubblicato oltre 20 anni fa (D'Andrea, Niccolucci 2001), nel quale si metteva in luce quanto la tecnologia, in larga misura proprietaria, non fornisse gli strumenti per una gestione efficiente e trasparente dei dati. Nell'articolo si suggeriva l'impiego di sistemi di marcatura in linguaggio XML per descrivere il documento aderendo ai principi del W3C (2000) basati sul ricorso a documenti “...*human-legible and reasonably clear*...”. Le indicazioni contenute nel contributo solo in parte sono state confermate dall'introduzione delle tecnologie semantiche, ma la proposta di utilizzare formati aperti e trasparenti sembra oggi costituire un vantaggio di non poco conto rispetto ai formati chiusi. La trasformazione introdotta dall'Unione Europea nei primi anni 2000 nei documenti relativi al free software del V Programma Quadro sembra costituire uno dei più significativi passaggi che hanno condotto nel tempo alla definizione degli strumenti e degli obiettivi dell'Archeologia Digitale. Ad una fase iniziale, coincidente con l'assenza di supporti, di documentazione, strumenti formativi e assistenza per l'uso dei prodotti

open source, è subentrato un periodo di sviluppo e largo impiego di sistemi gratuiti, ma soprattutto affidabili ed efficienti.

I *legacy data* sono archivi non compatibili con le pratiche e gli standard digitali, ma non è opinione condivisa se essi includano anche i dati non digitali, come, per esempio, quelli cartacei da fonti storiche o da precedenti interventi. Se la definizione tecnica si riferisce ai dati generati da vecchi sistemi, è però chiaro che, per una ricerca archeologica che voglia impiegare al meglio tutti i dati disponibili in uno specifico contesto, il termine debba comprendere tutti i record, anche quelli predigitali. Questi dati cartacei da digitalizzare e quelli già digitali, ma codificati con vecchi sistemi, devono essere preparati e manipolati per un trattamento in ambienti informativi più moderni; inoltre, devono essere accuratamente analizzati gli strumenti adoperati, le procedure impiegate per la digitalizzazione e, soprattutto, le modalità di registrazione e formalizzazione del dato. In uno scenario caratterizzato in maniera sempre maggiore dalla navigazione tra i dati e dall'esplorazione di nuove forme di aggregazione dell'informazione, il tema dei *legacy data* diviene, pertanto, centrale. Analogamente, nell'ecosistema digitale un ruolo rilevante assumono quelle modalità di formalizzazione del dato in grado di trattare e conservare la parzialità e l'incertezza dell'informazione contro ogni forma di rigidità del software (Markos, Kalacy, Sarris 2022).

I *legacy data* includono ogni forma di descrizione elaborata ricorrendo ad una narrazione non strutturata (taccuini, diari di scavo), i primi database oramai non più supportati e mantenuti, i disegni con tocco realistico, le fotografie a stampa, le diapositive e, infine, mappe cartacee o digitali eseguite con le prime versioni CAD. I dati ereditati comprendono piante di distribuzione di oggetti o siti, vecchie raccolte di superficie, schede registrate su carta o con pioneristici strumenti digitali non più esistenti, liste di oggetti, report e, soprattutto, materiale edito, pubblicato talvolta in atti di convegni o riviste con limitata diffusione. Purtroppo, non sempre questi archivi, che rappresentano l'unica testimonianza di un paesaggio profondamente mutato, sono preservati e resi disponibili per successive indagini, mettendo così a rischio la conservazione di informazioni preziose. Questi dati non sono indispensabili solo per la ricerca archeologica. Essi sono una fonte unica e insostituibile per una corretta gestione del patrimonio archeologico all'interno di una

pianificazione paesaggistica sostenibile, per il monitoraggio delle risorse monumentali presenti in un determinato contesto territoriale e, infine, per evidenziare sopravvivenze di culture e modelli insediativi del passato nel tessuto abitativo moderno.

Poiché l'archeologia è una scienza che accumula informazioni e che esamina tutte le fonti a disposizione, gli studiosi devono obbligatoriamente riprendere la documentazione preesistente prima di avviare una nuova ricerca. La straordinaria varietà di fonti e supporti, spesso contenuti in archivi dimenticati, nonché l'ampia ricchezza delle testimonianze presenti in vecchi studi, obbliga i ricercatori a mettere a punto procedure automatiche o semi-automatiche che consentano il recupero e la lettura dei *legacy data*. Di frequente, però, la tecnologia da sola non è sufficiente a garantire il corretto aggiornamento degli archivi e l'integrazione di dataset registrati nel corso di precedenti interventi, in sistemi informativi più performanti. Frequentemente questo processo si basa su una rivalutazione del valore della conoscenza archeologica, cioè di quel sapere implicito, talvolta tecnico, che la vecchia documentazione contiene e che non è sempre immediatamente ricavabile dalla semplice lettura dei dati. Non si tratta di una ovvia reinterpretazione dei dati, che ha origine quando nuove evidenze vengono portate alla luce o sono formulate nuove domande di ricerca, quanto piuttosto di una più profonda comprensione dei metodi che hanno condotto alla registrazione di quella specifica informazione e alla formalizzazione del record indipendentemente dal supporto cartaceo o digitale.

Per riutilizzare questo corpus di informazioni, in gran parte digitali, occorre studiare gli archivi, la loro origine, strutturazione e funzione e, secondo una fortunata espressione diffusa nella recente letteratura sull'archeologia digitale, sbloccarli (Casarotto 2022). Se nella creazione dei record il ricercatore ha impiegato metodi noti e obiettivi altrettanto evidenti, allora un tale percorso di recupero dei dati è praticabile; una terminologia standard, descrizioni normalizzate, sebbene non più utilizzate, contribuiscono alla riuscita del processo di conversione degli archivi. Nuovi assemblaggi e significati sono possibili soltanto se le forme della precedente codifica dei dati sono individuabili; il rischio è che senza una struttura trasparente questi dati possano risultare inutilizzabili.

Poiché l'archeologia è una disciplina che studia i dati che essa stessa produce attraverso l'impiego di specifiche metodologie di indagine e raccolta delle informazioni, particolare cura è riservata agli archivi che costituiscono un patrimonio informativo da tutelare e conservare. Il riutilizzo dei *legacy data* non modifica questa attitudine, quanto, piuttosto, la circoscrive alla gestione del dato finale il cui impiego non è limitato all'utente umano, ma anche, e talvolta solo, ad una macchina. Ogni minimo cambiamento che interviene nella tecnologia o nel trattamento automatico dell'informazione può determinare l'obsolescenza dei nostri archivi che diventano inconsapevolmente vecchi.

Con l'espressione dati ereditati si intendono tutti quei dati che possono essere riutilizzati mediante particolari procedure informatiche o adattamenti funzionali che non sono ristretti ai semplici aggiornamenti del software. L'impiego di particolari tecniche nella conversione consente in teoria di integrare qualsiasi vecchio e nuovo archivio conservato e accessibile, ma la solidità del risultato finale non dipende dagli strumenti che la tecnologia rende disponibile, quanto piuttosto dalle esigenze poste dalle nuove traiettorie di ricerca. I *legacy data*, che in questa prospettiva diventano una fonte rinnovabile di conoscenza, si possono suddividere in dati codificati in un qualche linguaggio non più accessibile senza radicali interventi, e informazioni relative a precedenti indagini che un qualsiasi ricercatore intenda riutilizzare ricorrendo a nuove tecnologie e metodologie di analisi.

Tra i dati ereditati una particolare categoria di informazione è rappresentata dai diari di scavo al cui interno si trova la metodologia, l'evoluzione delle indagini, le ipotesi, le osservazioni e, talvolta, i dubbi registrati dall'archeologo durante le attività sul terreno. Questi giornali sono uno straordinario racconto che ci illustra la strategia che il ricercatore ha seguito nel corso dello scavo, le fasi di identificazione, rimozione e interpretazione di una unità stratigrafica, il riconoscimento di una fase cronologica. Spesso i diari descrivono in modo dettagliato anche i ragionamenti formulati dall'archeologo, consentendo di ricostruire i processi logici che hanno condotto lo studioso alle conclusioni. A differenza delle schede stratigrafiche, introdotte per generalizzare e normalizzare le modalità di registrazione dell'informazione, i diari di scavo contengono osservazioni preliminari, ipotesi, ma anche i dubbi e gli eventuali

errori, la ripartizione interna dello scavo e le motivazioni per ogni estensione, allargamento e/o approfondimento. La documentazione grafica e fotografica, generalmente allegata, arricchisce il contenuto del testo fornendo una dimensione visuale dello scavo e dei rinvenimenti.

Un qualsiasi intervento di digitalizzazione dei diari di scavo non può ignorare una attenta rilettura del metodo e del ragionamento che l'archeologo nelle varie fasi di indagine ha seguito. In questo senso l'operazione di conversione digitale dei quaderni di scavo non è un processo lineare, privo di salti, brusche frenate e ripensamenti. Se la fase di trascrizione appare una prevedibile copia digitale realizzata con un semplice software di scrittura, la trasposizione di questa narrativa all'interno di schede o in un record di un database risulta un'operazione delicata che rischia di semplificare il contenuto informativo di un dato in predeterminate voci o campi di una tabella. Il pericolo di generare dati errati nasce soprattutto dall'esigenza di dover dividere il filo della narrazione, di un ragionamento complesso e articolato, nella struttura rigida e formale di un database. Il processo che porta alla creazione di blocchi di informazioni chiuse, autonome e decontestualizzate è definito riduzionismo; una procedura corretta dovrebbe comportare una più ampia formalizzazione della conoscenza e non soltanto del singolo dato salvaguardando tutte le sfumature e i dubbi che sono spesso presenti nel record archeologico.

In un recente contributo dedicato all'importazione di dati CAD in una applicazione GIS (D'Andrea 2022a), è stata sottolineata la difficoltà di rendere automatico il processo di migrazione dei *legacy data* spaziali sebbene la tecnologia dei sistemi geografici consenta di visualizzare file CAD anche in formato proprietario. Il riuso di disegni (mappe, piante) si basa sulla presenza di oggetti discreti, cioè finiti, associati a descrizioni alfanumeriche ed interrogabili. Le primitive geometriche del rilievo CAD sono insufficienti poiché la formalizzazione dei programmi di grafica vettoriale non prevede obbligatoriamente, come nei GIS, la creazione di oggetti; gli elementi grafici CAD non hanno alcuna connessione concettuale con l'evidenza archeologica rilevata e, pertanto, sono esclusivamente una testimonianza visuale dello scavo, priva della necessaria descrizione. Soltanto un adeguato intervento umano, in grado di ricostruire le relazioni tra tutte le informazioni contenute nella documenta-

zione (testi, disegni, foto, etc.), può consentire il riuso funzionale di questa particolare tipologia di dati ereditati; sebbene gli archivi CAD contengano dati digitali accessibili e disponibili, il loro contenuto può risultare del tutto inutile o fuorviante all'interno di un GIS senza precise e dettagliate informazioni che illustrino l'articolazione dello scavo e l'interpretazione dei vari elementi grafici disegnati. L'assenza di schede, descrizioni o diari di scavo rende la fase di identificazione degli oggetti molto soggettiva e impossibile da automatizzare; una semplificazione nella creazione di oggetti può avere conseguenze negative sulla corretta organizzazione e gestione delle informazioni archeologiche (al-Rawash-deh Balqies Sadoun, Al Fukara 2012).

Mentre un disegno CAD adopera tratteggi, stili di linea, retini per indicare l'allineamento di un muro, la chiusura di un ambiente, lo sviluppo di uno strato e la ricostruzione di un settore dello scavo, il GIS richiede una più rigorosa e formale definizione delle varie geometrie, che si può realizzare solo attraverso la costruzione di elementi grafici associabili ad un unico contenuto alfanumerico. Senza una conversione concettuale si rischia di utilizzare il GIS come un più moderno motore CAD in grado di combinare differenti fonti cartografiche che, però, non possono essere interrogate utilizzando funzioni topologiche.

Un GIS visualizza un file CAD riconoscendo le caratteristiche geometriche degli elementi grafici senza alcuna interazione tra i dati disegnati e il mondo reale. Solo l'intervento umano è in grado di perfezionare il processo di migrazione attraverso la trasformazione dei punti, delle linee, dei poligoni e, infine, delle annotazioni in oggetti reali tratti dal mondo fisico.

Analoghe problematiche vengono sollevate nel caso di riuso di dati spaziali acquisiti in una epoca pre-GPS, cioè almeno fino al 2000, anno in cui il segnale venne reso accessibile liberamente e senza errori. Di frequente gli archivi includono dati da ricognizioni superficiali che non contengono alcuna informazione relativa alle metodologie di campionamento prescelte, ai sistemi di ispezione del terreno, al numero dei ricercatori coinvolti e, soprattutto, alle modalità di identificazione spaziale delle evidenze archeologiche registrate (Ullah 2015). Inoltre, i ricercatori possono aver adoperato carte locali con sistemi geografici che producono errori se riproiettati su coordinate codificate, ad esempio, in

WGS84, oppure possono aver arrotondato le misure memorizzate. L'assenza di metadati e *paradata* (infra 111) rende complesso il processo di verifica e riutilizzo dei dati geo-spaziali (Chapman 2001; Witcher 2008). L'uso dei GIS può aiutare a comprendere la sistematicità di eventuali errori e consentire la correzione, almeno in parte, delle imprecisioni registrate (Ullah 2015).

La digitalizzazione di *legacy data* spaziali è indissolubilmente legata a questioni di metodo che non sono risolvibili con l'implementazione di procedure automatiche e/o di software, per quanto sofisticate. Poiché il passaggio da elementi grafici, cartacei o digitali, a oggetti semantici richiede una attività umana che può essere paragonata alla fase di scoperta e registrazione del record, questo processo è ancora saldamente nelle mani dell'archeologo che deve districarsi tra sovrapposizioni stratigrafiche, non sempre ben definite, ed entità geometriche e/o semantiche che riproducono digitalmente le evidenze portate alla luce.

Il caso del riuso dei dati CAD in un GIS conferma l'importanza che hanno i dati ereditati nella ricerca archeologica, ma nello stesso tempo sottolinea il ruolo che l'intervento umano ha nella riscoperta di archivi obsoleti e/o dimenticati; analogamente l'analisi del processo di migrazione fa emergere la necessità di sviluppare una strategia che non consideri solo la parte informatica della conversione, ma valuti allo stesso modo importante quella concettuale che passa attraverso una riflessione sul metodo di acquisizione e restituzione dell'oggetto digitale. L'analisi rivela, in sintesi, l'esigenza di associare agli archivi digitali buoni metadati e *paradata* per garantire il futuro riutilizzo dei dati. Per questo motivo il ricercatore non deve limitare le sue competenze alla struttura informatica dei dati, ma deve possedere una approfondita conoscenza delle metodologie e delle tecniche che i ricercatori hanno impiegato nel tempo; solo in questo modo i dati ereditati diventano informazioni contemporanee e utilizzabili.

L'importanza progressivamente assunta dai *legacy data* deve spingere gli studiosi a pubblicare in forma accessibile i dati grezzi e i report preliminari inediti, in attesa della stampa finale delle ricerche. Nello stesso tempo, però, devono essere adottate adeguate procedure e misure affinché sia valutabile l'affidabilità di questa massa di conoscenze che costituisce una fonte di riferimento in molti casi imprescindibile per



l'avvio di nuove indagini. Spesso i report sono sintetici, non contengono informazioni precise sulla quantificazione e sulla natura dei materiali rinvenuti, sulle strategie di ricerca e/o campionamento impiegate e sono corredati da poche immagini e disegni preliminari. L'assenza di dati sulla classificazione degli oggetti recuperati, di riferimenti spaziali per le aree indagate o di tabelle riassuntive, impoverisce il contenuto informativo della letteratura grigia e fa sorgere il sospetto che, trattandosi in larga misura di relazioni legate ad interventi archeologici di emergenza, qualsiasi ragionamento o interpretazione dello scavo possa rivelarsi parziale ad una più approfondita analisi. Richards (2017) ha rilevato come molti archivi non soddisfino standard minimi di registrazione includendo soltanto testo e qualche foto, mentre di recente Huggett (2022), per valutare l'attendibilità delle relazioni, si è interrogato sulla natura dei report e sul processo di trasformazione dei dati originariamente raccolti in rapporti finali. Sulla base delle riflessioni dei due autori sarebbe fin troppo semplice concludere sulla inaffidabilità e scarsa qualità della letteratura grigia. Al contrario, la considerazione che emerge riguarda l'esigenza di comprendere appieno la natura dei dati presenti nei report. Un approccio consapevole deve prevedere una fase di valutazione attenta del testo e della documentazione associata, per apprezzarne più chiaramente parzialità, e eventuali incoerenze, e per riutilizzare le fonti in modo equilibrato.

Nell'esaminare e aggiornare vecchi archivi, i metodi propri dell'Archeologia Digitale devono necessariamente intrecciarsi con quelli delle tradizionali tecniche di acquisizione e registrazione dell'informazione archeologica; il prevalere di una delle due competenze può determinare, come abbiamo segnalato, rischi significativi tanto più gravi quanto più le informazioni circolano liberamente in un ecosistema digitale. Lo studio dei *legacy data* passa attraverso una *slow archaeology* che contribuisce a definire la qualità degli archivi e a favorire lo sviluppo di una ricerca realmente *data-intensive* (Sobotkova 2018; Börjesson *et al.* 2022; Huggett 2022). In sostanza un qualsiasi dataset erediterà i limiti della sua fonte originaria indipendentemente dal suo formato (Sobotkova 2018).

Il passaggio ad approcci automatici o semi-automatici per l'estrazione di conoscenza dalla letteratura grigia, che consiste soprattutto in report codificati in PDF, espone il ricercatore a forme di pericolosa sopravvalua-

tazione della fonte informativa, aspetto tanto più pericoloso quanto più si perde traccia del riuso dell'informazione nel tempo (Morgan, Eve 2012).

### 1.5. *Open Knowledge*

L'attenzione ad una corretta formalizzazione dei dati, inclusi quelli *legacy*, deve concretizzarsi in una particolare riflessione sulla natura del record archeologico e soprattutto sulla sua condivisione in rete. Un tale atteggiamento, orientato a garantire una maggiore collaborazione, sembra dare una soluzione all'interrogativo sollevato da J.C. Gardin (1999) sullo squilibrio tra il volume crescente di lavori editi in qualsiasi campo e la quantità quasi immutata di tempo dedicabile alla loro lettura. Poiché il ricercatore, leggendo una pubblicazione, non ha accesso ai dati scientifici di uno specifico contesto di ricerca, alcuni archeologi statunitensi (Kintigh *et al.* 2014) hanno proposto di rendere accessibili gli archivi; i dati primari o grezzi, accuratamente descritti, i report inediti e la documentazione associata (inclusa quella grafica e fotografica) costituiscono un supporto essenziale per qualsiasi analisi comparativa.

Lesigenza di pubblicare online i dati si ritrova anche in contributo di G.P. Brogiolo (2012) che, per contrastare i rischi di una archeologia senza dati (Brogiolo 1997), propone la creazione di un WebGIS, di un sistema WIKI e di archivi online in modo da rendere consultabili le schede degli scavi di emergenza. Poiché il 90% di questa tipologia di interventi resta inedito, Brogiolo ha proposto l'adozione di un approccio aperto e condiviso che avrebbe prodotto un decisivo cambiamento dei ruoli e delle responsabilità istituzionali dello Stato e dei suoi organi periferici. L'autore suggerisce l'elaborazione di un modello policentrico di organizzazione e gestione della conoscenza basato sulla messa in rete degli archivi delle Soprintendenze e sulla pubblicazione online dei dati prodotti dalla ricerca accademica e dall'archeologia professionale. Nella idea di Brogiolo l'informatica non è quindi solo uno strumento che accompagna o guida la ricerca, ma addirittura un approccio indispensabile alla sopravvivenza stessa della archeologia.

Della stessa opinione è A. Carandini (2008, 19) per il quale

dovrebbero esistere «sistemi informativi archeologici» unitari, statali, regionali e universitari, capaci di condividere un minimo di procedure essenziali.

Per alimentare questi sistemi Carandini prosegue auspicando la nascita di una soprintendenza del futuro

...costituita da pochi funzionari, ben formati archeologicamente e anche tecnologicamente ... i quali dovrebbero ronzare come api operose - esperte in AutoCad - attorno a quello straordinario favo digitale.

Per Carandini

Non servono schede di monumenti o di oggetti singoli ma appunto sistemi informativi territoriali in cui i dati siano geo-riferiti e vettorializzati ... che diano informazioni scientifiche ed essenziali.

Brogiolo e Carandini concordano sulla semplice constatazione che, per produrre conoscenza archeologica, occorre in modo prioritario mettere online gli archivi di dati e schede, sviluppando una nuova riflessione metodologica sulle forme di creazione, organizzazione e gestione del record archeologico.

Questo tema e le varie soluzioni progettate e messe a punto costituiscono uno dei punti più controversi della fase di digitalizzazione e soprattutto dell'*Open Knowledge*, un modo nuovo di fare ricerca finalizzato ad ottimizzare la produzione e a migliorare la circolazione dei dati. Per favorire uno scenario caratterizzato dalla collaborazione sono stati messi a punto negli ultimi anni specifici linguaggi per la formalizzazione, la descrizione normalizzata e la visualizzazione dei dati in rete; grazie alla diffusione del Semantic Web sono state progettate piattaforme che costituiscono l'infrastruttura concettuale per un accesso più evoluto e meno ambiguo alle informazioni disponibili.

Purtroppo, l'evoluzione del Web non sta avvenendo con una guida in grado di coordinare gli sforzi, ma in forma del tutto spontanea e con procedure non sempre facilmente comprensibili e quindi utilizzabili senza una conoscenza approfondita del codice informatico e, soprattutto, dell'architettura dei sistemi. Il ricercatore fa spesso fatica a districarsi tra soluzioni software e schemi messi a punto per strutturare archivi e file. Nonostante ciò, alcuni principi sono stati da tempo fissati e rappresentano un punto fermo per quanti vogliano standardizzare le loro risorse digitali e renderle accessibili nello spazio dell'OS. Prima di tutto,

però, è necessario che si consolidi nei ricercatori la consapevolezza che, in una rete pressoché infinita di connessioni, tutti sono produttori e consumatori di dati e che per rendere accessibili i dati si deve parlare una lingua comune, accettando anche qualche piccola modifica alle abituali pratiche di registrazione delle informazioni.

Molte delle innovazioni che la ricerca e la scienza hanno conosciuto negli ultimi anni provengono dalla introduzione di software in grado di realizzare calcoli prima d'ora inimmaginabili. Alcuni settori, come la computer graphics, hanno radicalmente trasformato l'approccio alla progettazione e riproduzione in digitale di parti dei processi costruttivi e/o descrittivi. Profondi cambiamenti innescati dalle tecnologie abilitanti sono visibili in quasi tutti i domini scientifici, inclusi anche i settori della ricerca umanistica. Tuttavia, i mutamenti a cui abbiamo assistito di recente non riguardano la messa a punto di soluzioni informatiche o di sistemi in grado di rivoluzionare i tradizionali approcci alla ricerca, quanto piuttosto l'emergere di una nuova filosofia sottesa all'idea stessa di scienza e comunità scientifica.

In questa metamorfosi un ruolo determinante è stato segnato dall'evoluzione delle tecnologie della comunicazione che, con la diffusione di Internet e delle connessioni ed interconnessioni globali pressoché infinite, hanno contribuito a definire gli orizzonti di un modo nuovo di fare scienza. I progressi di Internet, da semplice strumento di presentazione dei dati a sistema in grado di promuovere nuove forme di aggregazione, discussione e attività interdisciplinare, hanno scandito le fasi di questa trasformazione.

La ricerca archeologica, come altri domini disciplinari, ha tratto vantaggio da una trasformazione che ha modificato in profondità i processi di costruzione e condivisione della conoscenza. Non tutti però, ancora oggi, sono consci dei cambiamenti fondamentali che la tecnologia ha sostenuto anche in campo archeologico, principalmente nel settore della formalizzazione della conoscenza e dello stesso ragionamento archeologico.

Da questo versante i recenti sviluppi della rete sembrano dare ragione a Gardin che già negli anni '50, in una fase certamente pionieristica dell'informatica applicata alla ricerca archeologica, aveva compreso l'importanza del processo di rappresentazione dei dati attraverso co-

strutti che descrivono caratteristiche e soprattutto relazioni tra gli oggetti (Gardin 1959; Orlandi 2004; Moscati 2013).

Gardin aveva posto correttamente l'accento su una forma di descrizione normalizzata legata a standard in grado di superare registrazioni troppo personalizzate. Per l'archeologo francese la formalizzazione del ragionamento costituiva la principale sfida dell'archeologia informatica, concentrata sulla implementazione di procedure automatiche idonee a validare la logica del discorso della interpretazione archeologica. Con l'affermarsi delle tecniche di AI le intuizioni originali di Gardin potrebbero essere facilmente testate; agenti automatici potrebbero essere utilizzati per validare la correttezza formale di una qualsiasi affermazione scientifica suddivisa in costrutti semantici e connessa da relazioni.

Gli strumenti messi a disposizione dalla rete consentono, oggi, di approcciare in modo radicalmente nuovo al tema della codifica e soprattutto a quello della rappresentazione della conoscenza, un paradigma innovativo alla base delle più recenti evoluzioni della rete. Per Moscati (2019) i metodi digitali costituiscono, oramai, un modello irrinunciabile per la documentazione, gestione e analisi di dati; gli archeologi chiedono a tali metodi di contribuire ad una maggiore integrazione delle risorse digitali e, soprattutto, di supportare l'implementazione di sistemi o infrastrutture per condividere dati, metodologie e pratiche.

L'*Open Definition* della *Open Knowledge Foundation*, precisa che il termine aperto significa che chiunque può liberamente accedere, usare, modificare e condividere i dati per qualsiasi scopo, preservando, esclusivamente, i requisiti che conservino traccia della provenienza. Non si tratta, quindi, di limitarsi ad un conferimento con licenze pubbliche dei propri archivi, ma di rendere i dati realmente aperti, cioè in primo luogo comprensibili. In questo modo si realizza una accessibilità concreta e controllata dei dati e dei servizi, in contrapposizione ad un ambiguo concetto generico di dato aperto; una accessibilità trasparente e comprensibile consente il coinvolgimento di un'ampia serie di istituzioni pubbliche e private e, soprattutto, un'autentica partnership paritaria.

La documentazione archeologica digitale è in pericolo non soltanto per questioni legate ai *legacy data*, ma soprattutto per l'incapacità dei ricercatori di allineare i loro archivi alle esigenze dell'OS; senza interventi rapidi e, soprattutto, determinati l'archeologia rischia di perdere la

maggior parte del suo patrimonio di dati di ricerca e della sua capacità digitale (Richards *et al.* 2021).

L'introduzione massiccia delle tecnologie digitali nel processo di registrazione dello scavo in tutte le sue fasi ha trasformato gli archivi da contenitori di oggetti cartacei e digitali a depositi di documenti esclusivamente digitali e per i quali non esistono copie analogiche. Mentre i ricercatori hanno mostrato creatività e innovazione nell'utilizzo di sofisticati strumenti digitali, sperimentando, talvolta in forme entusiastiche, nuovi approcci per la documentazione archeologica, analoga attenzione non è stata riservata ad una politica di conservazione dei dati come avviene in altre scienze. Il pericolo di un tale atteggiamento è che i risultati di decenni di indagini potrebbero andare perduti a causa della mancanza di depositi specialistici (Wright, Richards 2018) e che senza interventi adeguati la vera sfida dei prossimi anni sarà diretta al recupero ed alla conversione degli archivi digitali pregressi.

### 1.6. Riflessioni di metodo

Le considerazioni contenute nel presente capitolo includono l'analisi dei cambiamenti osservati nel modo di fare di ricerca archeologica e le trasformazioni complessive di una società sempre più dipendente dalle tecnologie informatiche. Nello stesso tempo, mentre si affaccia un nuovo paradigma scientifico basato sullo sfruttamento intensivo dei dati e su una più intensa collaborazione, una maggiore condivisione e standardizzazione delle risorse è richiesta ai ricercatori.

Il confronto tra evoluzione degli strumenti e cambiamenti metodologici, fa emergere l'interrogativo su cosa resti del bagaglio tecnico tradizionale dell'archeologo di fronte ad uno sviluppo tecnologico, quasi illimitato, che estende le facoltà del ricercatore con strumenti innovativi (computer più potenti, stazioni totali robotizzate, sistemi GNSS, laser scanner, droni, etc.) e nuovi processi di conoscenza basati su architetture open e cooperative. Se gli obiettivi della ricerca archeologica restano immutati e orientati alla ricostruzione del passato, i metodi di indagine mutano, spesso parallelamente ad una più complessiva trasformazione della società.

Dopo oltre settanta anni di applicazioni informatiche in archeologia, si impone una riflessione sui cambiamenti intervenuti nella metodolo-

gia, soprattutto perché è necessario comprendere in che modo gli strumenti digitali abbiano modificato il perché e l'oggetto della disciplina, a partire dal come identificare, produrre ed elaborare i dati. Inoltre, è doveroso interrogarsi, in un volume dedicato ai dati digitali, quanto del cambiamento che, oggi, consideriamo nella pratica archeologica, sia il risultato di un atteggiamento consapevole oppure di un semplice ammodernamento acritico dei metodi e delle tecniche abituali.

Il rapporto tra teoria, metodo e metodologia è stato analizzato da numerosi studiosi. Renfrew e Bahn (2006) hanno, in particolare, sottolineato che la ricerca non può essere derubricata a mera acquisizione di dati, poiché la conoscenza dipende dallo sviluppo di nuove tecniche di indagine, dalla continua evoluzione della teoria e, quindi, dalla natura delle nuove domande storiche. Se le idee e le teorie sono in continua evoluzione, anche i metodi sono costantemente in trasformazione e non soltanto per i progressi della ricerca scientifica, ma anche grazie alle spinte culturali provenienti dalle trasformazioni tecnologiche e, più in generale, dalla società.

Nonostante i cambiamenti introdotti dalla digitalizzazione, l'Archeologia Digitale non è unanimemente riconosciuta come disciplina autonoma separata dalla metodologia archeologica tradizionale, almeno in ambito italiano. Questo limite, che è principalmente culturale e che relega all'ambito tecnico l'uso dell'informatica, crea una forte cesura tra le metodologie tradizionali in ambito umanistico e quelle computazionali. Come abbiamo in precedenza segnalato, il rischio è che un tale atteggiamento finisca con il favorire il subappalto della tecnologia in campo archeologico ad informatici, architetti e ingegneri, mentre sarebbe più opportuno formare specialisti, anche giovani laureati in archeologia, con competenze realmente interdisciplinari. Un intervento formativo in questa direzione produrrebbe anche il vantaggio di preparare una classe di professionisti in grado di gestire in forme innovative il nostro patrimonio culturale ed archeologico favorendo, in tal senso, l'amministrazione a regime dei tanti progetti di digitalizzazione che vengono periodicamente finanziati a favore di musei, parchi archeologici e istituzioni culturali (Guermandi 1996; Francovich 1999).

L'inserimento dell'Archeologia Digitale all'interno delle discipline metodologiche sembra essere il risultato di due distinti fenomeni solo in

parte contrapposti. Da un lato si registra l'uso ordinario degli strumenti digitali nella ricerca sul campo e in laboratorio; come le cazzuole o la macchina fotografica, il loro impiego non modifica la natura e la qualità dei rinvenimenti. Dall'altra parte, invece, il ricorso al digitale è riservato ad esperti provenienti da altri domini che mostrano limitate competenze archeologiche, ma sembrano maggiormente orientati a comprendere il ruolo che le tecnologie informatiche possono avere nel processo di scoperta.

Entrambi gli atteggiamenti devono essere esaminati per valutare al meglio la ragione stessa di una disciplina, come l'Archeologia Digitale, il cui status è ancora in via di definizione. Pertanto, prima di riflettere sui vari approcci che gli archeologi mostrano nelle varie indagini che usano strumenti digitali, si deve brevemente richiamare il significato di metodo e metodologia, evidenziando in che modo due discipline fortemente strutturate, come l'archeologia e l'informatica, possano integrarsi in una sintesi interdisciplinare particolarmente ricca di sfumature (D'Andrea 2022b).

Si definisce il metodo come una procedura di indagine della realtà finalizzata a dimostrare in modo logico gli enunciati di base di una scienza. In campo storico-archeologico il metodo si identifica in un processo di reperimento, analisi e sintesi delle fonti materiali e testuali. Il rispetto e la meticolosità nell'adozione del metodo conferiscono l'attendibilità scientifica di un insieme di dati.

La metodologia, cioè il discorso sul metodo, esamina le trasformazioni dei metodi su cui una determinata scienza o disciplina fonda la sua attività di osservazione e interpretazione dei fenomeni. La metodologia analizza anche il complesso dei fondamenti teorici sui quali un metodo è costruito con l'obiettivo di interrogarsi sul valore epistemologico dei risultati prodotti dalla ricerca, ovvero di analizzare, correggere e migliorare i procedimenti di indagine e scoperta. Nel caso della identificazione, raccolta e registrazione dei dati, il metodologo esplora le fasi del processo di acquisizione e interpretazione delle fonti materiali, alla ricerca di una corrispondenza simmetrica tra la regolare applicazione di un procedimento, con i suoi strumenti di selezione e classificazione, e i risultati ottenuti.

Quando la ricerca metodologica si pone ai confini di due distinte discipline, lo studio del metodo, o dei metodi, diventa più complesso



perché le tecniche e gli strumenti impiegati si intrecciano per produrre nuovi dati che appartengono soltanto in parte ai domini coinvolti. Per l'ambito della classificazione archeologica, particolarmente critico è, ad esempio, il procedimento della rappresentazione e formalizzazione dei dati, un campo che racchiude un articolato patrimonio di tecniche, conoscenze, linguaggi e terminologie che non può essere realizzato con la semplice giustapposizione dei due distinti ambiti del sapere archeologico e informatico.

Codifica binaria e descrizione archeologica devono necessariamente integrarsi in una sintesi che combina le esigenze dell'elaborazione informatica con la ricchezza semantica del dato archeologico che non deve essere semplificata per migliorare la performance computazionale. Il trattamento informatico, sebbene basato su una rappresentazione strutturata del record, deve includere la soggettività del dato informativo, la sua incertezza, il suo valore parziale ed eventuali ambiguità.

Un dato diviene una informazione scientifica se la sua identificazione ed il suo processamento avvengono attraverso l'impiego di un metodo riconosciuto, non soltanto perché si è adoperato un rigido formalismo e un linguaggio pseudo-scientifico. La digitalizzazione dell'informazione archeologica non si deve concretizzare nella riduzione del contenuto informativo di un dato eliminando tutti quei valori e parametri che creano inconsistenza in un archivio digitale. Al contrario, un trattamento informatico, che muova da una valida e ben testata metodologia archeologica, deve consentire la corretta gestione di un qualsiasi dato, anche quello impreciso e ambiguo senza doverlo necessariamente trasformare in una verità codificata.

Il metodo si pone, dunque, come un insieme di tecniche e procedimenti - condiviso da una data comunità scientifica - per acquisire dati e trasformarli in informazioni. Ad esempio, in archeologia, il metodo stratigrafico è impiegato per identificare la relazione di successione e anteriorità tra strati sovrapposti; l'uso della stratigrafia consente di datare, in modo relativo e/o assoluto, la ceramica o altri oggetti rinvenuti nei livelli che l'archeologo identifica e registra nel corso dello scavo. La sequenza che descrive l'ordine cronologico di successione degli strati consente di elaborare una seriazione delle classi ceramiche o di altri oggetti rinvenuti, datando, in questo modo, i livelli stratigrafici.

Mentre i principi della sovrapposizione stratigrafica elaborati da Harris (1983) forniscono ai ricercatori un metodo di indagine, la metodologia archeologica trasforma questi principi in procedure o protocolli o metodi che assegnano un valore di affidabilità scientifica alle informazioni identificate e registrate. La maggioranza degli archeologi concorda oggi sull'efficacia del metodo stratigrafico, sebbene qualche osservazione critica sia stata di recente sollevata sul valore del matrix (Zanini 2021).

Soltanto attraverso una indagine che evidenzii i livelli di costruzione, uso e abbandono, è possibile indagare le fasi di frequentazione di un sito, la storia di un monumento, lo sfruttamento di un paesaggio. Nonostante l'unanime riconoscimento dell'importanza del metodo stratigrafico alcune proprietà dell'oggetto appaiono indipendenti dal metodo di scoperta. In molti casi le informazioni sulla ceramica non derivano esclusivamente dallo scavo: la funzione, la composizione, la provenienza non sono suggerite dalla sovrapposizione degli strati. Se il contesto stratigrafico fornisce una ipotesi per la datazione degli oggetti, in altri è il reperto archeologico (le monete, particolari tipi di ceramica di importazione, etc.) a consentire di fornire la corretta cronologia di uno strato.

Questo esempio mostra la stretta relazione tra metodologia e metodo che potremmo anche definire tra lo studio della funzione di una procedura e la sua applicazione per produrre dati riconosciuti da una specifica comunità scientifica. Un ambito di ricerca metodologico destinato alla analisi dei metodi informatici applicati alla ricerca archeologica, non può, allora, che comprendere competenze miste, archeologiche ed informatiche, funzionali soprattutto alla individuazione di un più alto livello metodologico, fortemente innovativo, in grado di sintetizzare i vari approcci e generare nuovi dati.

La crescita delle soluzioni software e hardware, con la diffusione di strumenti trasportabili e di facile uso, non è stata, purtroppo, accompagnata dall'emergere di una metodologia mista, consapevole delle due facce del problema, quella algoritmica e quella archeologica. La rincorsa verso l'applicazione, piuttosto che in direzione della riflessione, ha generato un effetto distorsivo. Alcune norme, ad esempio, sono state messe a punto (carte di Londra, di Ename e di Siviglia) nel tentativo di fornire delle linee guida per la creazione di modelli 3D e per assicurare la long-

term preservation degli archivi, suggerendo regole per la condivisione ed il riuso dei dati (Moscati 2017). Anche il fenomeno dell'OS ha contribuito ad aumentare la consapevolezza degli addetti ai lavori nella manutenzione degli archivi proponendo codici di comportamento finalizzati a ridurre il rischio della dispersione dei dati.

Pressioni esterne, oltre che una maggiore attenzione disciplinare, hanno incoraggiato una riflessione finalizzata ad una più ampia comprensione e condivisione della relazione tra fonti archeologiche, applicazioni informatiche e risultati finali. L'esistenza di una disciplina autonoma che indaga l'impiego di tecnologie digitali in archeologia, non appare giustificata soltanto dalla presenza di uno specifico campo di interesse, a metà strada tra la ricostruzione storica e l'applicazione di tecniche di trattamento automatico di dati, quanto dall'emergere di proposte, problematiche e considerazioni che mettono insieme studiosi con competenze e saperi trasversali accomunati da un particolare sguardo sul mondo. La sfera di intersezione tra i due domini raggruppa un ambito di studi che si focalizza sulle applicazioni digitali per la descrizione, rappresentazione e soprattutto spiegazione di fenomeni e processi antropici e sulla identificazione delle pratiche più adatte alla formalizzazione degli oggetti digitali (testuali, spaziali, multidimensionali, etc.) anche nell'ottica dei più recenti principi introdotti dal Semantic Web e dall'OS.

Riflessioni metodologiche sui metodi e sui paradigmi scientifici si intrecciano per fornire allo studioso le migliori linee guida per un corretto impiego della tecnologia disponibile. Nello stesso tempo le attività di ricerca per la creazione di best-practices evidenziano il contributo che le applicazioni informatiche forniscono in archeologia alle innovazioni di metodo e alle trasformazioni della metodologia.

In un contributo dei primi anni 2000 (D'Andrea, Niccolucci 2001) era stato già definito, a grandi linee, il profilo di una Archeologia Digitale identificando alcuni particolari settori di interesse:

- l'informatica del ragioniere, ovvero la tecnologia dei database relazionali progettati per gestire la contabilità finanziaria dei conti correnti e utilizzata, invece, per strutturare informazioni archeologiche;

- l'informatica del geografo, orientata a riadattare ai contesti archeologici le funzioni e le analisi spaziali presenti nei GIS;
- l'informatica dell'ingegnere, finalizzata a supportare un esercizio di immaginazione creativa a metà tra la ricostruzione scientifica e la comunicazione rivolta al grande pubblico.

Alcune delle linee di ricerca enunciate in quel contributo, oramai in parte superate considerando i tempi dell'evoluzione tecnologica, sono cresciute e maturate nel corso della realizzazione di numerosi progetti e interventi internazionali multi e interdisciplinari. Altre linee si sono, nel frattempo, affermate e altre ne verranno nei prossimi anni sulla scia dell'evoluzione tecnologica.

Per sfruttare in modo positivo e duraturo le novità introdotte dai cambiamenti tecnologici, è necessario proseguire sul terreno della riflessione sulla metodologia archeologica; troppo spesso le trasformazioni, pur timidamente presenti soprattutto nel campo della valorizzazione e della comunicazione, non sono state sempre accompagnate da una adeguata valutazione del ruolo che gli strumenti hanno avuto nei mutamenti disciplinari preferendo un approccio pratico, efficiente ed efficace. Nel campo del rilievo, ad esempio, si è passati da quella tradizionale capacità di vedere, che caratterizzava il mestiere sul campo del rilevatore/architetto/archeologo e che consentiva di identificare e disegnare informazione archeologiche, alla produzione di nuvole di punti 3D estremamente dettagliate e complesse, ma non in grado di rappresentare oggetti estraibili in forma automatica.

In un universo caratterizzato dalla produzione, dalla condivisione e dall'uso di dati digitalizzati, digital-born, o *legacy* gli archeologi appaiono, purtroppo, troppo spesso concentrati sul proprio dominio, ignorando alcune delle elementari regole di base che l'informatica impone per la formalizzazione, l'organizzazione e il trattamento delle fonti.



## Capitolo 2

### L'INFORMAZIONE TESTUALE

#### 2.1. La nascita dell'Archeologia Digitale

I dati elaborati dagli archeologi per i loro studi sono il risultato materiale, residuale e spesso parziale, di comportamenti antropici o naturali. Ma attraverso quale linguaggio (naturale, formale) i ricercatori descrivono l'oggetto? E in che modo quell'oggetto diviene una informazione utile per la successiva interpretazione? La riflessione sulla formalizzazione o codifica del dato e della conoscenza è un processo avviato alla metà degli anni '50 e finalizzato a comprendere la natura del record archeologico.

La codifica dei dati testuali è stata probabilmente una delle prime tematiche trattate dall'Archeologia Digitale. Commentando criticamente, agli inizi degli anni '70 dello scorso secolo, il panorama dei progetti nei quali era presente un trattamento informatico dei dati, R. Whallon (1972, 40) concludeva che

*This review, brief and incomplete as it is, should nonetheless give some indication of the tremendous expansion of computer use in archaeology since the last review of the subject four years ago in this journal. Impressionistically, the most frequent and important uses of the computer in the day-to-day archaeological research at the University of Michigan are in the area of data manipulation and simple descriptive statistics, sorting, counting, frequency distributions, cross-plots, parameters of distributions, associations and correlations, and so on.*

Nel segnalare l'esplosione dell'uso del computer in archeologia Whallon individua nella descrizione e manipolazione dei dati una delle principali aree di ricerca insieme alla successiva elaborazione statistica dei record.

Agli inizi degli anni, '60 del '900, l'archeologia conosce un decisivo cambiamento nelle sue fondamenta all'interno della corrente nota come New Archaeology sorta con l'obiettivo di trasformare l'archeologia in una

disciplina fondata sul modello ipotetico-deduttivo derivato dalle scienze esatte, sull'utilizzo di tecniche sperimentali indirizzate a garantire la riproducibilità dell'esperimento. All'interno di un più ampio processo di revisione dei metodi, caratterizzato principalmente dal ricorso ad un approccio scientifico più rigoroso e formale e da un metodo di tipo generale piuttosto che particolare, le tecniche informatiche vennero subito considerate come necessarie alla creazione di strutture formali oggettive in grado di eliminare ogni forma di soggettività dai dati primari.

Le prime applicazioni informatiche in archeologia, sviluppate agli inizi degli anni '60 (Lock 2003; D'Andrea 2006; 2015b), si concentrarono in alcuni centri di ricerca ubicati in Francia, Gran Bretagna e negli USA e furono indirizzate alla sperimentazione di procedure per la codifica e la classificazione di complessi archeologici particolarmente ricchi di rinvenimenti. Fin dalle origini, il legame tra archeologia e informatica si è caratterizzato per il tentativo, portato avanti agli inizi da pochi ricercatori, di comprendere fino in fondo le tecniche computazionali, fornendo così un fondamento teorico per il cambiamento dei metodi, delle procedure e delle strategie in archeologia.

L'avvio di una revisione teorica e metodologica così profonda non poteva che accompagnarsi alla creazione di adeguati strumenti di confronto e verifica. Numerosi convegni e incontri professionali furono organizzati tra la fine degli anni '60 e gli inizi del decennio successivo (Whallon 1972), mentre tra il 1965 e l'anno successivo furono fondate due riviste: la *Newletter of Computer Archaeology*, su iniziativa del Dipartimento di Antropologia dell'Università dello Stato dell'Arizona e diretta fino al 1971 da R. Chenhall, e la *Computers and the Humanities* fondata da J. Raben, professore di inglese al Queens College dell'Università di New York.

L'attenzione crescente verso l'uso del computer testimonia la progressiva importanza assunta dai metodi informatici in archeologia al punto che lo sviluppo dell'Archeologia Digitale può essere considerato come un processo in divenire in cui, tranne poche eccezioni, il computer è stato forzatamente piegato per irrobustire una teoria o per avvalorare un nuovo approccio metodologico. Così è stato per la cornice teorica della New Archaeology, interessata alla messa in discussione dell'approccio storico-culturale e, successivamente, della Post-Processual Archaeology, fino ai nostri giorni nei quali l'uso delle metodologie compu-

tazionali si collega ad un più generale diffusione delle tecnologie dell'informazione e della comunicazione con la definizione di una società pienamente digitale.

L'uso del computer in archeologia ha contribuito, certamente, ad agire come elemento di novità nel dibattito degli ultimi 70 anni influenzando anche gli approcci più recenti del contestualismo, della complessità, della riflessività e del cognitivismo (D'Andrea 2006), ma le aree di sviluppo iniziali furono l'*information retrieval* ed il *data processing*, entrambe inestricabilmente collegate all'approccio della teoria processuale finalizzato ad una semplificazione della realtà.

Se i primi tentativi di applicare il computer nella ricerca archeologica furono dedicati all'impiego della tassonomia numerica per lo studio della preistoria britannica, l'analisi quantitativa si estese presto anche agli aspetti spaziali (Clarke 1968, 1972). Il successo degli studi statistici si deve, in larga misura, alla diffusione del paradigma riduzionista (*data minimal*) condiviso tra i sostenitori della New Archaeology. I dati, rappresentati in una forma oggettiva slegata dalla teoria, rivelavano proprietà misurabili e registrabili in modo empirico in quanto ridotte ad unità minime di conoscenza.

Il riduzionismo si applicava ad un discorso quantitativo che lentamente si diffondeva in altri paesi fino a quel tempo scarsamente disponibili a contaminare la cultura umanistica con l'uso dei calcolatori; in Francia la scuola dell'*Analyses des données* aggiunse ai metodi di classificazione multivariata tecniche di tipo multidimensionale, come l'analisi delle componenti principali. Del tutto diverso è, invece, il caso delle banche dati e del relativo processo di formalizzazione che, sempre, in Francia, ma già verso la fine degli anni '50, era stato affrontato con la progettazione di procedure per la descrizione, l'archiviazione e il recupero dei dati.

Uno degli aspetti più controversi e discussi nei primi anni della informatizzazione dei dati archeologici fu, senza dubbio, quello della necessità di disporre di un linguaggio di codifica aperto e flessibile alternativo a schemi chiusi, costruiti sulle esigenze dei singoli casi trattati (Cowgill 1967; Chenhall 1968). Nello stesso tempo il dibattito si estendeva alla valutazione di quali caratteristiche e proprietà dell'oggetto dovessero essere estratte affinché le successive analisi fornissero conclusioni credibili sul piano della ricostruzione storica.



La scelta ponderata di mettere a punto linguaggi adeguati alle variabili selezionate aveva l'obiettivo di rendere oggettivi i dati primari eliminando ogni interpretazione soggettiva. Per la prima volta, il computer offriva, rispetto al tradizionale approccio storico-culturale, la possibilità di inferire analitiche ricostruzioni socioculturali prive di qualsiasi forma di influenza soggettiva da parte del ricercatore. Tuttavia, l'assenza di sistemi di descrizione e classificazione universali limitava una più estesa diffusione dell'informatica in archeologia (Whallon 1972); la persistenza di campanilistiche e personalistiche tipologie rendeva impossibile la costruzione di banche dati generaliste.

Tra gli anni '60 e '70 dello scorso secolo nasceva e si diffondeva l'Archeologia Digitale concentrando analisi e riflessioni teoriche ancora su pochi campi di interesse come la statistica inferenziale, finalizzata alla ricerca, e il controllo di grandi quantità di dati. Molte delle problematiche emerse durante i primi tentativi di uso del computer per il trattamento dell'informazione testuale permangono anche nella più recente progettazione di sistemi per la gestione dei beni archeologici. Mentre le applicazioni di tipo statistico appaiono, oggi, in gran parte legate al mondo dell'archeometria, si registra ancora una costante attenzione al tema delle banche-dati e degli archivi.

Uno degli argomenti cruciali, sollevati oltre 70 anni fa, è quello della mancanza di una terminologia standard, solo in parte risolto dalla creazione e diffusione di thesauri. La principale preoccupazione, espressa dagli archeologi, riguardava l'uso di un lessico facilmente riconoscibile in grado di supportare lo studioso nella ricerca dei dati. Per raggiungere l'obiettivo vennero organizzate conferenze ed elaborati sistemi di archiviazione e liste terminologiche abbastanza flessibili da poter raggruppare una ampia serie di collezioni di dati (Whallon 1972).

I temi della codifica erano stati già trattati negli anni '50 dal *Centre d'Analyse Documentaire pour l'Archeologie* (C.A.D.A.) che aveva in Gardin il suo più importante riferimento per le questioni della formalizzazione e descrizione dei dati. Lo studioso francese (Gardin 1971) osservava che per costruire una banca dati fosse innanzitutto necessario un certo grado di accordo sulla natura dei dati tra gli specialisti; soltanto dopo poteva essere scelto un adeguato linguaggio descrittivo per identificare le proprietà da registrare. Mentre nelle scienze esatte e in altre di-

scipline umanistiche tali requisiti venivano riconosciuti e utilizzati, in campo archeologico si registrava una maggiore diffidenza nell'uso di strumenti, anche solo terminologici, comuni. Per Gardin soltanto un radicale cambiamento dei metodi e delle abitudini professionali avrebbe incrementato l'uso del computer in archeologia. Nonostante ciò, la previsione dello studioso francese è ottimistica:

Before twenty years are out, we shall see the rapid, large-scale proliferation of international banks of archaeological data similar to those already taking shape in certain countries for museum collections (Gardin 1971, 198).

Al C.A.D.A. va riconosciuto il merito di aver posto per la prima volta la giusta enfasi su argomenti quali i principi formali e logici in base ai quali devono essere costruiti rigorosi sistemi di descrizione e di aver messo a disposizione della comunità scientifica alcuni tracciati per l'uso pratico nelle analisi. La distinzione introdotta tra banche dati generaliste, tipiche della catalogazione museale, e settoriali, specifiche, invece, per la ricerca, è un altro elemento chiave dell'apporto teorico del gruppo di ricerca francese.

Gli argomenti trattati dai pionieri dell'Archeologia Digitale in relazione alla codifica del dato archeologico in forma testuale, numerica o simbolica, non appaiono molto distanti da alcune riflessioni e criticità registrate di recente nello stesso settore applicativo nonostante l'evoluzione delle metodologie ed una maggiore competenza informatica degli addetti ai lavori. Nei seguenti paragrafi, vedremo come dalla presunta oggettività del record digitale si sia passati all'esigenza di comprendere meglio la natura del dato studiando e cercando di formalizzare l'intero percorso di conoscenza e non soltanto le proprietà del singolo oggetto.

## 2.2. I dati testuali

Nel corso dello scavo l'archeologo trasforma gli strati e i materiali che rimuove in fenomeni e concetti per noi visibili e, quindi, comprensibili. Il dato archeologico assume la funzione di elemento base del processo interpretativo, e, pertanto, la sua osservazione e misurazione influenzano l'intero percorso di ricostruzione storica. Generalmente la distinzione tra dato e record archeologico coinvolge l'insieme di tecniche adoperate per

l'acquisizione e la registrazione dell'informazione e l'approccio teorico che l'archeologo proietta su ciò che emerge da uno scavo. Il dato rappresenta la natura empirica e storica dell'oggetto o eco-fatto rinvenuto: esso ha una sua materialità e/o forma, superficie interna/esterna, decorazione, etc. che possono essere descritte e misurate utilizzando un linguaggio scientifico convenzionale. La collocazione della traccia archeologica in uno specifico contesto culturale trasforma il semplice dato in un record che assume un significato in quanto evidenza di fenomeni passati.

Mentre la documentazione dell'oggetto può costituire una fase embrionale del processo di conoscenza, la classificazione e la catalogazione della traccia e la sua trasformazione in record fanno riferimento ad un mondo noto o del quale si conosce, anche se parzialmente, il modo di rappresentazione. In ciò risiede la differenza tra l'archeografia (dal greco ἀρχαῖος e γράφειν) e l'archeologia.

Il termine archeografia, per identificare la fase della ricerca archeologica destinata alla registrazione e documentazione del dato, è stato introdotto da Moberg (1981); l'espressione identifica un settore di studi dedicato all'analisi delle modalità di acquisizione e restituzione della documentazione archeologica (D'Andrea 2022c). In quanto descrizione accurata delle antichità, l'archeografia viene comunemente considerata il primo livello scientifico dell'archeologia moderna introdotto da Goethe, Vulpinus, Busching e Lindenschmidt fondatori del Romisch-Germanische Zentral Museum a Mainz (Schnapp 1996). L'archeografia si interroga sulle modalità di rappresentazione (grafica, fotografica, schedografica, etc.) delle tracce archeologiche. Mannoni (2000) ha precisato che l'archeografia si propone di studiare i metodi che servono ad acquisire e descrivere in modo oggettivo i dati stessi; l'identificazione di uno specifico settore di studi indirizzato all'indagine sui metodi di documentazione avrebbe il vantaggio di creare dati oggettivi su cui elaborare ricostruzioni prive della soggettività dello scopritore.

Come ci ricorda, tuttavia, T. Kuhn (1962) non si può ottenere una oggettività in assenza di un linguaggio descrittivo indipendente da un modello ed una teoria già noti e, quindi, tutte le presunte rivendicazioni di oggettività dell'archeografia restano disattese. Oggi, al contrario, siamo consapevoli che le diverse tecniche di acquisizione e registrazione dei dati, messe a disposizione dell'archeologo, forniscono una serie am-

pia di informazioni, talvolta eccessivamente ridondanti al punto da definire questa massa di dati (testuali, tabellari, spaziali, fotografici, grafici, etc.) come un vero e proprio diluvio difficile da arginare (Bevan 2015).

La vasta e quasi infinita disponibilità di dati impone una riconsiderazione della funzione dell'archeologo nel processo di costruzione del record archeologico e di condivisione e riuso dell'informazione (D'Andrea 2016; 2022b; 2022c) soprattutto considerando che molto spesso gli studi archeologici vengono effettuati sulla riproduzione grafica, in larghissima misura a stampa, di oggetti, monumenti e/o contesti.

Dopo un lungo periodo nel quale le indagini sul terreno sono state considerate un esperimento distruttivo, oggi gli studiosi ritengono che lo scavo perda la sua originaria connotazione di attività negativa di rimozione delle tracce del passato per trasformarsi in un momento di forte inventiva e creatività finalizzata alla costruzione di un apparato documentario composto da vari media che fungono da base scientifica per le successive fasi interpretative. Se il dato archeologico è il risultato di una trasformazione, la documentazione non può rappresentare l'esito di un processo oggettivo (Olsen *et al.* 2012), ma piuttosto uno sviluppo selettivo che tramuta la pesantezza dei materiali nella leggerezza delle parole, per usare una felice espressione di Carandini (2000, 18).

L'esigenza di registrare le varie fasi dello scavo impone all'archeologo un percorso di documentazione nel quale la continuità della storia della terra è scomposta in segmenti distinti; la discretizzazione archeologica, cioè la definizione dei limiti di ogni traccia archeologica, avrebbe l'obiettivo di isolare unità finite associabili a specifici media attraverso i quali il passato diviene conoscibile. La conoscenza archeologica risulta alla fine l'atto conclusivo di una attività che non scopre una realtà storica, ma la traduce in forma grafica e/o testuale, in modo spesso parziale e frammentario, affinché sia intellegibile. In questa prospettiva la documentazione non costituisce la testimonianza del lavoro dell'archeologo che ha distrutto la realtà antica stratificata, ma sarebbe il risultato di un potere immaginario in grado di rendere visibile all'archeologo gli avvenimenti del passato.

La produzione di nuovi dati, spesso legati all'ultima tecnologia disponibile, nasconde l'assenza di ogni attenta valutazione sulla forma di scrittura del dato archeologico, soprattutto se si considera quanto il digitale estenda i tradizionali limiti della ricerca e dei suoi metodi di indagine.

### 2.3. La formalizzazione

La scoperta archeologica non è soltanto l'esito di un processo scientifico basato sull'analisi e l'interpretazione dei materiali portati alla luce e del contesto che ha restituito l'evidenza archeologica esaminata. Spesso lo studio dei resti viene affrontato dall'archeologo con uno specifico background culturale che condiziona il risultato finale dell'indagine. Contenuto informativo dei dati e approcci teorici si mescolano e si sovrappongono in una sintesi che perde qualsiasi condizione oggettiva. In talune circostanze la stessa metodologia di scavo e la descrizione dei ritrovamenti può costituire un fattore che introduce elementi di forte soggettività nella classificazione dei materiali; anche la selezione delle tracce ritenute significative e diagnostiche influenza il successivo percorso ricostruttivo ed interpretativo.

Se la metodologia di scavo ha la funzione di rendere comprensibile il metodo di scoperta e, quindi, di porre tutte le evidenze in un contesto di riferimento tecnico-operativo le cui fasi di distruzione degli strati sono comunemente accettate, la registrazione dei dati viene circoscritta a quegli elementi che l'archeologo ritiene più significativi per le successive analisi. Pertanto, la procedura di registrazione rappresenta un momento critico della prassi scientifica del ricercatore; da un lato si distrugge concettualmente e fisicamente una realtà antropica e naturale stratificata e dall'altro se ne crea un surrogato digitale adoperando un linguaggio descrittivo collegato, talvolta, ad un particolare quadro teorico.

Per gli storici delle scienze (Kuhn 1962) il processo di codifica dell'informazione non è mai il risultato di una azione oggettiva e/o neutrale perché non esiste un linguaggio che sia in grado di spiegare un fenomeno senza sovrapporre alla descrizione stessa del fatto esaminato una specifica teoria. Da ciò si ricava che i dati hanno un significato solo all'interno di un modello di riferimento e, quindi, l'osservazione sistematica, la misurazione e la quantificazione non determinano automaticamente l'oggettività del dato registrato. Se ad un insieme di conclusioni scientifiche è possibile far coincidere più di una costruzione teorica, allora l'intero processo interpretativo non può che rappresentare il risultato dell'impiego di un linguaggio parziale.

L'assenza di una forma di rappresentazione universale in grado di descrivere e decifrare qualsiasi fenomeno, si traduce in campo archeolo-

gico nella adozione di protocolli comunemente accettati. G. Lock (2003) sostiene che la maggior parte degli archeologi fanno riferimento a pratiche e metodi tradizionali sperimentati da lungo tempo piuttosto che seguire in modo consapevole approcci teorici. La professionalizzazione dell'attività archeologica e la comparsa di quella commerciale, legata agli interventi di emergenza, ha probabilmente contribuito a ridimensionare l'attenzione al ruolo che i diversi approcci teorici possono avere sulla registrazione del dato finale. Nello stesso tempo, poiché il computer è entrato a far parte degli utensili di base dell'archeologo, ogni approccio critico-riflessivo sull'uso delle tecniche digitali nelle indagini sul terreno è stato abbandonato in favore di una migliore gestione e qualità dell'informazione (D'Andrea 2015b).

In realtà la necessità di una adeguata riflessione sulle forme di registrazione dei dati emerge in tutta la sua criticità come abbiamo visto in precedenza nel caso del riuso dei *legacy data* e dei vecchi archivi digitalizzati; questa considerazione è ancora più importante se i dati sono stati conservati in una forma tabellare o in schede che possono contenere un implicito modello di riferimento.

La corretta riflessione sulla natura dell'informazione archeologica, e del discorso interpretativo che su di essa lo studioso può costruire, rappresenta un elemento centrale per la scelta della migliore procedura da adottare per la strutturazione dei dati. Ai tradizionali metodi di indagine si affianca, nel percorso digitale, un altro aspetto teorico e concettuale altrettanto importante, quello della codifica, spesso derubricata a pratica neutrale e non meritevole di ulteriori approfondimenti. Da una narrativa, o da un insieme di record descritti in linguaggio naturale, il ricercatore deve ottenere una struttura semplificata basata su regole rigide che danno una forma ai dati esaminati.

Dal punto di vista del trattamento informatico, l'elaborazione di un dato digitale avviene dopo la conclusione di un processo di codifica che converte la materialità dell'oggetto, o del fenomeno, in unità minime, fisiche o concettuali, caratterizzate da parametri e valori che possono essere gestiti e manipolati da un calcolatore mediante appositi programmi. Questa operazione è il risultato di un procedimento, spesso empirico, eseguito dal ricercatore per la identificazione delle categorie nelle quali classificare gli oggetti, qualunque sia la loro natura. In tal senso, la

codifica è una attività di selezione delle variabili da registrare e calcolare in funzione di un determinato obiettivo; la natura intenzionale e soggettiva della scelta si riflette necessariamente sui dati.

I principi su cui si fonda il procedimento sono quelli della logica formale che consentono di tradurre i contenuti concreti del linguaggio comune in simboli che li rappresentano, passando da una conoscenza intuitiva ad una rigorosa; il passaggio si fonda sulla definizione di un modello della realtà descritto attraverso l'identificazione di alcune proprietà distintive degli oggetti osservati. Spesso il concetto di simbolo è stato interpretato, soprattutto agli albori dell'archeologia digitale, in chiave quantitativa e la formalizzazione si è tradotta in una discretizzazione numerica delle variabili più vicina al linguaggio matematico dei computer che alla descrizione archeologica; così sono state preferite proprietà o variabili in grado di rappresentare le dimensioni di oggetti o fenomeni ricorrenti oppure le occorrenze; più complessa risulta, senza dubbio, la strutturazione rigorosa di elementi incompleti, parziali, ambigui o qualitativi.

La formalizzazione è, in sintesi, un processo che definisce un insieme di regole che connettono dati e modelli; sebbene il procedimento sia utilizzato per modellizzare una realtà complessa attraverso un rigoroso formalismo di tipo logico-matematico, gli esiti finali della formalizzazione non sono oggettivi, né neutrali, poiché le caratteristiche distintive del modello astratto sono selezionate dal ricercatore in base ad uno specifico obiettivo di studi.

Per Orlandi (1997) alla base del processo di formalizzazione si devono identificare due componenti: i dati che rappresentano singoli elementi della realtà e i simboli che devono esprimere in termini informatici quella specifica realtà. La codifica digitale si traduce nell'utilizzo di uno specifico software, in prevalenza un database relazionale, la cui architettura in tabelle si adatta con maggiore fedeltà al tipo di realtà fisica che l'archeologo indaga e descrive.

## 2.4. Dal Semantic Web alla Scienza Aperta

Come abbiamo sottolineato in precedenza lo sviluppo dell'Archeologia Digitale segue coerentemente l'evoluzione delle tecniche informatiche; in particolare, per quanto riguarda il trattamento del record scritto, le

trasformazioni nell'elaborazione dell'informazione testuale derivano dalla espansione delle reti telematiche e dalla messa a punto di strumenti per l'indicizzazione, la ricerca e l'aggregazione dei dati.

Lo sviluppo di Internet dalla sua comparsa ad oggi è stato scandito da significativi cambiamenti concettuali e strutturali connessi principalmente alla definizione di una nuova interazione caratterizzata, sempre di più, dalla partecipazione attiva degli utenti che producono e scambiano dati.

Il Web 1.0 per lungo tempo ha rappresentato un mondo statico nel quale pubblicare documenti creati con un linguaggio semplice, l'HTML, ma efficiente. L'uso della rete era limitato, essendo appannaggio di pochi centri di ricerca, per la maggior parte militari; lo scambio dei dati era unilaterale con fruitori in gran parte passivi. Le relazioni tra documenti erano assicurate da link ipertestuali che collegavano, in una forma pre-determinata dall'autore della risorsa Web, contenuti differenti anche ospitati in altre pagine.

Con la diffusione delle connessioni e l'aumento degli utenti, la rete è diventata non solo un mezzo per veicolare i dati, ma anche un enorme mercato che ha visto all'attivo differenti protagonisti come gli sviluppatori di siti web da un lato e le nuove piattaforme commerciali per la vendita online dall'altro. L'introduzione del Web 2.0, che aveva l'obiettivo prioritario di connettere le persone, ha ulteriormente modificato la funzione della rete consentendo ad una grande massa di utenti di pubblicare dati e documenti (video e foto) e di scambiare contenuti in forma interattiva.

I motori di ricerca e le piattaforme social sono i più significativi risultati di questa trasformazione della rete finalizzata a supportare fruitori, prodotti e servizi. Con il Web 2.0 l'utente può accedere in remoto ad applicazioni sviluppate direttamente sulla rete che favoriscono una separazione tra il produttore di tecnologia, che mette a disposizione l'infrastruttura digitale, e il consumatore che può creare contenuti e condividerli con altri utenti della rete. Questa caratteristica contribuisce alla nascita di una intelligenza collettiva, collaborativa e diffusa di cui, per esempio, Wikipedia è il migliore esempio. D'altro canto, la creazione di piattaforme per il mercato concorre alla fortuna di molte aziende che investono nell'ecosistema delle informazioni.



Il Web 2.0 è, a ragione, considerato un ambiente prevalentemente commerciale nel quale le imprese possono espandersi facilmente raggiungendo mercati e consumatori potenzialmente infiniti. Anche la ricerca trae vantaggio da questa trasformazione soprattutto grazie alla diffusione di strumenti collaborativi che accrescono la cooperazione e le attività di co-working. L'eliminazione di barriere e la velocità nella trasmissione contribuisce allo sviluppo di forme nuove di lavoro accentuando apporti interdisciplinari. La rete riduce le distanze favorendo una più rapida circolazione delle idee e la conoscenza è fruita in forma attiva poiché l'utente può facilmente interagire con i produttori di contenuti.

Il Web 2.0 ha mutato i canali tradizionali della comunicazione scientifica che, però, resta ancora affidata allo scambio rapido di informazioni tra utenti. La rete supporta un processo di democratizzazione della scienza e del sapere che diventano accessibili, comunicabili e riutilizzabili soprattutto grazie alla diffusione di tecnologie open-source e *cloud-computing*. Nel paradigma del *citizen science*, l'attività dello scienziato è inserita nei processi di formazione del consenso pubblico cui spetta un inedito ruolo di valutatore della ricerca oltre quello talvolta di protagonista nello svolgimento della ricerca.

Senza dubbio il Web 2.0 ha svolto una importante funzione nella ottimizzazione della scienza, trasformando la rete in luogo e veicolo per idee, riflessioni e, quindi, partecipazione diretta. Le nuove sfide generate dalla affermazione di un nuovo modello di conoscenza hanno posto le basi per un ulteriore salto qualitativo determinato in larga misura dall'esigenze di gestire in modo efficiente, e possibilmente automatico, l'ingente quantità di risorse presenti nel Web. L'estensione più recente della rete diventa uno spazio nel quale Internet acquista nuovamente la sua originaria dimensione di luogo per l'accesso e lo scambio di dati con l'aggiunta del contesto interpretabile anche da una macchina.

Considerata la straordinaria quantità di risorse disponibili, gli strumenti tradizionali di ricerca si sono, presto, rivelati insufficienti per estrarre dati ragionati, in relazione tra di loro e corrispondenti alle esigenze dell'utente. Così il Web 3.0, o Semantic Web, da semplice contenitore di testi, video e immagini si è trasformato in un enorme database di informazioni che possono essere estratte anche automaticamente. Alla base di un tale profondo mutamento si pone lo sviluppo di strumenti che

associano, mediante particolari dati definiti metadati (*infra* 108), la risorsa ad uno specifico contesto eliminando o riducendo qualsiasi ambiguità.

L'introduzione del Semantic Web ha dato origine ad una modalità differente di produzione e fruizione di contenuti che, per essere pienamente condivisi, devono essere codificati in accordo ad una specifica struttura con formati aperti e connessi ad altri documenti digitali; la rete di dati e le relazioni reciproche diventano interpretabili anche da una macchina. L'informazione diviene, dunque, accessibile, disambiguata, perdendo o riducendo il suo carattere polisemico, e interoperabile, cioè integrabile in altri sistemi. L'internet dei documenti si è trasformato in un Web dei dati che ha incoraggiato un modo diverso di fare ricerca, favorendo forme ampie di cooperazione e condivisione dei saperi.

Lo scenario, originato nel dominio della conoscenza dai cambiamenti introdotti dal Web, viene definito con il termine Open Science (David, Spence 2003) per descrivere la proprietà dei dati scientifici generati dal settore pubblico in opposizione all'estensione percepita dei diritti di proprietà intellettuale nell'area dei beni informativi. Gli economisti considerano la conoscenza scientifica scaturita dalla ricerca pubblica come un bene comune; il riconoscimento del carattere pubblico della ricerca implica che tutti possono utilizzare la conoscenza, una volta edita, senza costi aggiuntivi, generando rendimenti sociali più elevati. La Scienza Aperta nell'era dell'informazione digitale abbraccia, dunque, l'idea che la conoscenza abbia le caratteristiche di un oggetto che va oltre il semplice concetto di bene comune sviluppato nel XVIII secolo, poiché l'accesso consentito dalle reti informatiche amplia le possibilità di arricchire i beni pubblici estendendone la fruizione verso una platea più ampia di utenti.

La scienza moderna è caratterizzata da una serie di principi di cui il più importante è la considerazione del valore universale delle scoperte scientifiche, vale a dire che l'innovazione è vista come il prodotto della collaborazione sociale e il suo sfruttamento è assegnato alla comunità.

La corsa per essere i primi a rivendicare il riconoscimento, la cosiddetta regola di priorità, nella scienza è stata tradizionalmente un forte incentivo per i ricercatori per rendere pubbliche le loro conoscenze. Mentre questo sistema funziona con l'attuale sistema di revisione tra pari e la pubblicazione in riviste esclusive, spesso a pagamento, la rivoluzione

informatica ha mutato, se non gli ideali alla base della scienza, certamente il sistema di produzione e promozione della conoscenza scientifica.

La diffusione delle tecnologie digitali per la comunicazione, conservazione ed elaborazione dei dati ha incoraggiato una radicale trasformazione della natura della scienza e dell'innovazione che si fondano su infrastrutture digitali, metodi e servizi Web e ambienti di ricerca collaborativi e virtuali. Il cambiamento è il risultato della combinazione di più effetti concomitanti che includono, da un lato, gli avanzamenti delle tecnologie e, dall'altro, l'emergere di un mutamento culturale contraddistinto da una richiesta di maggiore efficienza, efficacia e trasparenza.

Una Scienza Aperta favorisce nuove pratiche scientifiche, discipline e paradigmi che rispondono alle sfide di un mondo globalizzato nel quale i cittadini partecipano nello stesso tempo come produttori e consumatori della conoscenza. La filosofia dell'apertura della scienza è ancora più profonda di quanto a prima vista possa apparire: l'obiettivo è rifondare l'intero processo di ricerca e di diffusione dei risultati, incentivando l'uso di media digitali e incoraggiando la collaborazione.

I pericoli di una archiviazione digitale incontrollata sono stati fortemente ridotti negli ultimi anni grazie alla diffusione di un nuovo modo di organizzare la ricerca basato sul concetto di produzione, condivisione e riutilizzo, in modo trasparente, dei dati primari. La Commissione Europea, in un documento dal titolo *Digital Science in Horizon 2020*, dedicato a promuovere lo sviluppo di una Scienza Aperta, ha analizzato i benefici di un approccio condiviso e quindi più democratico.

Cinque sono i pilastri dell'OS:

- l'accesso aperto alle pubblicazioni di ricerca su riviste o archivi di articoli disponibili gratuitamente migliora la rilevabilità della conoscenza scientifica su una nuova scala, rendendo necessari nuovi strumenti per la formalizzazione e l'interoperabilità dei depositi;
- la messa a disposizione dei dati alla base delle pubblicazioni consente la riproducibilità della ricerca da parte di altri, la verifica dei risultati e il riutilizzo dei dati per nuovi scopi;
- la trasparenza che arricchisce la conoscenza attraverso la pubblicazione di esperimenti e risultati che non sono pubblicati nelle riviste scientifiche tradizionali, ma in luoghi condivisi e accessibili;

- la creazione di nuove soluzioni tecnologiche, combinate con l'accesso aperto alle risorse di ricerca, va nella direzione di semplificare i riferimenti incrociati tra pubblicazioni, dati e autori supportando una nuova tracciabilità dei flussi di ricerca e la creazione di nuove connessioni tra temi di indagini e attori. Tali collegamenti offrono nuovi mezzi per valutare il valore della scienza seguendo il riutilizzo del lavoro scientifico, sostituendo le tradizionali metriche basate sull'editore (*impact factor*) con quelle orientate alla ricerca (*h-index*);
- l'apertura delle risorse di ricerca facilita il controllo e il *feedback* da parte dei colleghi e persino del pubblico già durante il processo di ricerca, permettendo di affinare l'approccio e di migliorare la qualità dei risultati attraverso contributi aggiuntivi.

Anche il settore della ricerca archeologica ha beneficiato di una tale profonda rivisitazione epistemologica che ha radici in un modo diverso di diffondere la scienza oltre le comunità specifiche di riferimento (Moscati 2017). L'espansione del fenomeno dell'OS ha contribuito ad una più decisa consapevolezza nell'uso degli strumenti informatici imponendo anche codici di comportamento finalizzati a ridurre il rischio della dispersione degli archivi digitali. L'affermazione del paradigma dei BD ha spinto verso una accelerazione di questa trasformazione incoraggiando molti ricercatori a scegliere soluzioni open (software e di formato) rispetto a sistemi di tipo proprietari.

La pubblicazione online dei materiali di ricerca costituisce, quindi, il principale banco di prova per la costruzione di un approccio che non si limita ad illustrare cosa il computer fa, ma che esamina, invece, quanto il lavoro del ricercatore possa risultare profondamente trasformato dalla contaminazione digitale. Se l'archeologia è una scienza che si basa sull'analogia e sul confronto, soltanto i dati scientifici, correttamente descritti ed archiviati, consentono di costruire quella base di conoscenza valida, analogica o digitale, alla quale ancorare qualsiasi ipotesi o ricostruzione storica. Il rischio, sempre presente nella documentazione e nei processi di registrazione, è che una forma errata di descrizione/rappresentazione possa compromettere qualunque interpretazione o teoria.

Nei seguenti paragrafi vedremo alcune delle soluzioni messe a punto per supportare l'integrazione, la standardizzazione e la connessione tra

fonti digitali eterogenee. Alcune delle proposte presentate restano ancora circoscritte ad una cerchia di esperti, in gran parte informatici ed ingegneri della conoscenza. In questo panorama la maggioranza degli archeologi appare orientata al mantenimento degli approcci e dei metodi più tradizionali di ricerca. In realtà, come è stato evidenziato introducendo i cambiamenti sostenuti dalla *digital transformation*, gli stimoli esterni, di tipo tecnico e culturale, produrranno nuove traiettorie di sviluppo coinvolgendo anche il futuro della ricerca archeologica.

#### 2.4.1. L'architettura del Semantic Web

La Scienza Aperta, oltre che su concetti e visioni futuribili, si basa sull'implementazione di una serie di tecnologie il cui fine non è quello di supportare lo scambio di informazioni, ma di consentire lo sviluppo di strumenti affinché le macchine possano autonomamente estrarre e dedurre conoscenza dai dati disponibili in rete.

Mentre per gli esseri umani combinare informazioni è una attività semplice, spesso eseguita a livello intuitivo o sulla base di una conoscenza pregressa, una macchina non è in grado di comprendere immediatamente le relazioni tra fonti eterogenee, incomplete e conservate in formati differenti. I motori di ricerca hanno la capacità di rispondere in modo efficace alle particolari esigenze di selezione delle risorse in rete, ma non possono interpretare il ragionamento umano elaborato per definire una specifica interrogazione, né possono suggerire abbinamenti con altri concetti/espressioni.

Attualmente la componente semantica di una ricerca emerge attraverso il ricorso ad algoritmi che identificano i contesti in base alle parole adiacenti; i motori eseguono analisi testuali che identificano le parole associate ad una richiesta per cercare di affinare i risultati. Conoscendo il meccanismo di funzionamento, l'utente che deve ricercare un termine preciso generalmente aggiunge altri vocaboli per circoscrivere la ricerca.

Il limite dell'attuale Web, nella selezione di specifiche informazioni, consiste nella semplice osservazione che di frequente le parole da ricercare sono caratterizzate da una ambiguità semantica (ad esempio: Afrodite statua, dipinto, mito, centro estetico, esercizi commerciali, etc.). Per ridurre questi rischi, o quanto meno per fornire alcuni accorgimenti per ottimizzare le interrogazioni, Google ha sviluppato specifici strumenti,

tra cui le Google *Dorks*, che hanno l'obiettivo di creare query in grado di restringere i risultati della ricerca. Per quanto efficaci, i comandi non possono aggregare informazioni, né realizzare interrogazioni semantiche. Poiché le Google *Dorks* non consentono l'esecuzione di ricerche semantiche, il Google *dorking* è piuttosto una tecnica di ricerca avanzata o una strategia utilizzata per cercare dettagli nascosti che potrebbero essere disponibili sui siti Internet.

L'estensione della rete, insieme con l'incremento pressoché infinito delle risorse digitali, ha dimostrato l'insufficienza dei tradizionali motori di ricerca nella esecuzione di banali compiti di selezione e soprattutto di organizzazione delle informazioni. Allo scopo di supportare la nuova fase della crescita di Internet, è stata messa a punto l'architettura di un Web basato sullo sviluppo di un sistema di rappresentazione della conoscenza, capace di disambiguare il mare di informazioni sparse in rete; questo approccio comprende un meccanismo di raffigurazione delle connessioni tra concetti/termini/dati codificato mediante un linguaggio che modella le relazioni. Il Web Semantico, cioè elaborabile autonomamente da una macchina, richiede un insieme strutturato di informazioni e di regole di inferenza.

L'infrastruttura opera partendo dai metadati che costituiscono informazioni strutturate in classi di oggetti per uno specifico obiettivo e identificate univocamente da indirizzi Web permanenti o URI. Il linguaggio adoperato per modellare i metadati, denominato RDF, consente di memorizzare i concetti mediante un meccanismo di dichiarazione o *statement* rappresentato da triple che legano un soggetto (la risorsa Web) ad uno oggetto (il suo valore/contenuto) attraverso un predicato/proprietà. Affinché le asserzioni o triple rappresentino un significato in un certo dominio, le classi di oggetto devono essere codificate in una struttura formalizzata o RDFS che vincola il range delle dichiarazioni, dei valori e del formato.

La descrizione delle risorse, anche se formalizzata con metadati che esprimono concetti digitali presenti in rete tramite URI, non è sufficiente a rappresentare in forma comprensibile per una macchina il sapere di un dato settore della conoscenza. Per raggiungere questo obiettivo occorre che lo schema definisca anche delle regole per collegare, con relazioni e proprietà ben identificate, concetti differenti. In tal modo la tripla corri-

sponde ad una sequenza logico-sintattica determinata dal modellatore in uno schema che assume il nome di ontologia. Per potenziare l'espressività semantica ed eseguire dei ragionamenti automatici si adopera un linguaggio più complesso, l'OWL, che descrive le relazioni che intercorrono tra le risorse introducendo costrutti che vincolano le relazioni tra le classi.

L'architettura del Semantic Web è costruita su più livelli di cui quello dei metadati costituisce il gradino più basso e, quindi, indispensabile per la costruzione dell'intero processo. I metadati devono essere comprensibili per la macchina e organizzati in una forma che rispetti le convenzioni di una data comunità vincolando la semantica, la sintassi e la struttura delle descrizioni.

La sfida lanciata dal Semantic Web va nella direzione di identificare e progettare strumenti adeguati affinché i dati siano accessibili, processabili ed interoperabili per essere aggregati e ricercati. Per questo scopo sono necessarie infrastrutture di rete in grado di trasformare l'architettura concettuale del Web 3.0 in un sistema fisico capace di archiviare, gestire e rendere disponibili online i dati.

#### 2.4.2. Open Data e *Linked Open Data*

Con il termine dati si identificano in rete una pluralità di risorse digitali ospitate in qualche parte del Web, come numeri, immagini, statistiche, video, audio, presentazioni, testi. Se la ricerca deve fondarsi su dati aperti e disponibili nell'ecosistema digitale, una delle principali sfide poste dall'OS è come rendere gli archivi facilmente identificabili, essendo dispersi in uno spazio pressoché infinito di luoghi virtuali o in database chiusi. Affinché tale materiale sia accessibile, è necessario organizzare i dati in modo tale che se ne possa sempre identificare la provenienza, l'origine, gli eventuali diritti di proprietà e le forme del suo potenziale riuso ristretto o del tutto aperto.

Uno dei pilastri su cui si fondano il Semantic Web e l'OS è il concetto di dati aperti, la cui importanza è probabilmente maggiore rispetto agli aspetti della loro implementazione, caratterizzata da architetture, standard, linguaggi, protocolli e servizi. L'*Open Knowledge Foundation* definisce un dato aperto come

un dato che può essere liberamente acceduto, utilizzato, modificato e condiviso da chiunque e per qualunque scopo.

Si tratta quindi dell'insieme di informazioni rese disponibili da chi produce documenti senza alcuna restrizione di copyright, brevetti o altre forme di controllo che ne limitino la riproduzione.

I dati aperti devono avere specifici requisiti:

- essere utilizzabili con una licenza o una previsione che ne permetta l'uso anche per finalità commerciali e in forma disaggregata;
- essere accessibili attraverso reti telematiche pubbliche e private, in formati aperti e provvisti dei relativi metadati;
- essere disponibili gratuitamente, oppure resi accessibili a costi contenuti, per esempio per la loro sola riproduzione e divulgazione.

Il rispetto di questi semplici principi rende concreto uno dei principali obiettivi degli Open Data e cioè l'interoperabilità che può essere definita come la capacità di combinare dati provenienti da fonti differenti per produrre nuova conoscenza.

La filosofia dei dati aperti riflette l'idea che una maggiore trasparenza nella gestione della spesa pubblica e della ricerca scientifica, possa favorire un'ampia circolazione delle informazioni, promuovendo la partecipazione e il coinvolgimento di utenti meno esperti.

Gli Open Data sono una immensa risorsa informativa prodotta da istituzioni e privati cittadini in gran parte in forma spontanea e senza un disegno preciso, nonostante sia chiaro come il loro impiego possa stimolare un consistente valore aggiunto e contribuire a significativi vantaggi anche di tipo economico.

Per la creazione di risorse aperte non esistono specifiche regole, ma possono essere seguite alcune raccomandazioni che dovranno guidare il produttore nella organizzazione e pubblicazione dei dati. In particolare, occorre:

- scegliere la licenza migliore per rendere pubblici i dati;
- strutturare inizialmente piccoli dataset di Open Data; una partenza semplice può aiutare a comprendere al meglio i meccanismi per l'apertura dei dati;
- coinvolgere i potenziali utenti e fruitori degli archivi per migliorare l'organizzazione dei dati e, quindi, il loro potenziale riuso;



- esaminare gli eventuali effetti, legali ed etici, derivanti dalla pubblicazione dei dati;
- individuare il formato aperto più efficiente per la tipologia di dati che si vogliono aprire consentendo lo scaricamento dei file; se, per esempio, dati statistici vengono offerti in PDF l'utente dovrà riscrivere i valori con il rischio di introdurre errori;
- creare una pagina web dove ospitare gli archivi aperti inserendo, se possibile, un motore che ne consenta la ricerca;
- offrire, laddove possibile, la possibilità da parte di terzi di aggiungere dati al proprio archivio.

Un raccordo con la comunità di riferimento aiuterà certamente il produttore di contenuti nella implementazione di tutti i passaggi indicati nei suggerimenti per l'apertura dei dati.

Alcune particolari raccomandazioni riguardano anche gli aspetti tecnici degli Open Data che dovrebbero essere *machine-readable*, cioè comprensibili non solo per l'uomo, ma anche per una macchina.

La diffusione del paradigma dei dati aperti ha contribuito alla nascita di agenzie nazionali con l'obiettivo di coordinare le varie iniziative locali fornendo linee guida per valorizzare il patrimonio informativo pubblico.

In Italia l'AGID ospita il Catalogo Nazionale dei Dati Aperti e promuove l'interoperabilità per favorire i processi di accesso e scambio delle informazioni in coerenza con il relativo framework europeo. L'Agenzia ha rilasciato un profilo italiano, coerente con le regole e gli standard di interoperabilità definiti a livello internazionale ed europeo, che raggruppa l'insieme delle raccomandazioni utili per descrivere i dati in maniera omogenea e standardizzata; attualmente l'AGID rende disponibili circa 55.000 dataset anche di tipo geografico.

A livello europeo è stato implementato il portale ufficiale dei dati aperti della Commissione Europea che ospita 176 cataloghi, prodotti da 36 paesi per un totale di oltre 1.600.000 archivi accessibili liberamente e scaricabili in vari formati per le immagini, i dati testuali o tabellari, e per i dataset geografici.

La creazione e la promozione dei dati aperti si fonda su una filosofia e su una pratica che richiedono poche raccomandazioni: dati aperti, accessibili, senza restrizioni o controlli. Per queste caratteristiche il feno-

meno Open è considerato il risultato di un movimento eterogeneo contraddistinto dai principi della trasparenza, partecipazione e collaborazione, piuttosto che l'esito di un insieme strutturato di soggetti coinvolti in un coerente processo di pubblicazione di dati.

Nonostante l'ampia disponibilità di dati aperti e l'emergere di una concreta consapevolezza sull'esigenza di rendere fruibile il materiale presente in rete, la semplice apertura dei dati non è sufficiente a garantire un coinvolgimento ed una partecipazione altrettanto significativa e soprattutto l'aggregazione di fonti differenti per produrre nuova conoscenza. Per superare questo limite è stato identificato il modello dei LOD il cui obiettivo è produrre informazioni autodescrittive che possano essere facilmente reperite in rete e elaborate sia da essere umani che da agenti automatici. Se gli Open Data hanno consentito di rimuovere gli ostacoli legati all'accesso ed al riuso dei dati, i LOD sono la conseguenza dello sviluppo delle tecnologie messe a punto per il Semantic Web.

I LOD sono stati proposti nel 2006 allo scopo di garantire l'omogeneità nella descrizione dei dati identificando in particolare due aspetti critici nell'integrazione: l'identità (dove sono i dati) e l'interoperabilità (come aggrego i dati che ho trovato con altre risorse digitali). I due aspetti sono fortemente interconnessi perché, per risolvere il problema della interoperabilità, occorre aver la chiara identità del dato da aggregare, altrimenti i sistemi non sono in grado di dialogare tra di loro; una ulteriore complicazione deriva poi dal fatto che ogni sistema si basa su infrastrutture eterogenee e soprattutto su linguaggi differenti per cui, ad uno stesso oggetto, possono essere attribuite identità differenti determinando così evidenti rischi di inconsistenza o disallineamento dei dati.

I LOD costituiscono, non soltanto per la ricerca, una straordinaria opportunità per condividere e supportare nuova conoscenza, sfruttando l'ingente corpus di risorse digitali disponibili in rete. I dati devono, però, essere liberamente accessibili, comprensibili da una macchina e collegabili ad altri dello stesso tipo; soltanto in questo modo è possibile arricchire le nostre conoscenze e costruire una più ampia partecipazione rendendo i nostri materiali di ricerca visibili.

In uno spazio cooperativo il ruolo di chi partecipa ai benefici della Scienza Aperta non può essere circoscritto al semplice riutilizzo dei dati. Ognuno deve aggiungere e pubblicare i propri dataset aumentando

il valore complessivo di ciò che la rete offre e arricchendo le potenziali interrogazioni, connessioni e aggregazioni.

Questa filosofia, per quanto chiaramente condivisibile e frutto di una più generale trasformazione dell'idea stessa di scienza, comporta la conoscenza dell'architettura del Web Semantico e quindi come gestire e strutturare i dati secondo un formato standard riconoscibile e interpretabile dal calcolatore. Il Web dei dati richiede l'adesione ad una particolare organizzazione delle risorse che non snaturi il contenuto della conoscenza che vogliamo condividere, ma che, al contrario, richieda soltanto il rispetto di alcuni requisiti formali. I vantaggi che derivano da questo approccio sono del tutto superiori al tempo necessario per l'acquisizione di una competenza tecnica di base sulla strutturazione dei LOD.

Per aderire alla prospettiva dei dati aperti e collegati, l'accento deve essere focalizzato su come realizzare l'aggregazione tra le risorse digitali, poiché il Semantic Web non è un contenitore statico di informazioni, quanto il risultato di una rete di connessioni e di potenziali esplorazioni. Come abbiamo visto, i link tra i dati e le pagine web sono sempre esistiti fin dalla comparsa del Web 1.0. L'html, il primo linguaggio di Internet, consentiva di incorporare i collegamenti ad altri indirizzi, ad altre pagine contribuendo alla creazione ed alla estensione della ragnatela. Tutti i linguaggi attualmente disponibili supportano la funzione dei link che rappresentano, ancora oggi, la definizione di una relazione tra risorse. Nonostante la diffusione dei collegamenti ipertestuali, il Semantic Web ha definito un nuovo sistema di connessioni tra documenti che non si basa sul link esistente, ma che al contrario cerca di delineare un linguaggio formale per determinare un numero pressoché infinito di possibili interazioni. In questo caso il link fa riferimento non ad una pagina specifica, bensì ad un indirizzo permanente, definito URI, che può identificare qualsiasi specie di oggetto o concetto. Poiché l'html è un linguaggio limitato, è stata sviluppata una nuova forma di codifica basata sulle triple, cioè soggetto, predicato e oggetto, con l'obiettivo di rendere meno ambigui i dati.

Volendo ad esempio cercare sulla rete Andrea D'Andrea cerchereste un modo ottimale per ridurre il numero certamente ingente e poco significativo dei risultati. Generalmente si mettono i termini ricercati tra virgolette per individuare entrambe le parole con il rischio, però, di per-

dere risultati del tipo A. D'Andrea. Se invece Andrea D'Andrea fosse identificato da uno specifico URI, la ricerca risulterebbe semplificata ed efficiente. Una espressione del tipo Andrea D'Andrea insegna presso l'Università degli Studi di Napoli L'Orientale potrebbe essere rappresentata mediante la seguente tripla:

<https://orcid.org/0000-0002-5274-0786>  
*teaches\_at*  
[www.unior.it](http://www.unior.it)

Nella stringa, Andrea D'Andrea (Soggetto) è distinto dall'identificativo persistente o URI alfanumerico che lo identifica univocamente come autore scientifico sul registro ORCID, mentre l'università (Oggetto) presso la quale lavora è individuata attraverso il suo indirizzo web. Soggetto e oggetto sono messi in relazione attraverso la proprietà *teaches\_at*.

Vedremo più avanti, trattando il tema ontologie (*infra* 115), come le triple consentano di esprimere qualsiasi concetto in un dato dominio attraverso collegamenti tra URI che si riferiscono ad oggetti o classi di dati e relazioni o proprietà. I dati aperti diventano così autodescrittivi e disambiguati.

Per la creazione di LOD sono state definite quattro semplici regole:

- usare URI per nominare le cose.
- usare URI HTTP in modo che chiunque possa avervi accesso
- ottenere attraverso le URI informazioni utili descritte usando il linguaggio RDF/XML
- includere i link verso altri URI, in modo che sia possibile scoprire nuove informazioni.

La filosofia dei dati aperti e quella più complessa legata al Semantic Web, richiedono una precisa preparazione delle risorse da pubblicare. Per rendere più semplice la classificazione dei dati in base alle loro caratteristiche, Tim-Berners Lee ha proposto nel 2001 un modello a cinque livelli per esaminare e identificare l'accessibilità e la fruibilità tecnica dei dati aperti:

- ★ rendi i tuoi dati disponibili sul Web in qualsiasi formato e con una licenza open; si tratta di dati statici e non strutturati (documenti in formato Word, Pdf, oppure immagini in formato GIF, JPG, PNG). Sebbene disponibili online e distribuiti con licenza

aperta, questi dataset sono leggibili, conservabili e stampabili dagli utenti e non possono essere oggetto di ulteriori elaborazioni perché privi di un formato aperto;

★★ rendi i tuoi dati disponibili in una forma strutturata; questo valore identifica i dati strutturati e codificati in un formato proprietario (tabella in formato Excel) e, pertanto, non aperti. A differenza dei dati che rientrano nel precedente livello, questi possono di norma essere convertiti in un formato aperto;

★★★ rendi i tuoi dati disponibili in un formato non proprietario e aperto; si tratta di risorse strutturate e codificate in un formato aperto (ad es. CSV in sostituzione del formato XLS o XLSX di Excel). Le tre stelle identificano i dati aperti, sebbene a livello più semplice. Questo livello consente di elaborare i dati senza avere l'obbligo di utilizzare un software proprietario;

★★★★ usa URI e gli standard W3C per identificare cose e concetti, in questo modo gli utenti potranno più facilmente puntare ai tuoi dati; le quattro stelle classificano i dati strutturati e codificati in un formato aperto, così come il livello precedente, ma dotati di un URI (XML/RDF). Il livello indica tutte quelle risorse indirizzabili sulla rete e quindi accessibili online, attraverso uno specifico URL. Ciò consente ad una applicazione di poter accedere a questa tipologia di dati, visualizzarli e rielaborarli;

★★★★★ collega i tuoi dati ad altri dati consentendo di identificare meglio il contesto. In questa categoria rientrano i dati aperti che, in aggiunta al precedente livello, contengono collegamenti ad altri dataset (ad es. LOD).

Nonostante le rigide specifiche fornite per la modellazione delle informazioni e la scala di valori elaborata per valutare e classificare al meglio le risorse pubblicate, per promuovere una maggiore visibilità e accessibilità degli archivi contenenti materiali di ricerca è necessario che i dati siano conformi a terminologie standard.

#### 2.4.3. Thesauri e vocabolari

L'assenza di interoperabilità, cioè di integrazione fisica dei sistemi informativi, sembrava fino a pochi anni fa determinata dalla assenza di piat-

taforme, programmi e formati in grado di gestire fonti informative disponibili in rete. Oggi, le tecnologie semantiche mettono a disposizione dei ricercatori numerosi strumenti che contribuiscono ad assicurare la coerenza formale degli archivi e la loro consistenza. Tra gli strumenti di controllo un ruolo particolare rivestono i thesauri, i vocabolari e le tassonomie che riducono il rischio di uso di espressioni ambigue per rappresentare concetti analoghi, consentendo nello stesso tempo di migliorare la fase della ricerca delle informazioni

A differenza dei vocabolari, che contengono liste di descrizioni di concetti agevolando il compito di inserire valori omogenei nella codifica degli oggetti, i thesauri supportano ricerche complesse su ampie basi di dati, poiché usano relazioni associative tra i vari termini e possono ricercare sinonimi e varianti ortografiche.

In genere i vocabolari sono implementati in rapporto ad una specifica metodologia di classificazione e descrizione, ma falliscono quando si debba provare ad integrare contenuto e significato tra dati conservati in differenti repository. Il dizionario rappresenta uno standard all'interno di una distinta comunità di riferimento come il risultato finale di un processo di analisi storica ed anche metodologica sul quale si innestano diversi sottogruppi che specializzano ulteriormente nuovi raggruppamenti definendo terminologie aggiuntive.

Di frequente espressioni come vocabolario, dizionario, tassonomia e thesaurus sono adoperati come sinonimi. In realtà i vocabolari controllati, le tassonomie e i thesauri fanno riferimento a differenti modelli di rappresentazione della conoscenza. Un vocabolario controllato è una lista di termini resa disponibile da una autorità riconosciuta che ne garantisce l'affidabilità e la correttezza; i termini non sono ambigui e ridondanti. I termini di un vocabolario controllato dovrebbero rispettare almeno due regole:

1. se lo stesso termine è adoperato con differenti significati in vari contesti, esso non è esplicitamente qualificato a risolvere l'ambiguità;
2. se più termini sono adoperati per indicare lo stesso concetto, una delle definizioni deve essere evidenziata come termine preferito e le altre elencate come alias o sinonimi.

Una tassonomia è una collezione di termini, organizzata in una struttura gerarchica in cui ogni valore ha una o più relazioni del tipo genitore/figlio, intero/parte, genere/specie, tipo/istanza. Alcune tassonomie limitano, in genere, le relazioni al rapporto genitore/figlio, mentre altre sono poli-gerarchiche, cioè un valore può avere molteplici genitori mantenendo lo stesso significato; se un valore ha un figlio in un nodo, avrà lo stesso figlio se appare in un altro nodo della tassonomia.

Un thesaurus si distingue da un vocabolario in base ad una classificazione gerarchica dei termini espressa dalle seguenti relazioni generali:

- gerarchia: BT - *Broad Term*; NT - *Narrow Term*: indica un concetto che ha un significato più ampio o più ristretto di un altro termine. Si trova ad un livello superiore o subordinato nella categoria;
- associazione: RT - *Related Term*; indica similarità con altri concetti. Il termine associato non è un sinonimo e non esiste una relazione gerarchica;
- equivalenza: PT - *Preferred Term*: indica un termine con uno status preferito.

Le relazioni consentono di costruire una classificazione secondo gruppi concettuali, dal più ampio al più definito, strutturati anche su più livelli, sebbene non contengano norme che regolino l'uso dei termini per esprimere un significato all'interno di un dominio di interesse. Un thesaurus può essere anche poli-gerarchico, cioè un termine appartenente al livello ampio (BT) può anche apparire in uno o più livelli gerarchici sotto una o più classi o gruppi concettuali. I termini possono essere associati tra loro, ma non necessariamente connessi da una gerarchia. Ciò significa che due termini simili possono essere associati, anche se sono classificati sotto differenti BT. In questo caso sono definiti come RT, spesso utilizzati per aiutare le ricerche.

Gli standard ISO 25964 e 5964, rilasciati per la creazione di thesauri mono e multilinguistici, forniscono raccomandazioni su come implementare una risorsa, ma non danno indicazioni su come superare alcuni limiti evidenti nel campo della normalizzazione terminologica della documentazione archeologica; tra questi si segnala, ad esempio, la stretta relazione tra una lista di termini e la loro appartenenza, a volte esclusiva, a singole facies storico-culturali.

L'uso di una specifica terminologia implica la scelta di una precisa prospettiva di ricerca, nonché di un apparato lessicale connesso ad una distinta cultura. Nello stesso tempo, ad analoghe denominazioni culturali e/o terminologiche, possono associarsi ricostruzioni cronologiche con intervalli distanti, e infine, sistemi di classificazione della ceramica possono basarsi su parametri di tipo morfologico o funzionale. Varianti latine o greche e traduzioni in lingue moderne rendono ancora più confuso il quadro di riferimento.

L'uso di una nota e riconosciuta terminologia non appare sufficiente a rendere fruibile un archivio digitale. Alla attribuzione di un termine dovrebbe essere sempre associato un commento che illustri la scelta adottata per la definizione. Le ambiguità per la classificazione tipologica, in cui le variabili spazio-temporali appaiono condizionare la selezione dei termini adottati, si ritrovano anche nella identificazione e denominazione dei periodi culturali. Il rischio di generare confusioni e ambiguità utilizzando tradizionali strumenti terminologici è forte e sempre presente.

I thesauri esaminati in questo paragrafo sono, invece, la testimonianza di una traiettoria innovativa fondata sulle tecnologie del Semantic Web, per produrre fonti digitali standardizzate, integrabili e relazionabili.

La strada intrapresa dalle più importanti istituzioni nazionali e internazionali è quella della implementazione di un modello basato su una particolare forma di organizzazione della conoscenza, definita KOS, per rendere accessibili online le informazioni, i concetti, i termini e le loro relazioni. Gli strumenti a disposizione per una nuova strategia di condivisione dell'informazione derivano dal linguaggio RDF, dal modello SKOS e dai principi dei LOD che forniscono identificatori permanenti sotto forma di URL per i concetti documentati; i Web-services, implementati con interfacce e motori di ricerca semplici e accessibili, sono in grado di eseguire puntuali interrogazioni semantiche.

Negli ultimi vent'anni, ricercatori e operatori hanno affrontato molti dei problemi legati alla costruzione sistemi di organizzazione della conoscenza, fornendo l'accesso in rete a diversi vocabolari per descrivere la forma e il contenuto dei loro materiali. I dizionari geografici, come GeoNames (*infra* 99), sono diventati un'infrastruttura per il collegamento tra nomi di luoghi e coordinate geografiche. Il successo di queste risorse è il risultato di uno sforzo in termini di modellazione e rappre-



sentazione dei dati che consente di disambiguare termini e concetti e di metterli in relazione.

La tripla in precedenza illustrata (*supra* 75) potrebbe essere arricchita inserendo il nome della città dove ha sede l'Ateneo.

<https://orcid.org/0000-0002-5274-0786>  
*teaches\_at*  
[www.unior.it](http://www.unior.it)  
*is\_located\_at*  
[www.geonames.org/3172394/naples.html](http://www.geonames.org/3172394/naples.html)

L'affermazione strutturata consentirebbe ad un agente semantico automatico di inferire che il ricercatore identificato dal codice ORCID insegna a Napoli, città della Campania, in Italia.

Chiaramente le relazioni espresse in questo modo estendono il campo delle connessioni favorendo, ad esempio, il raggruppamento di tutti i ricercatori che insegnano a Napoli. Relazioni semantiche, gerarchiche e associative e uso di link permanenti sono alla base del linguaggio SKOS adoperato per la maggior parte dei thesauri.

Il modello dati SKOS è utilizzato per definire classi e proprietà indipendentemente dal dominio: ogni concetto è una risorsa RDF alla quale sono associati uno o più termini preferiti, valori alternativi e descrizioni o note. Le risorse possono essere organizzate a loro volta in relazioni gerarchiche, più ampie o ristrette, o non gerarchiche. La rapida espansione del Web di dati sta favorendo l'impiego di strumenti per l'organizzazione della conoscenza a partire dai thesauri che rappresentano risorse fondamentali per l'affermazione di un modello di scienza aperta; specifici registri disponibili online, come BARTOC, consentono di identificare il thesaurus più adatto alle proprie ricerche.

#### 2.4.4. FAIR Data

La condivisione dei risultati rende la ricerca più trasparente, promuove la collaborazione tra gli studiosi e riduce il rischio della obsolescenza dei formati. Inoltre, i dati aperti incoraggiano approcci interdisciplinari, ampliando il dibattito scientifico e, infine, soddisfano i requisiti degli enti finanziatori, aumentando la visibilità e le ricadute dell'investimento

pubblico. Ma come devono essere modellati i materiali scientifici per supportare un approccio *data-intensive*? Quali caratteristiche devono essere valorizzate, e in che modo, affinché i dati della ricerca siano accessibili in rete?

Fornire dati leggibili dalla macchina come livello principale per la *Knowledge Discovery* e per il corretto e sostenibile funzionamento di questi processi è una delle grandi sfide del Semantic Web e dell'OS. Per rispondere a questo interrogativo, un gruppo di ricercatori si è unito nel 2014 in una comunità, denominata Force11, con l'obiettivo di migliorare lo scambio e la condivisione della conoscenza fornendo linee per il riuso dei dati.

Nel 2016 il gruppo ha pubblicato un articolo che elenca i principi per la gestione e l'amministrazione dei dati scientifici (Wilkinson *et al.* 2016; 2019). La guida evidenzia che soltanto attraverso la definizione di un insieme minimo di pratiche e elementi concordati dalle proprie comunità di riferimento, i produttori e i consumatori di dati, sia automatici che umani, possono più facilmente scoprire, accedere, connettere e riutilizzare, con riferimenti e citazioni adeguate, una grande quantità di informazioni generate dalla conoscenza scientifica. Questi semplici principi dovrebbero supportare una variegata gamma di azioni consentendo di incrementare ed esplorare i dati suggerendo particolari scelte e implementazioni tecnologiche per la creazione, il consumo e la visualizzazione dei dati su Internet.

La premessa, su cui tali elementi si fondano, è che la ricerca sulla rete presupponga due distinti strati che devono essere chiaramente identificati e dichiarati: i dati, concepiti in modo semplice in quanto il risultato di un processo di acquisizione; e i metodi, spesso impliciti, con i quali i dati sono stati selezionati, elaborati, analizzati e, infine, interpretati. Il gruppo Force11 ha elencato quattro componenti che un moderno ambiente di pubblicazione dei materiali di ricerca dovrebbe possedere per supportare sia manualmente che in forma automatica il deposito, l'esplorazione, la condivisione e l'uso delle risorse digitali; i dati devono essere ricercabili, accessibili, interoperabili e riutilizzabili.

I principi FAIR identificano alcune specifiche aree di descrizione dell'informazione in rete che non richiedono, da parte del ricercatore, modifiche della formalizzazione o strutturazione del dato. Nello stesso

tempo questi principi dovrebbero ispirare la creazione di infrastrutture e specifiche tecnologie in grado di sfruttare le caratteristiche presenti negli archivi conformi ai requisiti FAIR.

Le quattro componenti principali sono state ulteriormente articolate in sub-requisiti per abbracciare un insieme di possibili scenari ed iniziative finalizzate ad indirizzare al meglio l'azione di un agente-semantic che individui e raccolga i dati necessari. L'idea sullo sfondo di un tale approccio è che una macchina dovrebbe lavorare come un agente-umano che in forma intuitiva comprende il significato semantico di una risorsa. Sia attraverso la lettura di un testo, di una espressione o di un termine, sia identificando una varietà di altri segnali visuali o audio, un utente, anche non particolarmente esperto, potrebbe facilmente cogliere il contesto nel quale è inserito un certo dato digitale. Questo procedimento non è immune dal rischio di errori nella interpretazione dell'oggetto, ma se le risorse sono ben inquadrare ed esplicitate il lavoro di un essere umano risulterà rapido ed efficiente.

La limitazione di un processo totalmente umano risiede nella scala e, quindi, nella velocità di una ricerca semantica eseguita su un numero pressoché infinito di potenziali fonti digitali. La sfida è, dunque, cercare di creare un ambiente nel quale agenti semantici possano riprodurre le forme istintive del ragionamento umano. Per raggiungere un tale obiettivo non è sufficiente programmare meccanismi automatici in grado di vagare nello spazio della rete alla ricerca di informazioni; occorrono dati, formati e protocolli, concettuali e fisici, che possano guidare e indirizzare l'attività automatica dell'agente-semantic nell'analisi approfondita degli archivi e del loro contenuto. È, dunque, necessario assistere qualsiasi forma automatica di estrazione della conoscenza fornendo dati *self-explaining* conservati in siti web accessibili.

Le informazioni consentono all'agente, in funzione della quantità di dettagli forniti, di avere la capacità, di fronte a una qualsiasi risorsa digitale, mai incontrata prima, di:

- identificare il tipo di oggetto, rispetto sia alla struttura che all'obiettivo;
- determinare se è utile nel contesto dell'attività richiesta interrogando i metadati e/o elementi di dati;

- determinare se è utilizzabile, in relazione al tipo di licenza;
- intraprendere azioni appropriate, più o meno allo stesso modo in cui farebbe un essere umano quando associa contenuti.

Se le informazioni non avessero tali prerequisiti un agente potrebbe determinare il tipo di dati di un oggetto digitale selezionato, ma non essere in grado di analizzarlo a causa della codifica in un formato non standard, oppure sarebbe incapace di elaborarne i contenuti mancando indicazioni relative alla licenza e al riuso dei dati. Purtroppo, una struttura formale che consenta ad una macchina di comprendere completamente il significato di una risorsa non è stata ancora progettata, tuttavia i principi FAIR possono aiutare gli agenti ad automatizzare molte delle attività di estrazione e aggregazione dei dati che caratterizzano i tradizionali compiti di un ricercatore.

Allo scopo di fornire una guida, il gruppo Force11 ha messo a punto una checklist di verifica di corrispondenza dell'archivio digitale ai principi FAIR.

In primo luogo, il processo di produzione di dati aperti deve rispettare il concetto della rintracciabilità. I dati e i metadati devono essere resi facilmente leggibili sia dalle persone sia dalle macchine poiché l'impiego di metadati codificati in un linguaggio macchina aiuta la ricerca automatica di set di dati e servizi. Le risorse devono essere:

- F1. Assegnate ad un identificativo globale unico e persistente; l'individuazione dei dati è una condizione necessaria per l'accesso ed il riuso degli oggetti digitali. La risorsa deve essere identificata in modo univoco ed immutabile con un identificatore persistente (PID) che continua a funzionare anche se la risorsa ha cambiato il suo indirizzo web;
- F2. Descritte con ricchi metadata (definiti ulteriormente al punto R1). L'utente deve scegliere uno schema di metadati esistente conforme agli standard internazionali per assicurare lo scambio dei dati;
- F3. Identificate chiaramente ed esplicitamente. L'autore delle risorse potrebbe essere identificato tramite il suo ORCID o risorsa VIAF o ISNI per consentire il collegamento, ad esempio, tra dataset e pubblicazioni;

- F4. Registrare o indicizzati in una risorsa accessibile in rete tramite un portale o un repository accreditato in grado di assicurare la disponibilità nel lungo periodo e una conservazione corretta e sicura.

Dopo che l'oggetto digitale è stato rintracciato, l'utente, automatico o umano, ha bisogno che i metadati assicurino l'accessibilità all'archivio. Pertanto, i metadati devono:

- A1. Esporre i protocolli di comunicazione necessari allo scambio dei contenuti;
  - A1.1. Adottare protocolli aperti, gratuiti e universalmente implementabili per garantire funzioni di harvesting che possono migliorare l'azione dei motori di ricerca;
  - A1.2. Accettare, se necessario, l'uso di protocolli per l'autenticazione e l'autorizzazione delle procedure;
- A2. Essere accessibili anche se i dati non sono più disponibili.

Lo step successivo riguarda come rendere interoperabili i dati per l'analisi e l'elaborazione degli oggetti digitali in altre applicazioni. Per garantire l'integrazione e l'armonizzazione, le risorse devono:

- I1. Usare un linguaggio di rappresentazione della conoscenza formale, accessibile, condiviso ed ampiamente applicabile;
- I2. Adottare un vocabolario conforme ai principi FAIR ed alla comunità di riferimento utilizzando termini e concetti noti e non ambigui;
- I3. Includere riferimenti ad altri (meta)dati. Se i dati citati hanno un PID le risorse e i dataset corrispondenti saranno più visibili e più facilmente rintracciabili.

Le risorse devono essere descritte in modo efficace per supportarne la riusabilità, la riproduzione e la combinazione in diversi contesti. I (meta)dati devono:

- R1. Essere riccamente descritti con una pluralità di accurati e rilevanti attributi;
  - R1.1. Essere rilasciati con una chiara ed accessibile licenza d'uso indicando le modalità dell'eventuale riutilizzo;

- R1.2. Essere associati a dettagli relativi alla provenienza digitale;
- R1.3. Soddisfare gli standard rilevanti per il dominio di riferimento.

I sub-principi forniscono all'utente chiarimenti nella comprensione dei principi generali che sono indipendenti dal dominio, modulari e astratti e quindi applicabili ad un range ampio di esigenze scientifiche. Le linee guida non forniscono alcuna informazione su come una infrastruttura debba essere costruita, quali protocolli selezionare o quale schema di metadati o identificatore. I dati che aderiscono ai principi fissati dal gruppo Force11 supportano gli agenti-semantici nei meccanismi di contestualizzazione degli oggetti digitali ed al loro riuso per nuove ricerche. Strumenti di *self-assessment*, disponibili online, aiutano il produttore di contenuti a verificare se il proprio dataset digitale incontri i requisiti minimi richiesti dai principi FAIR.

Nel 2016, all'interno della progettazione del programma europeo Horizon 2020 che ha promosso e sostenuto i principi dell'OS e dell'EO-SC (Mons *et al.* 2017) come approccio al processo scientifico basato sul lavoro cooperativo e sulla diffusione della conoscenza, i beneficiari di fondi comunitari hanno ricevuto l'obbligo di conferire i loro dati secondo le raccomandazioni FAIR; la Commissione Europea con questo programma ha fornito indicazioni preziose non soltanto per una gestione attiva e consapevole dei dati, ma anche per la definizione di politiche per la conservazione degli archivi digitali nel lungo periodo.

Considerata l'intersezione tra LOD e FAIR e il valore incrementale che i vari livelli della gerarchia proposta riflettono, i gradi delle due classificazioni sono stati combinati con l'obiettivo di raggiungere un più alto stadio di potenziale integrazione dei dati, rafforzando, soprattutto, la ricerca e il riuso degli archivi aperti (Hasnain, Rebholz-Schuhmann 2018).

Vedremo più avanti come il mondo dell'archeologia, all'interno di un più ampio passaggio verso forme più aperte della conoscenza, stia cominciando a adottare linee guida per rendere i dati compatibili con gli scenari definiti dall'OS.



## Capitolo 3

### OPEN ARCHAEOLOGY

#### 3.1. Introduzione

Gli scenari delineati dalla Scienza Aperta hanno prodotto una accelerazione nella riflessione disciplinare consentendo un salto metodologico, connesso soprattutto alla possibilità di verificare direttamente, come nel più tradizionale processo scientifico, la relazione tra fonti archeologiche, applicazioni informatiche e risultati finali (Marvick 2017).

La ricerca archeologica deve fondarsi su dati condivisi e, quindi, trasparenti per quanto riguarda il formato, l'accessibilità e il contenuto, conservando traccia di tutte le procedure impiegate per la digitalizzazione e il successivo trattamento informatico. Un modello tridimensionale, ad esempio, può essere correttamente interpretato, utilizzato e riutilizzato se si posseggono tutte le informazioni sull'oggetto fisico originale e sulle fasi della sua trasformazione in record digitale. Il produttore dell'oggetto 3D deve associare al documento grafico una scheda che registri gli strumenti adoperati, i processi messi a punto e tutte le condizioni, anche ambientali, verificatesi al momento della digitalizzazione e della post-elaborazione. Queste informazioni contribuiscono alla lettura del modello e rendono comprensibile qualsiasi ipotesi ricostruttiva. La risorsa 3D, per quanto accurata e complessa, diventa in tal modo un documento scientifico, contribuendo alla conoscenza archeologica e perdendo la sua connotazione di immagine suggestiva.

La Scienza Aperta agisce nella direzione di rimuovere ogni forma di proprietà e segretezza a favore di una gestione trasparente ed inclusiva dell'informazione scientifica; essa non è un fine, ma un mezzo fondato su accesso, dati e metodi aperti.

In una epoca totalmente digitale e nella quale si opera in ambienti virtuali, la richiesta della pubblicazione di dati aperti risponde a numerosi interrogativi. Mentre gli archeologi producono e gestiscono abitualmente grandi quantità di informazioni digitali complesse e altamente strutturate,



la diffusione e la comunicazione dei risultati delle ricerche rimangono ancorate alla tradizionale pubblicazione a stampa caratterizzata, visti i costi, da una logica di selettività nella presentazione dei materiali.

Numerose case editrici impongono ai ricercatori di non condividere i dati primari delle loro ricerche, precludendo così il riutilizzo e il riesame degli archivi. Se un tale comportamento, legato a finalità commerciali, poteva essere giustificato nell'era predigitale dai costi di produzione, tutto ciò appare, oggi, anacronistico e culturalmente controproducente in un mondo digitale con infinite connessioni globali. Poiché, di frequente, le pubblicazioni a stampa hanno un costo di vendita notevole, una parte consistente della conoscenza risulta negata alla maggioranza dei ricercatori.

Questo orientamento, ancora largamente prevalente in archeologia, evidenzia alcune significative criticità. La prima è di natura epistemologica poiché non è possibile parlare di ricerca senza poter accedere ai dati primari. La seconda è etica in quanto la maggior parte delle indagini sono svolte con fondi pubblici e, quindi, è legittimo avanzare un dubbio sulla natura aperta (pubblica) o chiusa (privata) dei dati raccolti. La terza è operativa dal momento che le pubblicazioni tradizionali paradossalmente richiedono tempo e impegno per convertire in formato cartaceo i dati acquisiti originariamente in digitale.

L'adesione ad una Scienza Aperta si delinea senza dubbio come un passaggio obbligato anche per il mondo della ricerca archeologica, ma questo cambiamento richiede un radicale mutamento nelle pratiche di acquisizione, raccolta, elaborazione e conservazione dei dati che consideri anche la naturale tendenza delle tecnologie a diventare obsolete in tempi relativamente brevi. Non è, dunque, sufficiente sviluppare competenze per digitalizzare i dati o per produrre archivi di risorse digital-born. È necessario, al contrario, partire dalla profonda consapevolezza che ogni trasformazione produce un salto qualitativo che, a prima vista, può indurre nel ricercatore il disorientamento tipico di chi deve abbandonare un porto sicuro per affrontare il mare aperto.

Nel dicembre 2012, un intero numero della rivista *World Archaeology* (44, 4) è stato dedicato agli scenari dell'archeologia aperta. Nell'editoriale che apriva questo volume tematico, M. Lake (2012) lamentava la scarsa attenzione che il fenomeno della Scienza Aperta aveva ricevuto in ambito archeologico, nonostante l'uscita del bel volume sull'archeologia

2.0 (Kansa *et al.* 2011) nel quale si cominciava a delineare il contesto di una nuova scienza digitale, aperta e collaborativa. Lake osservava come il mondo dell'archeologia restasse totalmente estraneo a questa rivoluzione culturale quando non dichiaratamente contrario. Una tale trasformazione non poteva rimanere circoscritta all'ambito di pochi adepti, peraltro già interessati alle implicazioni digitali in archeologia. Occorreva, a detta di Lake, attirare l'attenzione di quanti usano strumenti computazionali ed Internet inconsapevoli delle profonde ripercussioni che una tale tecnologia pervasiva produce sulla pratica archeologica, principalmente nel campo della creazione di contenuti scientifici.

L'impiego delle tecnologie digitali racchiude concetti come dematerializzazione, delocalizzazione, accesso rapido e non lineare, manipolazione, cambiamento nella produzione, forma, ricezione e uso dei media. L'informatica consente un effettivo risparmio di tempo principalmente nella organizzazione e gestione del dato testuale e grafico, indipendentemente dai software e dagli obiettivi della ricerca. Tuttavia, il tema centrale del ragionamento resta ancora quello della formazione del record archeologico e, quindi, della sua descrizione e documentazione.

Se per molti anni il dato archeologico è stato considerato il risultato di una attività distruttiva, la registrazione digitale ha spinto a ritenere l'informazione come il prodotto di un processo creativo di acquisizione di dati digital-born di alta qualità (Roosvelt *et al.* 2015). Negli ultimi anni questa visione efficientista del computer è stata negata da alcuni studiosi che, prediligendo una *slow archaeology*, rifiutano l'idea che l'informatica serva a velocizzare o semplificare l'acquisizione e la gestione dei dati (Caraher 2019).

Il riesame della funzione scientifica del dato ha spinto ad una sostanziale modifica degli approcci teorici. Nel corso degli ultimi 20 anni siamo passati da una archeologia dei modelli, fondata sulla rigorosa definizione di strutture preordinate di raccolta ed analisi su uno sfondo teorico caratterizzato da leggi generali, ad una archeologia dei dati, diretta ad una maggiore comprensione dei meccanismi di formalizzazione e trattamento del record archeologico e soprattutto delle forme di scrittura, riscrittura e gestione della conoscenza (D'Andrea 2016).

Gli archeologi devono abbandonare forme descrittive personali o troppo legate al dominio o sottodominio di appartenenza ricorrendo a thesauri e/o standard il cui scopo è rendere l'informazione realmente

accessibile e condivisibile. Lo sviluppo del Web Semantico e dei LOD sta spostando l'attenzione degli archeologi dalla semplice registrazione delle informazioni verso una più approfondita analisi del linguaggio scientifico adoperato e, più in generale, della conoscenza implicita ed esplicita presente nei dati. Lo sfondo di questo ripensamento disciplinare, ancora agli albori, è lo scenario dei BD, un ambito di ricerca caratterizzato da grandi volumi di dati accessibili liberamente ed interrogabili.

Questo approccio del tutto innovativo, ma ancora poco diffuso nei settori umanistici, sta traghettando l'archeologia verso orizzonti totalmente integrati nel Web ed in tutte le sue innumerevoli prospettive di ricerca, di gestione e di riuso del dato.

Siti come tDAR, ADS ed Open Context (*infra* 125) offrono, a costi contenuti, la possibilità di preservare e condividere in forma permanente, archivi digitali che comprendono informazioni grafiche e fotografiche, tabelle e testi, mettendo a disposizione di una ampia comunità, scientifica e non, dati di ricerche concluse o in corso. Anche la diffusione degli Open Data nelle Pubbliche Amministrazioni contribuisce a modificare il rapporto tra istituzioni pubbliche e cittadini, favorendo lo sviluppo di nuovi modelli di gestione più trasparenti e una diversa cultura del territorio nella quale anche l'archeologia trova una sua dimensione in una ottica di partecipazione e collaborazione.

I vantaggi di una tale rivoluzione sono del tutto evidenti: incremento nell'efficienza e nella diffusione della scienza riducendo, allo stesso tempo, i costi nel collezionare, creare, trasferire e condividere dati e materiale scientifico; il riutilizzo delle informazioni moltiplica, quindi, le opportunità di studio e analisi.

L'accesso libero ai dati della ricerca produce un altro significativo beneficio, poiché il ricercatore può avvalersi di un numero elevato di valutatori che verificano i risultati del suo lavoro.

Da una specifica prospettiva di ricerca, l'OSIG ha identificato nel 2017 tre aspetti determinanti per l'affermazione della Scienza Aperta: una analisi di dati, la definizione delle forme di accesso e l'individuazione di metodi digitali. Occorre, quindi, intervenire incoraggiando l'implementazione di repository per la condivisione di preprint, pubblicazioni e archivi comprendenti dati grezzi associati a metadati standard. Parallelamente si devono riconoscere i vari diritti d'autore connessi alla digitaliz-

zazione delle risorse che si vogliono rendere accessibili (D'Andrea *et al.* 2022). Infine, è necessaria una descrizione puntuale dei procedimenti di acquisizione, elaborazione ed analisi affinché i risultati della ricerca siano riproducibili. I ricercatori sono, perciò, stimolati a adoperare formati aperti, software non proprietari e a rendere gli script o i programmi, sviluppati per specifiche indagini, disponibili in archivi accessibili.

La crescente importanza della comunicazione via Internet, associata ad un incremento esponenziale dei dati digital-born per l'elaborazione della documentazione archeologica, ha spinto la comunità scientifica verso una maggiore enfasi sui concetti quali *Research Data Management* il cui fine è comprendere il potenziale che l'innovazione offre per migliorare la ricerca e la gestione (Kansa *et al.* 2020). Uno dei caposaldi di questa filosofia è senza dubbio il principio della condivisione che semplifica la riproducibilità delle analisi principalmente in quei settori dove il ricorso alle tecniche quantitative è determinante.

La consapevolezza dell'esigenza di trasferire dati grezzi, uniti alle elaborazioni e interpretazioni, produce nei ricercatori un duplice effetto: spingere a rendere fruibili gli archivi e, dall'altro, attendersi che ciascun archeologo contribuisca ad arricchire il corpus di informazioni accessibili liberamente. Il rischio ben presente in un tale approccio è che l'adesione dei ricercatori si realizzi senza una adeguata conoscenza delle strutture formali che presiedono al Web e alle metodologie di strutturazione dei dati, producendo al contrario una Babele di risorse digitali isolate e non integrabili.

Di recente il gruppo che coordina le attività del repository Open Context ha fornito alcune indicazioni per una procedura basata su passaggi in sequenza che hanno l'obiettivo di incoraggiare i ricercatori a sviluppare metodi migliori per modellare i loro dati e renderli così pienamente riutilizzabili. Le linee guida hanno il fine di trovare il più ampio consenso tra gli utenti per pratiche che includano anche lo sviluppo di thesauri e vocabolari comuni.

I dati dovrebbero essere strutturati in accordo alle seguenti linee generali (Kansa *et al.* 2020):

- garantire l'accesso alle fonti primarie, cioè tabelle o database, non come immagini incorporate nelle pubblicazioni, ma come file scaricabili;

- includere informazioni sul contesto di provenienza, es. coordinate temporali e spaziali del dataset. I siti e i contesti dovrebbero essere identificati esplicitamente per semplificare riferimenti incrociati;
- conservare per ogni progetto online traccia degli obiettivi del programma di ricerca, delle analisi eseguite e dei destinatari delle indagini. Un *data management plan* dovrebbe includere tutte le fasi della ricerca, dallo scavo alla pubblicazione finale, comprendendo anche le procedure di digitalizzazione adoperate;
- fornire il maggior numero di informazioni relative alle modalità di indagine del contesto/sito esaminato. Alcune ricerche dimostrano che differenti approcci applicati allo stesso dataset possono generare esiti diversi (SAA 2018). Una corretta *digital provenance* deve, pertanto, elencare dettagliate informazioni su quali metodi e tecniche siano state impiegate per estrarre i dati o un campione parziale di essi, allo scopo di ridurre ogni rischio di erranea valutazione degli archivi riutilizzati;
- adoperare identificatori per promuovere l'interoperabilità. L'attività del ricercatore non è limitata al semplice caricamento dei suoi dati sul Web. Egli deve utilizzare modalità descrittive non ambigue riferite, laddove possibile, a risorse già presenti in rete accrescendo il potenziale dei LOD che non costituiscono soltanto un riferimento all'interno del Web, ma rappresentano un identificativo permanente ad oggetti, cose e concetti;
- esplicitare il tipo di copyright per il riuso dei dati. La proprietà intellettuale può essere garantita condividendo i dati con licenze aperte, es. Creative Commons, che assicurino il riconoscimento dell'opera del ricercatore;
- integrare, fin dall'inizio, una condotta che assicuri la trasparenza dei dati e dell'intero processo di ricerca rendendo così più affidabili le analisi.

Adeguate i dati a queste linee guida contribuisce a migliorare la riproducibilità della ricerca archeologica e facilita l'integrazione e l'aggregazione dei dati. La pubblicazione online degli archivi al livello più granulare possibile risulterà, così, più importante della stessa edizione scienti-

fica favorendo, allo stesso tempo, una più ampia comprensione dei dataset condivisi.

### 3.2. Thesauri e strumenti terminologici

La multiculturalità, il multilinguismo e le regole, diverse da paese a paese, che presiedono alla gestione del patrimonio archeologico, hanno finora impedito una efficace condivisione dei dati. Nonostante l'azione di coordinamento che ciascuna nazione o istituzione culturale nazionale ha promosso attraverso la definizione di best-practices, mancano linee guida comuni per la costruzione di sistemi in rete accessibili e usabili. Anche l'implementazione di archivi online con software e procedure differenti ha contribuito ad accentuare la diversità tra le soluzioni computazionali esistenti, al punto che risulta arduo garantire l'interoperabilità sintattica e semantica tra collezioni digitali eterogenee.

L'adozione di standard, riconosciuti a livello internazionale, è apparsa come una delle possibili tecniche per la condivisione degli archivi digitali, sia per assicurare l'omogeneità della rappresentazione dei dati, che per consentire un riuso consapevole delle fonti digitali. In tale senso la standardizzazione ha rappresentato il primo decisivo passo, tecnologico e culturale, verso il superamento di ogni ostacolo ad una reale integrazione di distinti dataset. Tuttavia, la semplice rappresentazione del dato in base a specifici standard non ha consentito di superare alcune problematiche relative all'integrazione dei dati; in tal senso una precisa causa va riconosciuta all'ambiguità semantica delle fonti.

Ontologie, fondazionali e di dominio, thesauri, liste terminologiche e vocabolari controllati sono stati, quindi, rilasciati per disambiguare il contenuto degli archivi online. Accolte con perplessità dagli addetti ai lavori, queste tipologie di rappresentazione formale della conoscenza sono diventate estremamente utili con lo sviluppo del paradigma dei BD, parallelamente al crescere della consapevolezza della necessità di dover condividere i dati delle proprie ricerche.

Negli ultimi anni è emersa una maggiore attenzione dei ricercatori verso la creazione di thesauri, soprattutto, per due ragioni fondamentali: assicurare la coerenza formale dei dati e controllare l'uso di vocaboli appropriati per le descrizioni degli oggetti onde evitare i rischi di confusione derivanti dall'ambiguità e dalla incertezza terminologica. Quest'ul-

timo aspetto offre un ulteriore vantaggio poiché consente di migliorare la fase di interrogazione fornendo al ricercatore una lista di possibili valori alternativi da ricercare.

Per ridurre la difficoltà di operare specifiche ricerche su grandi archivi di dati, alcuni siti mettono a disposizione degli utenti una serie di filtri preimpostati che compensano precisione e genericità nella costruzione di una interrogazione complessa; i filtri di ricerca sfruttano la descrizione normalizzata dei dati fondata su vocabolari controllati o thesauri. La procedura assicura che i risultati delle interrogazioni siano abbastanza coerenti con le richieste formulate dal ricercatore che non deve necessariamente conoscere i termini adoperati per la codifica della risorsa digitale. I filtri eliminano i problemi legati alle varianti ortografiche, mentre, essendo i sinonimi in genere indicizzati, la ricerca risulta rapida e semplificata poiché l'utente non deve necessariamente testare tutte le varianti linguistiche di un dato termine.

L'incremento dei dati in rete, connesso all'esigenza di condividere ed integrare i dati, ha spinto verso un'ampia produzione di thesauri creati da istituzioni locali o internazionali il cui obiettivo è rendere la procedura di normalizzazione più semplice, priva di errori e facilmente indicizzabile. Liste terminologiche possono essere strutturate in thesauri multidisciplinari oppure in liste circoscritte ad uno specifico dominio.

Il thesaurus multilingua dell'UNESCO appartiene alla categoria dei vocabolari ed è composto da una lista controllata e strutturata di concetti utilizzati per l'analisi e il reperimento di documenti e pubblicazioni in differenti settori (istruzione, cultura, scienze naturali, scienze sociali e umane, comunicazione). Il vocabolario, pubblicato in inglese per la prima volta nel 1977, è continuamente arricchito e aggiornato; attualmente è disponibile in 5 lingue. I concetti riflettono l'evoluzione dei programmi e delle attività dell'UNESCO e sono raggruppati in sette aree tematiche, suddivise in micro-thesauri. L'archeologia è inserita nell'area Storia e nella macroarea Cultura. I termini relativi alla archeologia sono pochi e in gran parte riferibili a concetti generici; la voce contiene sei sotto-voce. Il thesaurus, conforme allo standard ISO 25964, è stato pubblicato come dataset SKOS strutturato secondo i principi dei LOD ed è scaricabile o consultabile in rete in vari formati. Il database include circa 4.500 concetti. Il thesaurus è ulteriormente articolato in 88 micro-thesauri e

95 collezioni. L'UNESCO ha reso disponibile la risorsa con la licenza CC-BY-SA 3.0 IGO che Creative Commons ha recentemente rilasciato per le Organizzazioni Internazionali.

Ad una categoria più vicina al mondo della terminologia degli oggetti archeologici, appartiene il thesaurus FISH sviluppato nel 2020 dal *Collections Trust* dell'*Historic England*. Il vocabolario include termini in inglese relativi a tracce antropiche o naturali recuperate durante il lavoro sul campo. La risorsa è disponibile online e, come il thesaurus dell'UNESCO, è strutturato come dataset SKOS secondo i principi dei LOD ed è visualizzabile in differenti formati. L'*Historic England*, l'*Historic Environment Scotland* e la *Royal Commission on Ancient and Historical Monuments of Wales* hanno prodotto molteplici thesauri conformi allo standard SKOS e pubblicati online come risorse del Semantic Web sulla piattaforma *heritagedata.org*. Il sito mette a disposizione anche una suite di controlli, accessibili con licenza CC-BY, che sono configurabili ed incorporabili direttamente nelle applicazioni basate su browser per fornire l'accesso ai vocabolari.

Tra i thesauri multilingua più adoperati dalla comunità degli archeologi si devono, senza dubbio, includere le risorse sviluppate dal Getty Research Institute: l'AAT, il TGN, l'ULAN, il CONA e l'IA. I vocabolari contengono una terminologia strutturata per l'arte, l'architettura, le arti decorative, i materiali d'archivio e i materiali bibliografici fornendo informazioni a catalogatori e ricercatori. Sono conformi agli standard internazionali e sono strutturati con riferimenti ad altre risorse in caso di sovrapposizioni semantiche. La grande varietà delle voci registrate, visualizzabili in vari formati e secondo le regole dei LOD, rende tutti i vocabolari una fonte straordinaria per la classificazione dei singoli termini e come base di conoscenza con molteplici interconnessioni all'interno e tra i vocabolari stessi. I thesauri del Getty sono costantemente aggiornati grazie alla collaborazione fornita dal Getty Program o da altre istituzioni; anche il Getty Conservation Institute e il J. Paul Getty Museum sono impegnati nella costante revisione delle liste e dei vocabolari. Tra i thesauri del Getty si segnala l'AAT che contiene oltre 70.000 concetti e circa 479.000 termini. Le risorse, rilasciate con licenza ODC-By 1.0, sono multilingue, multiculturali e crescono grazie ai contributi delle istituzioni e delle comunità di utenti esperti. L'AAT, nato agli inizi



degli anni '70 per rispondere alle esigenze delle biblioteche, si è poi esteso alla catalogazione degli oggetti museali, alle arti visive, all'architettura e all'archeologia. Sono esclusi i nomi propri e i soggetti iconografici per il dominio dell'arte e dell'architettura, inclusi negli altri vocabolari Getty: nomi di persone (ULAN), nomi geografici (TGN), titoli di opere d'arte (CONA), soggetti iconografici (IA). Sono esclusi dall'AAT anche i termini composti che combinano parole appartenenti a gerarchie distinte; ad esempio, cattedrale barocca è un termine composto che combina un tipo di opera e uno stile.

Tra gli strumenti terminologici adoperati in Italia si devono, certamente, includere i thesauri dell'ICCD disponibili sulla piattaforma github e sul sito del Catalogo dei dati aperti del Ministero della Cultura; tra i thesauri pubblicati si segnalano il dizionario dei reperti, della tecnica e della materia dei reperti archeologici anch'essi consultabili e scaricabili in vari formati e impostati secondo le norme del linguaggio SKOS. Il vocabolario dei manufatti mobili, aggiornato nel 2021, contiene 17.000 definizioni estratte dalla banca dati SIGEC a cui si sono aggiunti altri termini identificati nel corso di successivi progetti. Gli strumenti terminologici sono rilasciati con licenza aperta Creative Commons; nel quadro del progetto europeo ARIADNE, il dizionario relativo agli oggetti è stato rielaborato ed allineato all'AAT del Getty Museum (Felicetti *et al.* 2015).

L'esigenza di normalizzare la descrizione dei contenuti delle banche dati ha spinto i ricercatori ad adottare dizionari terminologici molto precisi o abbastanza ampi da consentire di rappresentare in modo omogeneo la totalità degli oggetti archeologici indipendentemente dal periodo di produzione o uso e dalla collocazione geografica o culturale. L'importanza assunta dai thesauri è, chiaramente, cresciuta in parallelo con la diffusione delle tecnologie semantiche che hanno messo a disposizione degli studiosi strutture formali molto flessibili. Linguaggi come SKOS e la modalità LOD, come abbiamo visto, offrono schemi di rappresentazione e buone pratiche per progettare, pubblicare e condividere in modo semplice i propri dati in rete.

L'incremento delle connessioni tra i dataset ha incoraggiato i ricercatori verso la creazione di allineamenti e *mapping* tra i vari dizionari con l'obiettivo di aumentare l'interoperabilità degli archivi digitali. La spe-

cializzazione delle informazioni e degli apporti scientifici provenienti da altri domini scientifici, connessi con le ricerche archeologiche sul campo, ha favorito anche la progettazione di nuovi strumenti terminologici. Nello stesso tempo, per superare la frammentazione dei vari vocabolari implementati, sono sorte iniziative per la progettazione di thesauri globali per le discipline umanistiche.

Il BBT, sviluppato dal gruppo di lavoro del *Thesaurus Maintenance* nell'ambito della infrastruttura europea DARIAH è una delle più interessanti iniziative in questa direzione. La versione più recente del BBT è stata rilasciata a febbraio 2022 e può essere scaricata in word, PDF o RDF. Nella introduzione si chiarisce che il modello proposto dovrebbe rappresentare l'ossatura portante di un thesaurus fondazionale al quale tutti i differenti vocabolari potrebbero facilmente allinearsi. Il lavoro di sviluppo si è concentrato soprattutto sulla identificazione di concetti *top-level* in grado di rappresentare una base comune per risorse terminologiche interdisciplinari ed integrabili ad un livello formale astratto più elevato. Uno dei principali vantaggi di un sistema aperto, così concepito, è quello di poter essere facilmente integrato e arricchito garantendo la compatibilità tra differenti thesauri e favorendo forme più ampie di collaborazione e scambio di dati.

Il thesaurus iDAI.world raccoglie e unifica tutti i vocabolari delle biblioteche e dei progetti di ricerca dell'Istituto Archeologico Tedesco secondo le linee guida della *backbone* di DARIAH. Sviluppato tra il 2015 e il 2019, il thesaurus viene continuamente integrato e aggiornato. I singoli concetti sono tradotti in arabo, cinese, inglese, farsi, francese, italiano, polacco, spagnolo e comprendono tutta l'antichità (dal paleolitico alla tarda antichità/epoca bizantina) e le aree di intervento dell'Istituto Archeologico Tedesco (America centrale e meridionale, Europa, Mediterraneo, Africa, Asia). Il thesaurus contiene: l'iDAI.gazetteer che collega i toponimi con le coordinate e funge da archivio di autorità per tutte le informazioni geografiche e i sistemi informativi dell'Istituto archeologico tedesco; l'iDAI.chronontology che combina termini temporali con date e consente un ulteriore collegamento a qualsiasi altro oggetto, luogo o risorsa web all'interno e all'esterno di iDAI.world; e, infine, l'iDAI.shapes che costituisce una sorta di meta progetto offrendo un'infrastruttura per database orientati alla forma.

Il dizionario Pactols, strutturato in linguaggio SKOS e progettato in base alle norme ISO 25964, è ispirato al modello concettuale del BBT di DARIAH. Disponibile in sette lingue, il vocabolario può essere scaricato con licenza ODbL. Il thesaurus, progettato dal consorzio francese MASA, è composto dai dizionari *Subjects* e *Places* ed è allineato a Wikidata, Geonames, IdREF ed a AAT. Pactols comprende 60.000 termini o concetti dedicati all'archeologia e alle scienze dell'antichità.

Un elenco parziale di thesauri per le Digital Humanities e i beni culturali può essere consultato sulla pagina dell'ACDH-CH nel sito di DARIAH. Il centro offre un servizio di *hosting* e visualizzazione di dizionari pubblicati dal consorzio CLARIAH, partner dell'infrastruttura europea DARIAH, e dal progetto PARTHENOS.

Tra i vocabolari ospitati si segnala anche Iconclass, un sistema di classificazione progettato per la descrizione e la ricerca di soggetti raffigurati in opere d'arte, illustrazioni di libri, riproduzioni, fotografie, etc. Iconclass è uno dei più vecchi thesauri la cui storia risale agli anni '40 del secolo scorso. Nella sua forma completa è stato stampato in 17 volumi tra il 1973 e il 1985. Incluso negli anni '80 nel database tedesco dell'HIDAMIDAS, è stato pubblicato online nel 1992 presso il dipartimento di *Computer & Letteren* dell'Università di Utrecht. Il thesaurus si adattò presto alle esigenze dei primi sistemi digitali sviluppati a partire dagli anni Settanta finalizzati alla costruzione di banche dati. L'importanza di un così ricco e strutturato strumento terminologico è ben evidenziata da Brandhorst (1993) il quale ha ricordato come la digitalizzazione di Iconclass abbia rappresentato un caso di studio di successo nella conversione da un sistema cartaceo ad uno digitale in grado di standardizzare il contenuto informativo delle immagini, evidenziando connessioni e relazioni tra concetti. Dal 2015 è disponibile in SKOS e scaricabile in differenti formati. Iconclass è strutturato in 10 categorie principali e 450 sottocategorie definite con un sistema di codifica alfanumerica per classificare circa 28.000 definizioni e 14.000 parole chiave. Numerosi database online, che sfruttano il sistema di concetti e gerarchie di Iconclass, sono diventati un banco di prova per gli algoritmi dell'AI e del *machine learning* nel campo dell'analisi tematica delle immagini. Un lavoro è stato sviluppato su oltre 87.000 immagini selezionate dall'archivio Arkyes e indicizzate attraverso le definizioni e i concetti di Iconclass.

Altri vocabolari, altrettanto importanti per la normalizzazione delle informazioni archeologiche e per ridurre i rischi di confusioni terminologiche, hanno una caratterizzazione geografica e sono utilizzati per disambiguare località e siti moderni e/o antichi.

Il TGN, edito dal Getty Museum, è un vocabolario poli-gerarchico e multilingua, che comprende nomi, descrizioni e altri metadati relativi a città, siti archeologici e caratteristiche fisiche attuali (montagne, fiumi, etc.) e storiche, importanti per la ricerca sull'arte, l'architettura e la cultura materiale. Il TGN può essere collegato ad un archivio GIS, a mappe o altre risorse geografiche. Il TGN, rilasciato con licenza ODC-By, cresce e si modifica grazie ai contributi della comunità di utenti e al lavoro editoriale del *Getty Vocabulary Program*. Il thesaurus è conforme agli standard ISO e NISO e contiene relazioni gerarchiche, di equivalenza e associative. I nomi di un luogo possono includere versioni dialettali locali, termini in inglese e in molte altre lingue, nomi storici, nomi in ordine naturale e in ordine inverso; uno dei nomi è contrassegnato come preferito secondo lo schema SKOS. Le voci contengono coordinate geografiche approssimative in WGS84 espresse in gradi e minuti o in gradi decimali; il database di TGN contiene oltre 3.000.000 definizioni di luoghi e 5.300.000 voci relative a nomi di città moderne o storiche. TGN si focalizza sul mondo storico e sui contesti storici per facilitare la catalogazione e la geolocalizzazione di manufatti e altre strutture antropiche e/o naturali; il thesaurus registra sia luoghi esistenti che perduti. TGN è collegato ad altri thesauri come NGA/NIMA, USGS-GNIS e, soprattutto, Geonames.

Mentre TGN è stato progettato e implementato considerando soprattutto i luoghi storici e con una interfaccia per ricerche testuali, Geonames è un database geografico, distribuito con licenza Creative Commons che contiene oltre 27 milioni di nomi geografici e oltre 12 milioni di caratteristiche uniche, di cui 4,8 milioni di luoghi popolati e 15 milioni di nomi alternativi. Le voci sono classificate in nove classi ripartite in 645 categorie. GeoNames integra dati geografici come i nomi dei luoghi in varie lingue, l'altitudine, la popolazione e altri dati provenienti da varie fonti. Tutte le coordinate sono espresse in WGS84. Il motore di ricerca è implementato sia per ricerche di tipo testuale che tramite coordinate. Il modulo per le interrogazioni avanzate consente di specificare il termine da ricercare circoscrivendo il continente o la caratteristica fisica

o amministrativa del luogo; attivando la funzione *fuzzy search* il motore può identificare anche voci corrispondenti a termini scritti in modo errato. GeoNames dispone di una procedura di modifica molto semplice che consente agli utenti registrati di variare le informazioni, spostare un punto su Google Maps e aggiungere nuovi nomi di luoghi al database.

Ai thesauri geografici che hanno una finalità meno generalista come GeoNames e più storica come TGN, si deve aggiungere Pleiades. Il sito offre agli studiosi la possibilità di utilizzare, creare e condividere informazioni geografico-storiche sul mondo antico. Attualmente, Pleiades copre il mondo greco e romano e si sta espandendo alla geografia del Vicino Oriente antico, bizantina, celtica e altomedievale. Tutti i contenuti pubblicati sono accessibili con licenza aperta. Il sito contiene poco meno di 40.000 luoghi, oltre 36.000 nomi e circa 43.000 località. Piuttosto che seguire un ordinamento alfabetico, Pleiades predilige una organizzazione delle risorse in luoghi, nomi, località e connessioni. I luoghi sono il principale costrutto organizzativo del dizionario. Il termine luogo va inteso come concetto applicato a qualsiasi area dove sono attestate manifestazioni umane, materiali o immateriali, in un dato contesto geografico reale. Un insediamento presente in un testo antico è un luogo, indipendentemente dalla sua localizzazione; un sito archeologico scavato è un luogo, ma anche una città moderna che si sovrappone ad un antico insediamento è un luogo. Qualsiasi elemento spaziale collegato al passato ed esaminato da uno studioso antico o moderno è un luogo. Le aree sono, quindi, identificate in una accezione ampia e possono rappresentare: luoghi di attività umana intensa, come insediamenti e santuari; caratteristiche ambientali (geologiche e fisiografiche) conosciute nell'antichità; costrutti politici, sociali o culturali come province e distretti; singole strutture o estensioni spaziali. I luoghi sono entità astratte che non sempre hanno una localizzazione su una mappa; pertanto, possono non avere attributi spaziali o temporali. Un luogo può esistere solo come nome in una fonte antica, senza alcuna attestazione materiale, mentre un sito archeologico noto può rappresentare un luogo privo di un nome antico. Le coordinate geografiche e i nomi antichi e moderni dei luoghi sono classificati con le entità concettuali delle località e dei nomi. Una località identifica una singola area di interesse e può essere associata a un luogo in uno specifico intervallo di date. Un luogo a sua volta può contenere

più località. Il collegamento tra luogo e località dipende dalle informazioni disponibili; per esempio, un luogo attestato dalle fonti antiche può risultare privo di evidenze materiali, sebbene gli studiosi siano stati in grado di individuare approssimativamente l'area di una eventuale presenza. I nomi sono collegati alla loro identificazione scritta e non alla posizione fisica nel paesaggio; appartengono ad un periodo o a più periodi. Come per le località, un singolo luogo può avere più nomi, ma un singolo nome può essere associato a un solo luogo. La regola vale anche se una stessa sequenza di caratteri è attestata come nome di un altro luogo; Pleiades tratta questi nomi identici come entità separate. Le connessioni consentono, invece, di esprimere e documentare le relazioni tra luoghi, gerarchie geografiche, reti e collegamenti. Analogamente ai nomi e alle località, anche le connessioni appartengono ad uno specifico periodo cronologico. Le connessioni partitive, di flusso direzionale e di prossimità, sono preferite alle geometrie di localizzazione (ad esempio, la *bounding box*) per identificare luoghi non localizzabili, aree con estensioni incerte e confini indeterminati. Le connessioni si estendono ai collegamenti temporali, politici ed economici e non soltanto geografici.

AAT, TGN e Pleiades includono anche voci che descrivono i differenti periodi storici dell'antichità. Più specifico per la cronologia è PeriodO, un repertorio di pubblico dominio che contiene definizioni relative a periodi storici, storico-artistici e archeologici. PeriodO semplifica il collegamento tra dataset che definiscono in modo differente gli stessi intervalli cronologici supportando l'analisi soprattutto nelle aree di sovrapposizione tra periodi storici differenti. L'uso del termine età del bronzo, ad esempio, non viene circoscritto ad una singola area, ma si estende ad un numero ampio di luoghi antichi. PeriodO documenta le definizioni dei periodi storici che hanno un nome, limiti temporali, un'associazione implicita o esplicita con una regione geografica e pubblicata in una fonte citabile. Le designazioni di periodo provenienti da uno stesso testo o dataset sono raggruppate in collezioni. Attualmente PeriodO identifica 1.800 definizioni temporali apparse in fonti pubblicate o citate da archeologi. Le descrizioni sono state raccolte da circa 70 fonti: circa 1.000 provengono da progetti che curano e pubblicano dati archeologici sul Web, come il database *Fasti Online* degli scavi archeologici e il *Levantine Ceramics Project*. Il set di dati di PeriodO permette il

confronto e la mappatura tra periodi definiti da diversi grandi database archeologici e documenta centinaia di determinazioni relative ad informazioni cronologiche presenti nella letteratura archeologica. Anche PeriodO modella i periodi in base ai concetti di SKOS rappresentati in formato RDF. Mentre l'estensione spaziale è espressa attraverso identificatori collegati a risorse esterne come DBpedia, i periodi sono descritti con una citazione testuale ripresa dalla fonte oppure con una approssimazione ricavata dalla modellazione dell'ontologia OWL-Time, ad esempio, attraverso un intervallo temporale descritto con valori compresi tra *time:IntervalStartedby* e *time:IntervalFinishedby*.

Accanto a thesauri che possiamo definire generalisti, sebbene circoscritti ad uno specifico dominio, sono presenti in rete altri strumenti destinati a illustrare la terminologia di un singolo ambito di studi; spesso si tratta di iniziative finalizzate alla digitalizzazione di corpora di oggetti, documenti e edizioni critiche. Tra questi si annovera, ad esempio, THOT che raccoglie un insieme di risorse per documentare e codificare con termini controllati il patrimonio culturale dell'antico Egitto. Il progetto, terminato nel 2018, mirava alla definizione di un formato di interscambio XML/TEI per consentire l'integrazione degli archivi. TOTH prendeva spunto da precedenti iniziative come il MET. Nelle intenzioni degli sviluppatori il thesaurus doveva essere progressivamente allineato ad altri strumenti terminologici come quelli pubblicati dal British Museum, dal Getty Research Institute, dal Deutsches Archäologisches Institut e, infine, dalla *Digital Topographical Bibliography*, il cui sito online non appare aggiornato dopo il 2017.

Diversità linguistiche e, soprattutto, differenze culturali nella creazione degli archivi digitali costituiscono un ostacolo per una reale fusione e condivisione dei dati. Il ricercatore può, oggi, utilizzare numerosi strumenti terminologici il cui obiettivo è ridurre le ambiguità lessicali, ottimizzare lo studio dei dati e abilitare ricerche incrociate tra repository differenti. Sono numerosi i vocabolari e i thesauri disponibili in rete: *Open Context*, ad esempio, offre agli utenti una lista di circa trenta vocabolari utilizzabili per l'annotazione delle risorse (Kansa *et al.* 2019); l'allineamento tra gli strumenti esistenti consente di salvaguardare l'integrità e la consistenza degli archivi, riducendo il rischio di interventi successivi necessari per mantenere la standardizzazione dei propri dati.

### 3.3. I dati FAIR in archeologia

Numerose iniziative sensibilizzano e sostengono a livello europeo lo sviluppo dei dati FAIR, ad esempio, EOSC, FAIRsFAIR e GO-FAIR (Dunning *et al.* 2019) con l'obiettivo di incoraggiare ulteriori e più specifiche azioni da parte di comunità e istituzioni di diverse discipline.

Nonostante ciò, l'importanza di strutturare i dati in accordo alle raccomandazioni FAIR non è ancora un obiettivo preliminare della ricerca e sono ancora pochi gli esempi che riguardano il mondo dei dati dell'archeologia. Benché sia stato riconosciuto il rilievo che le indicazioni FAIR hanno per rendere trovabili, accessibili, interoperabili e riutilizzabili gli archivi digitali, la discussione tra i ricercatori è ancora limitata alla generica affermazione sulla utilità di disporre di dati in rete condivisi in modo efficace.

In un recente volume della rivista *Internet Archaeology* (Jakobsson *et al.* 2021), interamente dedicato alle problematiche di archiviazione del dato digitale, ampio spazio è stato riservato al tema delle possibili applicazioni dei principi FAIR agli archivi archeologici. In modo particolare, i vari articoli registrano quali cambiamenti siano intervenuti, nei diversi paesi, nelle pratiche di digitalizzazione del dato affinché gli archivi risultino allineati alle regole FAIR. Questo aspetto non è soltanto connesso alle esigenze di ricerca, ma è anche correlato agli obblighi, amministrativi o etici, di conferimento dei dati digitali presso repository accademici, nazionali e internazionali (Geser *et al.* 2022). I contributi segnalano, all'interno di un quadro nazionale ancora poco coerente, norme per la creazione di archivi digitali aperti, mentre altri evidenziano gli sforzi effettuati in direzione di rendere FAIR i propri dati.

Gli oggetti digitali conservati nello portale svedese NHB (Löwenborg *et al.* 2021) sono in gran parte reperibili poiché hanno un identificatore persistente e sono ricercabili attraverso motori e aggregatori nazionali mentre i metadati sono recuperabili con il protocollo OAI-PMH. Il portale consente al pubblico di accedere e scaricare l'intero archivio dei siti e dei monumenti antichi, nonché il registro degli scavi e delle indagini eseguite a livello territoriale sia come documenti in formato GIS sia attraverso il servizio WMS (*infra* 141). Sebbene i metadati siano documentati con lo schema Dublin Core e CIDOC-CRM e, quindi, alle relative pratiche che variano da paese a paese, l'interoperabilità è bassa poiché le voci



riprese dai dizionari terminologici non utilizzano identificatori persistenti. Infine, manca il contesto di creazione del dato digitale, in generale, disponibile soltanto nei documenti associati al record principale.

Anche il governo olandese sta promuovendo l'uso dei dati FAIR (Hollander 2021) attraverso la costruzione di una infrastruttura comune alle scienze sociali e umane nella quale depositare gli archivi digitali. Il centro che gestisce il repository nazionale per i dati della ricerca è il DANS in collaborazione con il RCE. Il deposito dei dati archeologici è regolamentato da apposite leggi ed è obbligatorio; gli archivi devono essere conformi ad uno standard nazionale approvato nel 2016.

Lo stato della archiviazione digitale nel Regno Unito è alquanto complesso poiché convivono differenti Commissioni Reali che hanno il compito di tutelare il patrimonio archeologico dell'Inghilterra, della Scozia e del Galles spesso agendo in modo separato (Richards 2021). Il NRHE è il principale deposito per la maggior parte degli interventi archeologi sul terreno in Scozia (McKeague 2021). Come per il NHB e il DANS anche il NRHE fornisce strumenti per la ricerca delle informazioni, inclusi record spaziali, che rispettano i principi FAIR. Questo tipo di approccio favorirà la crescita esponenziale dei volumi degli archivi depositati incoraggiando la progettazione di nuove modalità di pianificazione per il conferimento di grandi quantità di dati, incluse le informazioni telerilevate. Più definita è la situazione dell'Inghilterra, contraddistinta dalla presenza e dalle iniziative dell'ADS il cui repository, attivo dagli anni '90 presso l'Università di York, ha rappresentato uno dei primi tentativi di implementare e gestire archivi digitali online. Nonostante, la lunga attività dell'ADS nella promozione e diffusione dei principi dell'OS, la realtà del Regno Unito in tema di adesione ai dati FAIR è disomogenea e dipendente dalle politiche delle singole università anche per la carenza di adeguati sostegni finanziari (Richards 2021).

Per quanto riguarda i dati archeologici depositati in Belgio, il *Flanders Heritage* ha istituito un portale centralizzato che, dopo l'entrata in vigore nel 2016 di una nuova legislazione sul patrimonio culturale, è stato reso conforme ai principi FAIR (Hacıgüzeller *et al.* 2021).

Anche esaminando lo scenario italiano si evidenziano elementi di forti contraddizioni caratterizzate dalla compresenza dell'archivio centrale ministeriale con repository locali, spesso, di tipo accademico. Dal

2016 l'archiviazione digitale e, soprattutto, la pubblicazione in formato aperto degli archivi è di competenza dell'Istituto Centrale per l'Archeologia (ICA) del Ministero della Cultura, ma soltanto nel 2019 è stata pubblicata, a cura dell'ICCD, la traduzione italiana delle linee guida per l'applicazione dei principi FAIR nel quadro di una collaborazione con il progetto europeo PARTHENOS. Mancano ancora norme per il conferimento dei dati relativi ai risultati dell'archeologia preventiva e, più in generale, si assiste ad una timida attenzione delle tematiche della accessibilità e della cura degli archivi digitali. Soltanto nel quadro del progetto Geoportale Nazionale per l'Archeologia (GNA), finalizzato alla standardizzazione della documentazione scientifica prodotta nell'ambito di tutte le ricerche archeologiche svolte in Italia, i metadati sono stati scelti con l'obiettivo dichiarato di aderire ai principi FAIR (Calandra *et al.* 2021). La situazione italiana molto frammentaria contrasta con l'impegno di numerosi ricercatori italiani che a livello europeo, nei progetti ARIADNE (Niccolucci 2017), E-RIHS (Castelli, Felicetti 2019) e PARTHENOS (Di Giorgio, Ronzino 2019), hanno contribuito alla diffusione dei FAIR data anche in campo archeologico (Niccolucci *et al.* 2018).

La breve rassegna sulle soluzioni sviluppate a livello nazionale per adeguare gli archivi digitali alle raccomandazioni FAIR mostra una realtà disomogenea non solo tra paese e paese, ma anche all'interno di ciascun singolo Stato tra iniziative centralizzate, in genere ministeriali, e progetti universitari spesso poco attenti al rafforzamento di una prospettiva unitaria.

L'adesione ai principi comuni dell'OS e alle indicazioni sui FAIR ha contribuito a consolidare un nuovo approccio scientifico condiviso e collaborativo e, nello stesso tempo, ha favorito l'emergere di proposte innovative rispetto alle raccomandazioni standard sui dati FAIR.

S. Hermon e F. Niccolucci (2021), ad esempio, ritengono che anche le ipotesi e i risultati dovrebbero essere conformi ai principi FAIR e, pertanto, propongono di estendere la R di *Reusable* in *Reusable, Relevant e Reliable*; per i due studiosi è necessario ricostruire una linea di verità e affidabilità scientifica che anche una macchina può comprendere e suggeriscono l'uso dell'ontologia CIDOC-CRM per supportare sul piano delle tecnologie semantiche questo ambizioso obiettivo. N. Barbuti (2021), a sua volta, aggiunge una quarta R ai principi FAIR dedicando una istanza

alla *Resilience*, mentre M. Figuera (2021) suggerisce un suggestivo approccio finalizzato al trattamento dei dati incerti introducendo un coefficiente di affidabilità espresso in termini di logica *fuzzy*.

Gli archivi digitali conformi alle regole FAIR rivelano una realtà abbastanza frammentaria per motivazioni che variano da questioni di formazione a specifiche competenze sui metadati e sui linguaggi semantici. Per superare diffidenze e problematiche legate alla formazione, alcune iniziative sono state promosse per fornire linee guida specifiche per il mondo dell'archeologia (Geser *et al.* 2022); ciò consentirà di creare archivi compatibili con i principi FAIR e, soprattutto, farà aumentare la consapevolezza da parte dei ricercatori sull'utilità dell'uso di tali indicatori per l'allineamento degli archivi digitali alle più recenti raccomandazioni dell'OS. Parallelamente occorrerà sensibilizzare i responsabili dei repository certificati sulla necessità di una validazione dei dati anche rispetto ai principi FAIR e non soltanto in relazione agli aspetti legati ai formati di deposito. L'ADS, ad esempio, descrive in dettaglio le modalità specifiche con cui garantisce la conformità a tutte le norme FAIR; per ogni principio o sotto-principio l'ADS illustra come si può ottenere la congruenza degli archivi. Le indicazioni segnalate dall'ADS sono una utile guida per chi voglia verificare la corrispondenza dei propri dati digitali alle raccomandazioni FAIR e, principalmente, per migliorare le politiche di cura e archiviazione dei dati digitali.

Gli archeologi sono tra i principali sostenitori dei dati aperti e la testimonianza di un atteggiamento positivo verso la condivisione è rafforzata proprio dall'ampio interesse per i principi FAIR. Tuttavia, la diversità dei tipi di dati archiviati e, in particolare, dei metodi utilizzati dagli archeologi per acquisire ed elaborare la documentazione archeologia, suggerisce che l'adozione di FAIR porrà sfide significative, spingendo verso una maggiore collaborazione tra i produttori e consumatori di informazioni digitali (Richards *et al.* 2021).

Tracciabilità e armonizzazione sono due componenti chiave del processo di creazione di una archeologia *data-intensive*; i principi FAIR contribuiscono in maniera determinante a migliorare la qualità, l'affidabilità e la rilevanza dei dati condivisi. Nello stesso tempo, una più attenta cura degli archivi aiuterà a preservare l'integrità nel lungo periodo dei dati e a strutturare le pubblicazioni scientifiche in un formato logico che

illustri compiutamente il ragionamento che sta dietro una interpretazione archeologica. Il lettore potrà, così, verificare i vari passaggi formali e seguire il ricercatore nel suo processo di costruzione della conoscenza archeologica, come aveva ipotizzato Gardin agli albori dell'Archeologia Digitale (Marlet, Rodier 2019).

### 3.4. Metadati

La corretta documentazione del processo di digitalizzazione di una fonte e la descrizione accurata di ogni singolo oggetto di una collezione di dati contribuiscono a identificare in modo efficace un insieme di risorse digitali. La registrazione di tutte le caratteristiche e le proprietà di una raccolta di oggetti digitali distingue un archivio, che può essere sfruttato in modo estensivo e continuativo, da un insieme di documenti digitali interpretabili soltanto dal gruppo che li ha prodotti. Su un piano astratto, la memorizzazione di una fonte digitale (un database, un filmato, un GIS, un'analisi statistica, etc.) dovrebbe sempre essere associata alle ragioni della sua digitalizzazione e ad alcuni elementi di base quali: l'origine, la provenienza, il contenuto, la struttura e i termini e condizioni applicabili al suo uso e riuso.

Le informazioni che accompagnano una risorsa digitale risultano ancora più rilevanti se si pensa che nel Web circolano dati descritti nella maggior parte dei casi ricorrendo a strutture formali o database differenti tra loro nelle voci utilizzate e nel livello di analisi ed interpretazione che tali dati consentono di ricavare. L'esigenza di garantire una coerente ed efficiente normalizzazione ha, nel tempo, spostato l'interesse degli studiosi dalla definizione di ristretti modelli di rappresentazione, universalmente riconosciuti, all'adozione di schemi di metadati in grado di consentire l'interoperabilità e l'integrazione tra tipi differenti di risorse online (testi, audio, video, immagini, etc.).

Fino a pochi anni fa, la parola metadati era un termine limitato alla comunità degli informatici e degli specialisti dei dati geospaziali. Definiti originariamente con l'espressione "dati sui dati", oggi si preferisce indicare i metadati come contenitori o portatori di dati distinguendo l'informazione dal mezzo di trasporto. Uno degli obiettivi principali dei metadati è quello di modellare, attraverso specifici descrittori, gli attributi essenziali di una classe di oggetti. I metadati raccolgono informazioni su qualsiasi tipo di documento digitale, anche non testuale, codifi-

cato in qualsiasi formato consentendo la scoperta, l'accessibilità e la fruibilità di file depositati in un archivio.

In base alle finalità del processo di archiviazione gli schemi possono essere generali o specifici con un set minimo di descrittori. I tipi più comuni di metadati includono informazioni relative:

- al contenuto di un file o di un progetto: possono includere valori da un vocabolario controllato per la classificazione e l'indicizzazione delle risorse;
- al contesto generale al quale il documento appartiene;
- all'amministrazione del record con dati sull'invio, l'accesso, i diritti; sono, spesso, definiti meta-metadati poiché registrano tracce della creazione e modifica del dato;
- al processo di digitalizzazione che identifica il software adoperato, la dimensione del file e altre informazioni dettagliate;
- alle azioni da intraprendere per assicurare la conservazione del documento digitale.

I metadati mirano a rendere le risorse digitali facilmente identificabili, recuperabili e utilizzabili attraverso la memorizzazione di informazioni descrittive e contestuali, ma, nello stesso tempo, costituiscono la spina dorsale della *digital curation*, poiché, senza di essi, una risorsa digitale potrebbe essere introvabile o indistinta.

Gli schemi possono essere raggruppati in tre categorie:

- i metadati descrittivi: identificano in modo accurato l'oggetto digitale, a livello di collezione o progetto e di singola risorsa, che risulta così facilmente recuperabile e consultabile; descrivono, ad esempio, elementi quali la cronologia, parole chiave e posizione geografica del documento;
- i metadati amministrativi e tecnici: incorporano le informazioni relative ai formati di archiviazione dei file alla loro modifica e preservazione nel lungo periodo garantendo l'autenticità e l'integrità del contenuto;
- i metadati strutturali: forniscono informazioni per la localizzazione del documento all'interno del sistema di conservazione anche legando un documento ad un fascicolo informatico.

Fornire l'accesso agli archivi digitali e garantirne la conservazione nel lungo periodo è un compito difficile da realizzare senza precise indicazioni, finanziamenti e regole chiare per tracciare il ciclo di vita del dato. Poiché la politica di conservazione di un archivio digitale investe gli elementi di un'organizzazione, indipendentemente dalla sua complessità e dimensione, l'identificazione di specifici piani di azione risulta determinante. Per ridurre i rischi connessi alla perdita dei dati, ogni particolare gruppo di utenti ha sviluppato un suo schema in modo da consentire la migliore descrizione possibile di un tipo di risorsa per le esigenze amministrative, di ricerca e di conservazione. L'elaborazione degli schemi è avvenuta con il consenso delle varie comunità che hanno avuto, nel tempo, il compito di presentare, approvare e pubblicare gli elementi dei metadati.

Gli schemi contengono generalmente definizioni semantiche rappresentate attraverso linguaggi standard adoperati per descrivere la struttura e il contenuto del record, assicurando così la coerenza formale necessaria per la condivisione e la ricerca (anche automatica) dei dati, per la gestione del processo di creazione, registrazione e definizione della provenienza, per la memorizzazione dei processi tecnici e, infine, per l'amministrazione dei permessi di accesso. Ideati per l'indicizzazione e la ricerca di record in grandi archivi centralizzati, ben presto i metadati sono stati impiegati per la ricerca delle informazioni presenti sul Web.

Con l'obiettivo di migliorare il processo di integrazione degli archivi aperti, nel 2001 è stato messo a punto il protocollo OAI-PMH per semplificare e rendere efficiente la diffusione e l'interoperabilità dei repository; il protocollo si basa sulla distinzione tra i fornitori di dati (*content* o *data providers*), il cui obiettivo è rendere disponibili le risorse associate ai dati descrittivi, e i *data service* che aggregano i dati e indicizzano le risorse consentendo ricerche *cross-archive* e *cross-domain*.

Il contenuto dei metadati può essere ulteriormente arricchito attraverso l'uso di elenchi di parole e thesauri specifici e standardizzati che consentono di completare in modo controllato elementi dello schema di metadati, migliorandone l'usabilità e l'affidabilità.

L'esigenza di adeguare i dati ai principi FAIR sta incoraggiando gli archeologi ad un ripensamento sulla funzione dei metadati, sebbene, in assenza di schemi internazionali, ogni paese abbia messo a punto un suo specifico schema specifico. L'introduzione dei LOD e delle tecnolo-

gie semantiche ha contribuito a consolidare l'importanza del valore scientifico degli archivi digitali standardizzati favorendo una maggiore diffusione dei metadati e di strumenti per la creazione di dati affidabili ed autorevoli (Baca 2016).

Ancora una volta l'esperienza maturata presso l'ADS, poi seguita da tDAR, può essere considerata come modello di riferimento internazionale per l'uso e l'applicazione di metadati standard. Lo schema adoperato a livello di descrizione del progetto è costituito dagli elementi del Dublin Core, mentre, per la conservazione dei file, la scelta di associare a ciascuna risorsa un singolo record di metadati dipende dalla natura dei dati o da gruppi di dati che condividono formati e caratteristiche comuni. Diversamente dall'impostazione scelta dall'ADS, il tDAR mantiene un record di metadati per ogni file, sebbene ciò richieda tempo aggiuntivo e, spesso, una duplicazione delle informazioni.

I metadati sono di straordinaria importanza per la conservazione del record digitale nel lungo periodo; non soltanto è possibile avere traccia del record, ma, soprattutto, è possibile seguire l'intero ciclo di vita del dato e, se necessario, aggiornarlo per renderlo compatibile con software più recenti. I metadati aiutano, in definitiva, i ricercatori a rendere trasparenti e accessibili i loro archivi migliorando i flussi di lavoro e di gestione dei dati di ricerca e fornendo, nello stesso tempo, descrittori semantici che agevolano l'integrazione e l'aggregazione delle fonti digitali. L'uso di termini che derivano da thesauri internazionali arricchisce il contenuto informativo della risorsa e favorisce la creazione, anche automatica, di connessione tra i dati in rete.

Se agli studiosi è richiesto di associare agli archivi digitali metadati standard liberamente accessibili, dal lato dei sistemi informativi occorrerà definire opportune strategie per garantire un deposito di dati conforme ai principali standard internazionali.

La situazione dell'archiviazione digitale, come abbiamo visto, non è omogenea a livello internazionale e all'interno dei singoli paesi, per cui sono i repository internazionali e le infrastrutture di ricerca sovranazionali a supplire all'assenza di politiche coordinate sulla catalogazione dei dati, sulle licenze di condivisione, sulla implementazione dei LOD e, infine, sulle migliori strategie di cura degli archivi digitali. Probabilmente, in questo modo, si potrà superare la resistenza di larga parte dei

ricercatori che, in quanto produttori di contenuto, prediligono utilizzare metadati locali piuttosto che schemi standard che richiedono un consistente investimento in termini di tempo, energia e attenzione (Faniel *et al.* 2021). Nello stesso tempo, le tecnologie semantiche potranno contribuire a migliorare non soltanto la comunicazione tra gruppi di ricerca, ma anche lo scambio di informazioni tra differenti specialisti che collaborano alle stesse indagini e che spesso riutilizzano la documentazione prodotta da altri ricercatori.

### 3.5. *Paradata*

Nell'ambito dei metadati un particolare ambito di descrizione dell'informazione riguarda la registrazione degli aspetti tecnici della risorsa, dalla digitalizzazione alla sua elaborazione e post-elaborazione. Questa particolare categoria di descrittori, definiti *paradata*, conservano dati utili relativi alla metodologia di acquisizione dell'oggetto digitale ed alla qualità e alle motivazioni della sua digitalizzazione. Pur essendo simili ai metadati amministrativi o contestuali, i *paradata* si occupano soprattutto di illustrare i metodi, fornendo informazioni utili per una corretta valutazione della affidabilità della risorsa a partire dagli strumenti e dalle tecnologie adoperate.

La registrazione dei parametri selezionati per la digitalizzazione di una qualsiasi risorsa può aiutare il ricercatore nella progettazione della fase di acquisizione di altri oggetti individuando, e, quindi, riducendo, potenziali rischi di errori strumentali e/o umani. Inoltre, i *paradata*, descrivendo la metodologia, le apparecchiature e i processi di digitalizzazione forniscono informazioni che appaiono essenziali per il corretto riutilizzo dei dati.

Di recente è stata evidenziata l'esigenza di arricchire i metadati con i *paradata* allo scopo di rendere trasparente il processo di conoscenza del dato in modo da consentirne una corretta interpretazione e un più efficiente riutilizzo degli archivi (Huvila 2022); soltanto la comprensione dei meccanismi di creazione e cura della risorsa digitale possono rendere concreto lo scenario di una scienza realmente *data-intensive*.

Vedremo nel paragrafo sul 3D (*infra* 160) come i *paradata* siano collegati alla diffusa consapevolezza che si debba garantire la trasparenza dei dati condivisi soprattutto laddove la ricostruzione, che si propone in



rete, contenga ipotesi che non sono immediatamente percepibili dalla semplice osservazione dell'oggetto (Bentkowska-Kafel, Denard, Baker 2012; Moore, Rountrey, Ketter 2022).

I *paradata* documentano il processo di interpretazione del record archeologico fornendo all'utente scopi, contesto e affidabilità dei metodi adoperati per la visualizzazione dell'oggetto 3D. Con l'obiettivo di spingere i ricercatori a fornire informazioni complete sul ciclo di vita della fonte, i *paradata* sono stati inclusi nella Carta di Londra per assicurare la trasparenza delle fasi di creazione e visualizzazione degli oggetti.

L'enfasi sull'importanza di associare la descrizione delle scelte adottate dal ricercatore in termini di metodologie e ipotesi ricostruttive adoperate è connessa alla diffusione delle tecnologie per la creazione e/o modellazione di oggetti 3D e, nello stesso tempo, di software per la visualizzazione. L'uso dei *paradata* avrebbe l'obiettivo di offrire all'utente strumenti per decodificare il linguaggio di cui lo studioso si è servito per realizzare il modello digitale e, soprattutto, per comprendere i vari passaggi logico-deduttivi impliciti nella ricostruzione scientifica dell'oggetto.

I principi di trasparenza e correttezza su cui si basano i *paradata* e che riguardano le motivazioni contenute in una risorsa, dovrebbero scoraggiare la pubblicazione in rete di ricostruzioni fantasiose o del tutto prive di valore scientifico e analogamente rimuovere ogni forma di giudizio preconcepito o ambiguità.

Uno degli obiettivi dei *paradata* è permettere di separare in modo netto il dato digitale, inteso come registrazione di base non manipolabile, dalle sue trasformazioni che possono assumere significati differenti nel tempo. I *paradata* utilizzati per la descrizione delle varie trasformazioni digitali dipendono dai dati grezzi, dai processi di selezione e decisione, dai ragionamenti e dalla combinazione motivata di differenti oggetti in un unico record. Mentre i metadati conservano informazioni oggettive, i *paradata* vengono adoperati per memorizzare la soggettività presente nella costruzione del record archeologico.

Una questione importante è il livello di granularità delle informazioni da associare al record digitale. Alcune sono certamente essenziali come l'identificazione del/i creatore/i della risorsa e delle differenti repliche oppure la descrizione delle prove documentali sulle quali poggia la ricostruzione archeologica. L'illustrazione del processo ricostruttivo

sarà chiaramente in connessione con il livello di dettaglio scelto dal ricercatore e più analitiche sono le informazioni, tanto più profonda e complessa si mostrerà la verifica della documentazione e delle varie fasi di esame e comprensione del dato.

Oltre agli aspetti quantitativi relativi alla particolarità del singolo strato informativo descritto, si deve considerare anche la qualità dell'informazione. Qual è il peso attribuibile alle specifiche parti registrate rispetto all'intero oggetto digitale? Ogni singola porzione ricostruita ha lo stesso valore delle altre? Le prove documentali sono identiche per tutti gli elementi o in talune circostanze i dati possono essere meno affidabili? Queste sono solo alcune delle tante domande che devono essere esaminate per ottenere un compromesso efficace tra qualità e quantità che permetta di utilizzare i *paradata* in modo efficiente senza soffocare gli obiettivi della ricerca.

La pubblicazione dei *paradata* permette al pubblico di ricavare una comprensione accurata e completa dei processi, in particolare di quelli decisionali e interpretativi, coinvolti nella produzione e visualizzazione in rete di un oggetto 3D. La registrazione di *paradata* adeguati è certamente preferibile alla pubblicazione online di ricostruzioni scintillanti, ma mute e non verificabili.

Se c'è unanimità da parte della comunità scientifica nel considerare la descrizione dei processi decisionali un passaggio determinante per la validazione di un oggetto digitale che si ponga l'obiettivo di fornire informazioni sul passato, meno consenso si registra su cosa documentare per assicurare la leggibilità ed il riuso della fonte digitale. Sul tema è di recente intervenuto Huvila (2022) il quale ha rimarcato che la proliferazione di schemi per la registrazione delle informazioni non rivela automaticamente quali elementi descrittivi debbano essere catturati con i *paradata*. L'estrema varietà ed eterogeneità di metodi e dati non aiuta ad indentificare una lista di informazioni elementari e prioritarie; spesso dietro una diversità di pratiche scientifiche possono celarsi impostazioni e approcci che tendono a considerare importanti una serie specifica di informazioni non valutate nello stesso modo da altri studiosi.

Alcune informazioni sul processo di digitalizzazione possono essere acquisite autonomamente; per esempio, ogni scatto fotografico contiene dati EXIF che includono i parametri della macchina fotografica (lun-

ghezza focale obiettivo, esposizione, data, etc.) associati talvolta a coordinate geografiche rilevate con un GPS integrato nel dispositivo (Huvila 2022). Anche i computer conservano dati sulla creazione e modifica dei file. L'uso di strumenti adeguati consente di poter acquisire e trattare automaticamente differenti tipologie di *paradata* già disponibili in alcune tipologie di file (Felicetti, Lorenzini 2011). Inoltre, la lettura delle foto può fornire tracce utili su azioni, strumenti, condizioni climatiche che contribuiscono ad una più ampia definizione del contesto nel quale si è realizzata la digitalizzazione della risorsa.

Börjesson *et al.* (2022) hanno analizzato un database relazionale relativo ad un intervento di scavo, segnalando la difficoltà di individuare quali informazioni possano essere estratte ed utilizzate per compilare i campi di un record di *paradata*. La struttura dell'archivio può illustrare alcuni aspetti relativi all'intervento segnalando, ad esempio, quale insieme di campi sia stato oggetto di particolari tipologie di dati. Altri elementi che possiamo dedurre dal database relazionale riguardano la ripartizione dell'area in distinte zone di intervento, i metodi di indagine e la loro eventuale trasformazione, e, infine, la formulazione di ipotesi e ricostruzioni. Il dataset, pur mostrando incoerenze nella nomenclatura e nelle relazioni tra le varie unità, si rivela una preziosa traccia per la comprensione dello sviluppo dell'indagine. Tuttavia, senza un consistente intervento umano, non è possibile trasformare automaticamente parte della struttura di un database in *paradata*.

In alcuni casi, dallo studio della terminologia selezionata nel database, è possibile dedurre la presenza di differenti archeologi che potrebbero aver lavorato in aree o in stagioni diverse, oppure aver registrato informazioni che appaiono disomogenee soltanto perché legate alla diversa formazione degli specialisti coinvolti nel progetto. Un archivio digitale può, dunque, apparire come un prodotto disorganizzato o, meglio, organizzato in modo non sistematico perché compilato da vari ricercatori in tempi diversi. Tutto ciò rende complessa e non automatizzabile l'identificazione dei valori da associare ai *paradata*. Spesso, inoltre, mancano informazioni puntuali sulla corrispondenza tra i dati raccolti, le schede cartacee registrate direttamente sul campo, le informazioni caricate nel database e talvolta modificate per esigenze di normalizzazione dei dati, e infine, gli archivi pubblicati in rete o a stampa. All'interno del

processo di documentazione e registrazione appena delineato si perde la connotazione di dato primario (D'Andrea 2016).

In conclusione, nonostante un ampio consenso sui dati aperti, sulle raccomandazioni FAIR e sulla utilità di buoni metadati e *paradata* per descrivere gli archivi online, troppo spesso i dataset in rete risultano ancora incompleti, inaccurati e inconsistenti, sia sul piano dei formati che delle terminologie, rendendo in questo modo problematico e in parte rischioso, il riuso acritico dei dati e compromettendo l'accessibilità e, quindi, l'esistenza stessa degli archivi digitali nel lungo periodo.

### 3.6. Ontologie di dominio e *task ontology*

L'inserimento di metadati, *paradata* e termini ripresi da thesauri per documentare e popolare gli archivi digitali sono necessari per rendere l'informazione in rete coerente con i principi dell'OS e conforme alle raccomandazioni e linee guida messe a punto in tema di scoperta, accessibilità e riuso dei dati. L'uso di strumenti semantici consente, nello stesso tempo, di verificare la consistenza del processo di codifica digitale dei dati contribuendo alla normalizzazione degli archivi e riducendo i rischi di ambiguità lessicali e concettuali.

Una pluralità di soluzioni, informatiche, tecniche e metodologiche, sono state sviluppate per ottimizzare le fasi di registrazione ed estrazione della conoscenza garantendo, nello stesso tempo, l'uso prolungato e la manutenzione della risorsa digitale. Tuttavia, affinché un uomo o una macchina possano comprendere il ragionamento scientifico incorporato in archivi che contengono dati eterogenei, non sempre le tecnologie descrittive delle risorse digitali (metadati e *paradata*) appaiono sufficienti.

Come abbiamo già visto (*supra* 60) l'archeologia fonda la rappresentazione della conoscenza su strutture formali e/o simboliche attestate all'interno di un linguaggio scientifico noto ai ricercatori. La descrizione e la successiva classificazione del dato primario non è la conclusione di un processo oggettivo e neutrale poiché lo studioso interviene nella definizione dell'oggetto proiettando le sue conoscenze e competenze.

Ogni oggetto contiene, dunque, un sapere implicito iscritto nella identificazione del record che condiziona anche la sua successiva interpretazione. Se il ricorso ai *paradata* assicura la registrazione delle informazioni relative alla digitalizzazione del record e alle motivazioni alla

base del processo di scoperta e identificazione del dato archeologico, più complessa è la formalizzazione dei ragionamenti, tecnici e concettuali, che possono celarsi in una risorsa digitale. L'estrazione della conoscenza implicita assicura non soltanto la correttezza formale dell'archivio, ma consente anche di rendere i dati realmente interoperabili e, soprattutto, integrabili indipendentemente dagli standard e dalle strategie di indagine selezionate.

La ricerca archeologica produce grandi quantità di dati non limitati alle strutture e ai materiali mobili riportati alla luce nel corso dello scavo o delle ricognizioni di superficie, ma anche ai risultati di ulteriori indagini specialistiche che contribuiscono alla conoscenza completa dei siti investigati. Un archivio archeologico è, molto spesso, composto da una molteplicità di informazioni che concorrono a connotare la complessità e l'eterogeneità delle informazioni registrate. I dati chiaramente si riferiscono ad uno stesso contesto, ma la selezione, l'aggregazione e l'interpretazione delle informazioni ha origine in specifici saperi e particolari tecnologie e metodologie di analisi.

Un insieme di conoscenze di diversa provenienza e finalità produce archivi eterogenei che vengono generalmente gestiti dai ricercatori con strutture formali differenti, difficilmente gestibili in modo unitario dai tradizionali sistemi informativi. L'estrema specializzazione delle informazioni se da un lato garantisce la creazione di dati di qualità, dall'altro rende più problematica, e, talvolta anche incompatibile, l'integrazione di tutte le informazioni acquisite.

In passato, i ricercatori hanno affrontato e risolto il tema della complessità degli archivi e della eterogeneità delle risorse progettando database relazionali in grado di contenere in uno o più sistemi tutte le informazioni acquisite.

Mentre una tale soluzione ha contribuito a migliorare il flusso delle informazioni all'interno del gruppo di ricerca, nuovi problemi sono sorti parallelamente all'emergere dell'esigenza di condividere i dati in forme comprensibili e accessibili. La mancanza di un modello completo e unitario di rappresentazione dei dati, indipendentemente dalla loro posizione e dal tipo di oggetto documentato, solleva due problemi: il ricercatore potrebbe non avere accesso completamente al flusso di informazioni non essendo in grado di individuarle; gli archivi non possono es-

sere integrati a causa delle differenti soluzioni adottate per la codifica e la gestione dei dati.

Poiché è impensabile che la comunità degli archeologi sviluppi un unico sistema di rappresentazione della conoscenza, occorre utilizzare le tecnologie semantiche per modellare e strutturare il sapere archeologico in modo flessibile e tale da essere adattabile ad ogni specifica struttura dati e al suo impiego. Mentre, a livello generale, per garantire un accesso ai dati è sufficiente standardizzare l'informazione utilizzando schemi di dati noti compilati con valori normalizzati interrogabili con linguaggi specifici, più complessa è l'implementazione di un sistema in grado di integrare i dati facendo ricorso ai consueti database relazionali.

Per superare i limiti dei database sono stati introdotti nuovi formalismi che consentono di catturare la semantica e la conoscenza racchiusa nel processo di codifica digitale delle fonti archeologiche. Il procedimento garantisce l'interoperabilità semantica cioè l'armonizzazione degli archivi attraverso un linguaggio logico costituito da classi e relazioni e composto da sostantivi e verbi nella forma di frasi Soggetto-Verbo-Oggetto. La rappresentazione si basa su un modello di codifica che rispecchia il linguaggio naturale. L'interoperabilità semantica non viene raggiunta in modo tecnico mediante l'uso di super-standard in grado di includere ogni standard esistente, ma assicurando la codifica delle informazioni su modelli formali di tipo più astratto e generale.

Tali modelli sono definiti ontologie (Gruber 1995; Guarino, Giarretta 1995; Guarino 1998) e, in base ai differenti livelli di generalizzazione ed astrazione della realtà, si dividono in:

- *Top-level ontologies*: descrivono concetti molto generali come lo spazio, il tempo, l'oggetto, il soggetto, l'evento e l'azione indipendentemente dal dominio. Sono anche dette fondazionali o di primo livello e, almeno in teoria, possono essere adoperate da comunità attive in differenti domini; consentono l'integrazione a un livello estremamente elevato includendo regole logiche che supportano la creazione automatica di ragionamenti;
- *Domain and task ontologies*: identificano rispettivamente il vocabolario relativo a un distinto dominio o a una attività generica mediante la specializzazione dei termini e delle relazioni presenti

al *top-level*; sono implementate per modellare settori molto specifici della conoscenza garantendo, pertanto, una interoperabilità più ristretta;

- *Application ontologies*: descrivono concetti che dipendono sia dal particolare dominio che dal task e sono spesso specializzazioni di entrambe le ontologie; corrispondono a ruoli giocati da certe entità quando realizzano una data attività.

Un'ontologia formale deve rappresentare accuratamente le aree di conoscenza comuni di un particolare dominio e le relazioni che gli utenti definiscono tra le entità. Inoltre, essa deve basarsi su classi e relazioni sufficientemente astratte per consentire la rappresentazione di stati caratteristici di mancanza di informazioni, cioè strutture informative che permettano di rappresentare non solo dati molto accurati, ma anche informazioni più generali, parziali o incerte.

Da questa prospettiva l'ontologia costituisce un mezzo per identificare e modellare livelli di conoscenza dettagliati o generali in base alle informazioni disponibili. I dati possono così essere interrogati a livelli elevati di astrazione o di specializzazione, assicurando la massima quantità di dati in riferimento al particolare valore di granularità richiesto dall'utente.

L'ontologia si configura come un prodotto dell'ingegneria della conoscenza che formalizza attività di analisi concettuale e/o di modellazione di dominio con un approccio fortemente interdisciplinare nel quale principi filosofici e linguistici si integrano per descrivere una determinata realtà su un piano alto di generalità ed astrazione.

A livello di implementazione fisica l'espressione ontologia viene adoperata per descrivere un sistema formale semantico caratterizzato da specifiche proprietà collegate tra di loro attraverso una teoria logica. Più semplicemente, si può definire come un sistema che codifica le regole formali, spesso implicite, che vincolano una data realtà. Un'ontologia formale è composta dai seguenti elementi:

- un ambito che indica l'oggetto che si intende rappresentare;
- le classi che mediante sostantivi descrivono gli oggetti specifici del mondo reale al quale l'ambito si riferisce;

- le relazioni che attraverso l'uso di verbi definiscono il collegamento tra le varie classi in quello specifico dominio;
- un insieme di regole logiche che consentono di dichiarare l'esistenza di classi, superclassi e sottoclassi che indicano una struttura gerarchica espressa, ad esempio, dai seguenti tipi di relazione:
  - a) *is\_a* che identifica sottoclassi e/o sottotipi;
  - b) *part\_of* che descrive in che modo i concetti sono tra di loro interrelati.

L'ontologia può essere descritta come una particolare forma di rappresentazione astratta della conoscenza costruita utilizzando fatti, relazioni ed assunti generali e asserzioni relative ad una specifica situazione. A differenza dei tradizionali database, lo schema ontologico è più flessibile e, soprattutto, facilmente riutilizzabile poiché, grazie al principio della ereditarietà, una classe condivide le proprietà e gli attributi espresse per una superclasse.

La struttura si basa su due componenti: le informazioni dipendenti dallo stato e la *core-ontology* che include istanze formali indipendenti dallo stato.

La formalizzazione di una ontologia si ottiene adoperando diversi formati: i due più popolari sono RDF e OWL che consentono di codificare e memorizzare, in forma di grafo, le cosiddette triple adoperate per identificare una relazione di tipo Soggetto-Verbo-Oggetto comprensibile anche da una macchina. Tuttavia, poiché la maggior parte dei dati è contenuta in database relazionali, le ontologie vengono popolate attraverso un processo di mappatura semantica nel quale l'utente costruisce una equivalenza tra le tabelle dei database e le classi dell'ontologia collegandole attraverso le proprietà in modo da poter costruire le triple leggibili dall'uomo e processabili dalla macchina (Sahoo *et al.* 2009; van Helden, Hong, Allison 2018). Il *mapping* può essere realizzato con una trasformazione automatica operata con opportuni software che supportano l'utente nella identificazione delle classi e delle relative relazioni, oppure attraverso l'esecuzione di query che estraggono e mappano, a richiesta dell'utente, le tabelle dei database sulla struttura del modello ontologico.

Attualmente, per la creazione di dati semantici nativi, si possono utilizzare alcune piattaforme, liberamente accessibili, come ResearchSpa-



ce, Arches e WissKi che guidano i ricercatori nel processo di creazione e gestione dei loro archivi digitali fornendo anche indicazioni sui modelli ontologici da utilizzare.

Sebbene ogni ricercatore abbia in teoria la possibilità di implementare il suo specifico modello concettuale adattandolo, se necessario, alla rappresentazione formale prescelta per la creazione dell'archivio digitale, è preferibile ricorrere ad ontologie standard che garantiscono il successivo *mapping* su schemi concettuali più astratti, assicurando la corretta interazione e integrazione delle risorse.

Per l'ambito umanistico sono disponibili numerosi schemi concettuali, tra i quali lo studioso può scegliere quello maggiormente idoneo agli obiettivi del proprio progetto. Per l'area archeologica e museale, l'ontologia di primo livello più diffusa è il CIDOC-CRM a cui si aggiungono numerose estensioni dello stesso modello concettuale che approfondiscono alcune specifiche aree di ricerca.

Non mancano altre proposte di ontologie sviluppate per singoli progetti come, ad esempio, il modello Horai che semplifica la gestione documentale delle attività eseguite nel corso di uno scavo archeologico assicurando la ricerca e la conservazione delle informazioni (Allepuz *et al.* 2021). Un ulteriore modello che riguarda specificamente l'ambito del patrimonio culturale italiano è ArCO che comprende una rete di schemi che descrivono il dominio dei beni culturali e i dati estratti dal Catalogo Generale dei Beni Culturali dell'ICCD, trasformati in RDF secondo i paradigmi dei LOD. ArCO riprende direttamente alcuni patterns, classi e proprietà degli schemi Cultural-ON, per i luoghi della cultura e degli eventi culturali, e OntoPiA che rappresenta la rete di ontologie e vocabolari della Pubblica Amministrazione italiana; inoltre, indirettamente riprende assiomi da altre ontologie, tra le quali si segnalano DOLCE (Borgo *et al.* 2022) e EDM; in tutta la rete ArCO riusa dodici *Ontology Design Patterns*.

Gli elementi principali che uno schema ontologico deve modellare sono lo spazio/luogo, il tempo/periodo e le cose, evidenziando i criteri che li identificano e le relazioni che vengono stabilite con altre entità, classi o sottoclassi. Per strutturare il ragionamento occorre, quindi, riorganizzare gli elementi definendo i concetti su cui la conoscenza si poggia in modo indipendente dall'applicazione; classi, assiomi e relazioni sono costruiti facendo ricorso alla logica matematica, filosofica e infor-

matica. Soltanto dopo aver delineato il processo di modellazione del dominio si può passare all'implementazione informatica. Invertendo le due fasi c'è il rischio che la ricerca dello schema migliore prevalga sulle esigenze del settore scientifico introducendo, in questo modo, elementi di determinismo tecnologico nel ragionamento.

Per quanto riguarda la sfera archeologica il problema non è sostituire i concetti esistenti con definizioni semantiche più precise riducendole ad una forma logicamente corretta, ma rivelare, cioè reingegnerizzare in forma esplicita, la natura, spesso, multiforme della conoscenza a cui ci si può riferire secondo l'utilità dell'oggetto archeologico nel suo rispettivo contesto d'uso.

Per l'implementazione di archivi semantici esistono strumenti e servizi (infrastrutture o repository) che sono allineati a *data-model* idonei per il dominio dei beni culturali; superando i vincoli dei tradizionali database i ricercatori possono impiegare modelli ontologici standard per costruire in modo semplificato una rete di connessioni e conoscenze dinamiche.

### 3.6.1. Il CIDOC-CRM

Il CIDOC-CRM (Doerr 2003; Niccolucci, Hermon, Doerr 2015) è una ontologia formale messa a punto per facilitare lo scambio di informazioni e l'integrazione tra fonti eterogenee relative ai beni culturali. Sviluppato dal 1996 dal CIDOC dell'ICOM, il CIDOC-CRM è stato accettato come standard ISO21127:2006. In particolare, il modello si propone di identificare un insieme di definizioni semantiche per armonizzare dati eterogenei derivanti da molteplici sorgenti locali in un insieme coerente di informazioni globali. La sua prospettiva è sovranazionale e, quindi, indipendente da qualsiasi specifico contesto. Il CIDOC-CRM non definisce, contrariamente dall'implementazione dei database, la struttura dei documenti, né propone alcuna terminologia, né cosa e in che modo il patrimonio culturale dovrebbe essere documentato; esso non consente di implementare specifiche procedure, ma fornisce un modello concettuale finalizzato alla creazione di documenti descrivendo la semantica sottostante ai dati culturali.

Diversamente dai thesauri e dai metadati, che specializzano e standardizzano i valori da adoperare o le informazioni associabili ai dati, il CI-

DOC-CRM consente di estrarre la cornice concettuale implicita nel processo di formalizzazione dei dati. Lo scopo del CIDOC CRM può essere sinteticamente definito come una forma di specializzazione della conoscenza, originariamente indirizzata alla catalogazione museale e successivamente estesa alla classificazione delle fonti e dei processi archeologici.

Il CIDOC-CRM è un dominio ontologico nel senso usato in informatica espresso con un *object-oriented semantic model*, comprensibile sia agli esperti di beni culturali che agli informatici. A differenza del Dublin Core, che offre uno schema rigido per la descrizione di informazioni associabili ad un oggetto digitale, il CIDOC-CRM si configura come un approccio semantico finalizzato all'integrazione dei dati attraverso una coerente formalizzazione concettuale della conoscenza racchiusa nel modello-dati. L'integrazione è raggiunta attraverso l'uso di uno schema generale (*core*) che definisce classi e proprietà e sul quale si possono modellare differenti rappresentazioni e specializzazioni.

I concetti fondamentali di spazio, tempo, attore e oggetto ruotano intorno all'elemento centrale della *Temporal\_Entity*; le entità principali si relazionano le une alle altre attraverso l'introduzione di un evento che evidenzia il carattere dinamico e flessibile del processo. Il modello evento-centrico adottato dal CIDOC-CRM è essenziale nella formulazione di asserzioni che rientrano nel campo della conoscenza storico-archeologica. Il processo di analisi, documentazione e catalogazione della cultura materiale è configurato come il risultato di un evento che mette in relazione oggetti, persone, idee, luoghi e tempi. Il modello evento-centrico fornisce una più accurata descrizione del passato o del ciclo di vita attuale di un oggetto culturale in quanto offre la possibilità di definire informazioni più dettagliate sull'origine e sull'interpretazione del record archeologico. In un tale senso la corretta definizione semantica dell'oggetto, dalla scoperta alla classificazione, agisce analogamente ai *paradata* consentendo all'utente finale di ricostruire e di contestualizzare il dato e comprendere il ragionamento del ricercatore.

In termini di concettualizzazione, l'attività di documentazione può essere facilmente schematizzata: essa registra un evento che si è verificato nel passato (la deposizione stratigrafica archeologica) e, allo stesso tempo, riporta l'azione dello studioso che scopre attività antropiche o naturali. Qualsiasi attività di scavo con la relativa metodologia possono essere facilmente descritte seguendo un semplice formalismo concettuale.

La versione 7.2.2. del CIDOC-CRM, rilasciata nell'ottobre del 2022, include 81 classi, identificate dai numeri da 1 a 99 preceduti dalla lettera E (*entity*), e 161 proprietà che presentano un numero da 1 a 199 preceduto dalla lettera P (*property*); Entità e Proprietà sono organizzate in gerarchie come avviene in RDFs.

Il sito web, che ospita il modello con tutte le versioni finora rilasciate, fornisce all'utente numerosi assiomi, già predisposti, per la rappresentazione grafica e concettuale di molte delle asserzioni utilizzate nel campo, ad esempio, della classificazione e datazione. Inoltre, sono illustrati numerosi progetti che attestano l'efficacia e la validità del modello proposto e la sua applicabilità in differenti settori. Il CIDOC-CRM è stato allineato ad altri schemi ontologici in modo da poter utilizzare le potenzialità espressive contenute in modelli concettuali che coprono l'area della conoscenza di altri domini e può essere arricchito utilizzando, per le istanze, i LOD (van Ruymbeke *et al.* 2018).

Il vantaggio nell'utilizzo di CIDOC-CRM è evidente: mentre forza gli archeologi a concettualizzare il proprio lavoro attraverso l'identificazione della metodologia e delle modalità di acquisizione e registrazione della documentazione, garantisce, al contempo, la verifica e il riuso dei dati. Gli svantaggi sono, invece, quelli relativi al tempo da spendere per identificare la corretta formalizzazione del flusso di lavoro e del ragionamento archeologico; l'approccio deriva dall'ingegneria della conoscenza, un dominio alquanto distante dalle attitudini dell'archeologo che segue, spesso, un suo particolare istinto e non un rigido protocollo.

L'importanza del CIDOC-CRM è attestata dall'uso del modello in numerosi progetti (STELLAR: Tudhope Binding, May 2008; ARIADNE: Aloia *et al.* 2017, SITAR) oppure come *backbone* per specifiche applicazioni (ARCHES: Myers, Dalgity, Avramides 2016; BEARCHAEO: Lombardo *et al.* 2023).

Il CIDOC-CRM, come abbiamo visto, può supportare l'archeologo nella creazione di risorse semantiche favorendo in questo modo l'interoperabilità nello scambio di dati tra sistemi informativi eterogenei. Di recente, per agevolare la formalizzazione e l'implementazione di archivi semantici sono stati messi a punto servizi online che offrono la possibilità di gestire e migliorare la qualità dei dati aperti anche sfruttando le potenzialità di un approccio basato sull'uso di modelli ontologici.

### 3.7. Repository e infrastrutture

Metadati, *paradata*, ontologie, FAIR data, dati aperti e LOD costituiscono gli strumenti concettuali di base del Semantic Web favorendo l'apertura, l'integrità e la riproducibilità della scoperta. Una nuova cultura della ricerca si è sviluppata negli ultimi anni, accelerando soluzioni e approcci in tema di accesso e gestione dei dati scientifici. Nonostante i profondi cambiamenti, anche di tipo comportamentale, che hanno investito numerosi settori della ricerca, sono ancora molti gli studiosi che mostrano difficoltà nell'accettare percorsi finalizzati a rendere condiviso e riutilizzabile il proprio sapere.

Nel clima attuale, la pratica più diffusa resta quella della pubblicazione tradizionale che è ancora il luogo privilegiato di comunicazione dei ritrovamenti. Occorre invertire la tendenza esortando gli studiosi ad intraprendere approcci più aperti, rendendo più trasparenti e comprensibili i processi di scoperta e sostenendo la verifica e il confronto dei dati e delle conclusioni. I ricercatori dovrebbero condividere materiali, protocolli e dati che hanno prodotto nel corso dei loro studi, sostenendo una espansione ed una accelerazione della scienza.

Per sollecitare il cambiamento verso una cultura della ricerca maggiormente aperta e condivisa, è necessaria l'approvazione di comportamenti che rendano praticabile lo scenario dell'OS, soprattutto attraverso lo sviluppo di strumenti tecnologici che facilitino l'adozione di flussi di lavoro trasparenti, collaborativi ed integrabili. Non è sufficiente mettere in condivisione dati ricorrendo a piattaforme che supportano la condivisione e il lavoro a distanza durante il ciclo della ricerca; occorre, al contrario, definire forme del tutto innovative per svolgere attività di alto livello in tutti i campi.

Per eseguire questi particolari compiti sono stati messi a punto repository e infrastrutture di ricerca che rappresentano piattaforme online ad accesso controllato che abilitano l'utente alla elaborazione di alcune funzioni, tra le quali il caricamento di dati utilizzando specifici metadati o *data-model* o attraverso un processo di *mapping* tra un sistema locale di metadati e lo schema o gli schemi selezionati dalla comunità che gestisce gli ecosistemi digitali.

I repository costituiscono il primo gradino delle innovazioni tecnologiche sviluppate per rendere concreta la prospettiva della Scienza Aperta.

Il repository è una piattaforma che permette di memorizzare una significativa quantità di informazioni sulle quali eseguire numerose operazioni, tra le quali protezione, classificazione, elaborazione, ricerca e duplicazione dei documenti. La gestione delle risorse è centralizzata e viene realizzata in un ambiente accessibile da più macchine hardware; l'accesso ai dati ed alle funzioni è generalmente controllato da password. Questi depositi virtuali assicurano una corretta gestione del flusso documentale mediante metadati e modelli-dati conformi a standard internazionali, e meccanismi di classificazione e ricerca per facilitare il recupero e la visualizzazione dei dati. L'obiettivo del repository è curare e preservare collezioni di oggetti e documenti in vario formato promuovendone il loro uso educativo, scientifico e culturale per un ampio pubblico o per un ambiente professionale.

Un deposito digitale è un archivio che fornisce l'archiviazione e l'accesso a lungo termine ai dati di ricerca che costituiscono la base per attività di tipo scientifico. Numerosi repository sono elencati sul sito *re3data.org*, un registro globale di archivi di dati di ricerca di tutte le discipline accademiche lanciato nel maggio 2013 per supportare l'utente nella ricerca dello spazio e dell'ambiente appropriato per i propri dati di ricerca.

Per depositare i dati su *re3data.org* occorre che l'archivio:

- 1) sia gestito da una istituzione sostenibile (ad es. biblioteca, università);
- 2) siano chiare le condizioni di accesso ai dati e al repository, nonché i termini di utilizzo;
- 3) pubblici esclusivamente dati di ricerca.

Impostando la voce *archaeology* per la ricerca, il sito mostra 45 progetti censiti e solo in parte certificati. Alcuni repository ospitano dati locali elaborati nel corso di particolari attività di ricerca, mentre altri consentono a diversi ricercatori di poter caricare e pubblicare i propri archivi. Alcuni repository sono sviluppati per conservare risorse digitali specifiche, come i modelli 3D oppure dati archeometrici. Oltre ai già segnalati ADS, Open-Context, tDAR, IANUS, DANS e Arachne, *re3data.org* riporta anche il progetto italiano MAPPA, un archivio archeologico digitale che pubblica dataset e report archeologici. Il MOD è il repository del progetto MAPPA con l'obiettivo di conservare e disseminare la docu-

mentazione archeografica e la letteratura grigia prodotta nel corso delle indagini archeologiche. Attualmente il sito ospita 146 archivi interrogabili e navigabili. Ogni documento digitale presente nel repository è identificato da un DOI ed è rilasciato con licenza CC-BY 3.0, il modo per consentire contemporaneamente la corretta circolazione dei dati e la tutela della paternità intellettuale. Il MOD nasce all'interno della lunga esperienza dell'Università di Pisa sulla pubblicazione e condivisione aperta dei dati. Per archiviare i dataset e renderli facilmente reperibili, MOD propone un adeguato sistema di metadatazione; lo schema attualmente in versione beta, descrive il retroterra storico e culturale di provenienza del dato archeologico ed è conforme allo standard ISO19115, messo a punto per identificare informazioni e servizi geografici.

Altri registri, disponibili in rete per la ricerca di archivi di dati, tra cui *Open Science Framework*, segnalano repository come *Sustainable Archaeology* (Canada) e *FAIMS* (Australia).

Alcune piattaforme, come Sketchfab, Smithsonian 3D Digitization e 3DHOP, censite in *rezdata.org*, sono utilizzate per la presentazione online di oggetti archeologici 3D. 3DHOP è un archivio di restituzioni 3D e soprattutto un framework open-source per la creazione di presentazioni Web interattive di modelli 3D ad alta risoluzione; 3DHOP converte i modelli nel formato Nexus che consente di visualizzare oggetti 3D con differenti livelli di dettaglio all'interno di un browser che sfrutta l'approccio WebGL (*infra 00*); il codice garantisce una ottima interattività anche in presenza di bassa larghezza di banda. Per la sua estrema affidabilità e semplicità di uso, la tecnologia 3DHOP è utilizzata da alcuni importanti archivi multimediali come l'inglese ADS, il tedesco TOPOI e l'infrastruttura europea ARIADNE.

Diversamente dai repository, le infrastrutture di ricerca comprendono, oltre alle necessarie componenti tecnologiche per la memorizzazione e condivisione dei dati, anche l'accesso a servizi messi a punto da ricercatori con qualificate ed elevate conoscenze di dominio. Le infrastrutture possono essere definite anche come ecosistemi condivisi, eterogenei, aperti che coinvolgono diverse capacità informatiche ed una comunità di utenti con specifiche funzioni ed applicazioni; questi sistemi riuniscono risorse rendendole accessibili e disponibili in un lungo termine per ricerche singole o di gruppo.

Lo studioso ha a disposizione una scrivania virtuale dove può trovare agevolmente dati di ricerca e strumenti per elaborarli, metodologie di indagine e risultati su argomenti simili. Questo approccio coinvolge le scienze dure e le discipline umanistiche. Le infrastrutture di ricerca offrono al ricercatore una vasta gamma di risorse digitali, grezze o elaborate, garantendo diversi tipi di analisi; esse funzionano come reti di conoscenze accademiche, che mirano a una maggiore efficienza della ricerca di alto livello.

Affinché i dati possano essere utilizzati è necessario che gli archivi siano interoperabili attraverso l'uso di metadati; per garantire l'integrazione le infrastrutture possono utilizzare anche forme di rappresentazione della conoscenza, come le ontologie, che migliorano l'integrazione dei dati e la loro ricerca e aggregazione.

Sulla garanzia del ciclo di vita dei dati, le infrastrutture di ricerca apportano un significativo contributo sostenendo un insieme specifico di processi che aiutano il ricercatore a gestire in modo più efficace i risultati delle proprie attività.

Le fasi del ciclo di vita dei dati che le infrastrutture supportano sono, abitualmente, le seguenti:

- 1) Creazione: pianificazione della gestione dei dati (formati, archiviazione, etc.);
- 2) Elaborazione: acquisizione, trascrizione e traduzione con la convalida e pulizia dell'archivio;
- 3) Analisi: interpretazione e conseguente creazione di nuovi dati derivati;
- 4) Conservazione: migrazione delle risorse nel formato migliore e su un supporto adeguato dopo la creazione di metadati;
- 5) Accesso: distribuzione e condivisione dei dati dopo il controllo dei diritti d'autore;
- 6) Riutilizzo: pianificazione di nuove indagini dopo la revisione dei dati primari, derivati e dei risultati.

Uno degli aspetti più delicati nella gestione dei dati di ricerca è garantire il reperimento della risorsa che deve essere identificabile in modo non ambiguo. In passato la migrazione dei siti o il cambiamento di standard o software determinava la perdita di collegamenti e relazioni, rendendo



impossibile il tracciamento automatico del dato e, quindi, la sua accessibilità. Oggi, grazie agli identificatori persistenti o PID è possibile ridurre il rischio della perdita del dato. Numerose agenzie di registrazione digitale assicurano che le risorse siano stabili e visibili anche ad agenti automatici consentendo una identificazione permanente in rete dell'oggetto digitale sia che si tratti di un singolo documento o di un archivio. Alcuni standard forniscono un riferimento duraturo agli oggetti digitali; tra questi il più noto è il DOI. L'esistenza di identificatori persistenti che tracciano i set di dati, incoraggiano i ricercatori nell'accesso, nel reperimento e nel riutilizzo degli archivi online.

Un ulteriore elemento di efficienza che le infrastrutture offrono al ricercatore è la protezione del diritto d'autore attraverso l'uso di licenze aperte come le Creative Commons che consentono il riutilizzo illimitato delle risorse garantendo il mantenimento dell'attribuzione originaria all'autore.

La Commissione Europea sostiene dal 2006 la creazione di infrastrutture di ricerca (ESFRI) all'interno di un quadro di profonda trasformazione delle pratiche tradizionali di ricerca. Uno dei punti salienti dell'intervento comunitario riguarda la sostenibilità delle infrastrutture per garantire la long-term preservation degli archivi digitali. Le piattaforme di ricerca sfruttano in modo prioritario le tecnologie semantiche per la raccolta, l'archiviazione, la gestione, l'accesso, l'elaborazione e la condivisione dei dati, nonché strumenti di comunicazione e collaborazione per la ricerca. L'accesso facile e controllato alle risorse e agli strumenti digitali integrati contribuisce alla creazione di comunità di ricerca; se l'architettura modulare dell'infrastruttura è flessibile, gli utenti utilizzeranno servizi e risorse assicurando, così, la sostenibilità del sistema e dei relativi dati conservati nel lungo periodo.

Archivi, standard, buone pratiche di archiviazione rappresentano la componente principale delle infrastrutture di ricerca. Anche il coinvolgimento degli utenti rientra nell'ottica della sostenibilità poiché attivando i ricercatori nella adozione di strumenti aperti e interoperabili si promuove il principio stesso della infrastruttura intesa come una confederazione di parti costitutive, tutte determinanti, ognuna delle quali può risultare utile per specifici utenti. La sostenibilità riduce il rischio di perdita o obsolescenza dei dati, assicurando, allo stesso tempo, la proprietà intellettuale e i diritti d'autore dei ricercatori.

Sono numerose le infrastrutture che offrono dati scientifici e servizi per la ricerca archeologica; ARIADNE offre un ambiente di ricerca virtuale specifico per il mondo dell'archeologia, mentre CLARIN, DARIAH, E-RIHS, PARTHENOS e RESILIENCE si occupano del più ampio settore dei beni culturali.

ARIADNE è un progetto sostenuto dal 2013 dalla Commissione Europea e che coinvolge numerosi partner appartenenti a 16 paesi europei. L'obiettivo iniziale era lo sviluppo di un'infrastruttura di dati con i relativi strumenti, lavorando sull'organizzazione e l'integrazione delle conoscenze attraverso l'uso del CIDOC-CRM, di thesauri standard, e dei LOD e offrendo formazione, buone pratiche e accesso alle strutture di ricerca (Richards 2012; Niccolucci, Richards 2013). ARIADNE si proponeva di superare la frammentazione dei dati archeologici distribuiti in numerosi archivi digitali, promuovendo una cultura scientifica aperta e sovranazionale basata sulla condivisione e sul riutilizzo dei dati. Considerata la disponibilità di archivi digitali conservati presso centri dati nazionali, depositi istituzionali e di varie altre risorse in rete il primo passo di ARIADNE è stato realizzato nella direzione dell'integrazione. Gli archivi dell'ADS, del DANS, di Fasti Online, di ARACHNE, di IANUS, del SND e del MIBAC sono stati integrati e resi accessibili attraverso l'infrastruttura virtuale di ARIADNE (Aspöck, Geser 2014).

In rete sono disponibili infrastrutture di ricerca sviluppate a livello locale che offrono servizi circoscritti per specifiche funzioni, tra cui la possibilità di scaricare strumenti per configurare una propria infrastruttura. Per esempio, la piattaforma FAIMS è stata messa a punto con l'obiettivo di facilitare la personalizzazione dei dati raccolti sul campo fornendo un supporto per la creazione di dati FAIR nel corso delle attività sul terreno (Ross *et al.* 2022). L'infrastruttura si configura come una piattaforma personalizzabile anche per la compilazione di metadati associati ai record registrati. FAIMS risponde soprattutto all'esigenza di ottimizzare la fase di strutturazione e convalida dell'archivio, con le connessioni tra tutte le risorse digitali acquisite (foto, disegni, dati spaziali e alfanumerici, etc.), direttamente sul campo. Finora sono stati pubblicati soltanto due dataset utilizzando il sistema FAIMS; ostacoli sono emersi a causa della complessità del sistema e dalla mancanza di specifiche competenze da parte dei ricercatori che hanno preferito utilizzare la piattaforma per pubblicare più velocemente in modo tradizionale i risultati delle loro indagini.

### 3.8 Aggregatori

Un ulteriore livello di organizzazione dei dati online relativi ad uno specifico dominio di interesse è quello degli aggregatori. Anche le infrastrutture di rete, a differenza dei repository, funzionano come aggregatori integrando differenti archivi digitali; tuttavia, l'aggiunta di servizi e funzioni, per una particolare categoria di utenti, configura le infrastrutture di rete come una speciale tipologia di piattaforme per la pubblicazione e condivisione delle risorse.

Gli aggregatori, definiti anche *Digital Libraries*, garantiscono l'accesso a collezioni di dati conservati in differenti server attraverso un motore di ricerca in grado di recuperare i dati disponibili nelle differenti raccolte. Per integrare gli archivi e renderli disponibili tramite un singolo servizio di rete, è necessario che i dati contengano metadati codificati secondo schemi standard che assicurino l'interoperabilità; specifici protocolli di *harvesting*, tra i quali l'OAI-PMH (*supra* 109), raccolgono i metadati strutturati e li combinano e diffondono in forma di Open Access. Gli aggregatori includono differenti documenti digitali, tra cui testo, immagini, audio, video, ma le risorse restano presso il *data-server* o repository che li ospita; la funzionalità degli aggregatori è chiaramente subordinata alla tipologia di *data-model* o di schema che viene adoperato per rappresentarli. Per assicurare la raccolta dei propri dati, il ricercatore deve soltanto progettare ed implementare un *mapping* tra il suo schema locale ed uno standard (per esempio, il Dublin Core) e scegliere la licenza più idonea per il riuso dei metadati e della risorsa digitale condivisa.

Il più grande portale web europeo che aggrega dati digitali è EUROPEANA lanciato dalla Commissione Europea nel 2008 per condividere collezioni digitali relative al patrimonio culturale. Il progetto nasceva per dare una risposta alle esigenze di un pubblico sempre più familiare con i contenuti Web. La Commissione, nell'ambito del programma *eContentplus*, aveva proposto l'implementazione di un portale in grado di raccogliere milioni di oggetti culturali (libri, dipinti, film, documenti d'archivio, ecc.). Gli utenti avrebbe così avuto l'opportunità di esplorare in un ambiente Web multiculturale e multilinguistico, tutte le ricchezze del patrimonio culturale e scientifico europeo dalla preistoria ai giorni nostri. Il progetto consente un accesso integrato agli oggetti digitali forniti da numerose organizzazioni e istituzioni dell'Unione Europea. Lanciata

originariamente con il nome di EDLnet e con l'obiettivo di dare accesso a oltre 10 milioni di oggetti digitali entro il 2010, EUROPEANA è nata come prototipo di un servizio interdisciplinare incentrato sull'utente. Dopo lo sviluppo della prima interfaccia il portale è diventato pienamente operativo nel febbraio 2009 con il nome di EUROPEANA versione 1.0.

Considerato che uno degli obiettivi centrali di EUROPEANA consisteva nell'accesso al patrimonio digitale europeo, gli sforzi per la sua implementazione sono stati concentrati su due linee: aumentare il numero di documenti disponibili e sviluppare protocolli specifici per consentire la pubblicazione e il recupero dei dati. Mentre la Commissione Europea ha lanciato specifici bandi con l'obiettivo di contribuire al popolamento di EUROPEANA con nuove collezioni, archivi e mappe geografiche, gli sviluppatori hanno progettato un modello dati che fosse abbastanza rappresentativo da includere differenti tipologie di risorse e nuovi strumenti e servizi per ampliare le funzionalità del motore di ricerca.

EUROPEANA è caratterizzata da un accesso decentralizzato in base al quale le diverse categorie di documenti sono ospitati sulle reti delle istituzioni culturali che forniscono i contenuti. Poiché ogni organizzazione culturale (biblioteche, musei, archivi, ecc.) ha catalogato le proprie risorse ricorrendo a schemi locali, per aggregare i contenuti, EUROPEANA ha messo a punto un suo schema di metadati, denominato EDM in grado di raccogliere e indicizzare le descrizioni associate agli oggetti digitali. Lo schema inizialmente scelto per l'aggregazione delle collezioni costituiva un profilo applicativo del più diffuso Dublin Core. Il set di metadati ESE adottava un approccio minimalista e generico alla descrizione e integrazione dei differenti tipi di contenuto.

Con lo sviluppo del portale e l'arricchimento delle raccolte digitali sono sorte nuove esigenze che hanno favorito la revisione di ESE e alla creazione nel 2009 del nuovo modello EDM progettato per sfruttare i progressi tecnologici di Internet, soprattutto nel campo del Web Semantico. Il mapping tra ESE e EDM ha assicurato la migrazione dei vecchi dati alle nuove entità maggiormente aderenti ai principi delle ricerche semantiche. L'EDM consentiva una più sofisticata rappresentazione dei dati garantendo, nello stesso, tempo, collegamenti tra gli oggetti anche

attraverso l'impiego dei LOD.

La pubblicazione di archivi e collezioni in EUROPEANA assicura una ribalta internazionale e la circolazione, la diffusione e la conservazione dei dati digitali appartenenti alle istituzioni culturali e di ricerca europee. L'aggregazione di collezioni appartenenti a differenti paesi e tradizioni incoraggia la costruzione di percorsi che possono essere su base cronologica, nazionale, di contenuto o relativa a particolari tipologie di oggetti. Attualmente, il canale tematico della Archeologia Europea comprende oltre 2.000.000 oggetti che possono essere organizzati per la creazione di mostre virtuali; il canale comprende immagini, testi, oggetti 3D, audio e video riferiti a manoscritti, mappe, piante, fotografie, giornali e telegiornali. Uno specifico progetto, implementato tra il 2019 e il 2020, è stato ideato per aumentare la visibilità di un ampio corpus di contenuti digitali relativi al ricco patrimonio europeo composto da monumenti archeologici, edifici storici, paesaggi culturali e manufatti. Il progetto, denominato *European Archaeology*, espone oltre 374.000 documenti archeologici provenienti da 28 partner Europei e ha coinvolto il pubblico nella mostra online *Uncovering Hidden Stories. An Introduction to European Archaeology*.

Nell'ambito del progetto sono stati resi accessibili nel 2020 anche i contenuti conservati nel repository della rivista *Archeologia e Calcolatori*; si tratta di 930 risorse per un totale di 15.000 pagine. Inoltre, sempre nel 2020, è stato creato uno specifico servizio che consente di arricchire i metadati forniti a EUROPEANA attraverso un *tool* che supporta l'utente nella mappatura di oggetti archeologici utilizzando descrizioni normalizzate.

Un'altra iniziativa che reimpiega dati di EUROPEANA è EAGLE, una rete di best-practices, che raccoglie, in un unico database facilmente consultabile, oltre 1,5 milioni di articoli, attualmente sparsi in numerosi paesi dell'UE, oltre che nel Mediterraneo orientale e meridionale. EAGLE colleziona la gran parte delle iscrizioni superstiti del mondo greco-romano, complete di descrizioni e informazioni essenziali, nonché per tutte le più importanti, di una traduzione in inglese. Attraverso il motore di ricerca si accede al vasto database epigrafico di EAGLE. Oltre le tradizionali modalità di ricerca il motore consente anche di caricare un'immagine e cercare iscrizioni simili. Le ricerche possono esse-

re effettuate in sette lingue ricorrendo ad un dizionario multilingua messo a disposizione per semplificare l'allineamento dei metadati. Il vocabolario è stato creato riunendo le terminologie ed allineandole successivamente, laddove possibile, ad alcuni thesauri standard, tra cui l'A-AT, mentre è in fase di progettazione l'allineamento con il LIMC.

EAGLE ospita circa 350.000 testi disambiguati e oltre 455.000 edizioni critiche. La maggior parte dei documenti epigrafici disponibili è stata disambiguata. Alcuni progetti che aderiscono a EAGLE sono terminati, altri sono in continuo aggiornamento, fornendo nuove iscrizioni e modificando precedenti edizioni. Tutti i dati vengono aggregati due volte all'anno. I testi sono in maggioranza codificati secondo lo standard EpiDoc. Sfruttando i contenuti di EUROPEANA, EAGLE ha realizzato una mostra online dedicata all'epigrafia e composta da sale virtuali nelle quali il pubblico viene introdotto agli studi sui documenti epigrafici e sul contenuto. Dalla sezione tecnologie digitali per l'epigrafia si può accedere al museo virtuale dove sono presentati, in un ambiente espositivo ricostruito digitalmente, pannelli, carte geografiche, oggetti e modelli 3D di iscrizioni. Oltre al museo virtuale è stata predisposta anche una app per dispositivi mobili che consente a un utente di ottenere informazioni su un'iscrizione visibile scattando una foto con un cellulare o tablet e inviandola al portale EAGLE. L'app dispone di due modalità di riconoscimento: la prima, *Exact Match Mode*, cercherà informazioni associate all'immagine, la seconda fornirà un elenco di iscrizioni simili all'immagine inviata al server.



## Capitolo 4

### L'INFORMAZIONE SPAZIALE

#### 4.1. I dati spaziali

L'informazione spaziale, sia essa geografica o relativa alla descrizione di manufatti rinvenuti in un dato contesto, ha un ruolo centrale nella ricerca archeologia. Ricognizioni di superficie, scavi, prospezioni geologiche e geomorfologiche hanno come oggetto la raccolta di dati che contengono, tra l'altro, informazioni sulla localizzazione dei rinvenimenti, sulle aree di produzione e sui luoghi di conservazione. I dataset registrano dati spaziali che possono successivamente essere adoperati per complesse analisi territoriali o per più semplici carte distributive e/o tematiche.

Lo spazio, con la sua descrizione, misurazione e, infine, analisi, ha assunto progressivamente un peso determinante nella metodologia archeologica. Lo studio delle relazioni spaziali e, soprattutto, delle trasformazioni nell'organizzazione dello spazio, costituiscono un obiettivo cruciale dell'archeologia come disciplina storica, al punto che possiamo legittimamente parlare di archeologia spaziale. La componente dello spazio e della spazialità hanno sempre ricevuto una particolare enfasi nella analisi dei resti antropici e naturali che costituiscono la base dell'indagine archeologica sul terreno. Non soltanto è importante evidenziare l'uso dello spazio nelle tracce materiali, ma è anche rilevante registrare le forme della sua rappresentazione e manipolazione nel mondo figurato antico.

Prima delle riflessioni disciplinari maturate nell'ambito della New Archaeology, l'interesse degli studiosi verso la mappa era indirizzato soprattutto alla identificazione di quegli aspetti in grado di fornire informazioni su lunghe scale temporali, grandi temi e tendenze. Dopo la metà del secolo scorso le nozioni di tempo e spazio vennero esaminate contestualmente senza privilegiare l'una sull'altra come era stato nell'approccio storico-culturale.

In un recente contributo, dedicato all'analisi spaziale in archeologia, alcuni autori (Gillings, Hacıgüzeller, Lock 2020) hanno giustamente se-



gnalato le principali interazioni uomo-spazio. La prima è il movimento. La mobilità non soltanto rappresenta il viaggio in uno spazio fisico definito, ma, nello stesso tempo e su un piano più strettamente antropologico, descrive la relazione, talvolta simbolica, che l'uomo ha con il territorio circostante. Flussi e reti di connessioni forniscono informazioni sul paesaggio materiale e sulle forme del suo sfruttamento, anche immateriale, attraverso modelli economici, sociali, culturali e religiosi. Gli studiosi hanno anche evidenziato come, dal punto di vista della ricerca archeologica sul campo, l'identificazione dello spazio, degli assemblaggi, delle forme e della quantità e densità dei resti antropici consenta di ricostruire la presenza e lo sviluppo di pratiche quotidiane, nonché le interazioni con altri luoghi abitati e destinati a specifiche funzioni. Non soltanto le presenze suggeriscono ipotesi per comprendere il significato dello spazio; anche le assenze e i vuoti possono avere un loro preciso interesse archeologico e non essere fenomeni casuali. Lowenborg (2018, 37) sottolinea che

The saying 'the absence of evidence is not the evidence of absence' certainly applies to archaeology.

Infine, un elemento, altrettanto importante, anche per gli aspetti connessi alla strutturazione dell'informazione digitale, è quello della rappresentazione grafica dello spazio. Mappe, piante e modelli distributivi sono strumenti costruiti e adoperati dagli studiosi per concettualizzare, sintetizzare e spiegare l'informazione spaziale. La carta rappresenta anche il modo con il quale l'archeologo comprende lo spazio, e, pertanto, la progettazione di una mappa non appare un processo semplice e lineare, ma comporta decisioni che possono avere conseguenze sulla lettura e sul riuso delle carte.

Analizzare le interazioni uomo-paesaggio attraverso la cultura materiale comporta una definizione a priori della nozione di spazio che non può essere un contenitore neutro e universale dell'esistenza umana, come nella impostazione della New Archaeology. La critica post-processualista e i più recenti studi antropologici e etnografici hanno profondamente criticato l'idea occidentale e moderna di uno spazio statico, preferendo una concettualizzazione dinamica nella quale emergono luoghi culturali e relazioni sociali tra individui. Negli anni '90 vengono messe in discussione

le basi teoriche dell'archeologia spaziale di ispirazione processualista improntata al principio che una distribuzione di punti su una mappa possa riflettere un processo spaziale e una causalità da identificare (Gillings, Hacıgüzeller, Lock 2020). Tuttavia, dopo la modellistica spaziale elaborata negli anni '70 (Hodder, Orton 1976), la Post-Processual Archeology non sarà in grado di sostituire quell'approccio teorico, basato su strumenti anche statistici, con una nuova metodologia formale altrettanto solida e chiara; il rifiuto dell'analisi spaziale gemoetrica condurrebbe ad una narrazione testuale dello spazio in continua evoluzione e interpretazione.

Il paesaggio non è uniforme e i modi per descriverlo e rappresentarlo non sono omogenei; addirittura, nella impostazione post-processualista non è nemmeno utile proiettarli in forma di mappa.

Nella dialettica sulla definizione dello spazio e, soprattutto, della sua concettualizzazione e funzione quantitativa o qualitativa, statistica o narrativa, i metodi computazionali hanno influenzato il recupero di un approccio metodologico e teorico allo studio dei dati a connotazione spaziale ispirato prevalentemente dai formalismi della New Archaeology.

La fase attuale dell'archeologia spaziale è caratterizzata dall'uso delle tecnologie e, in particolare, di sistemi informativi il cui potenziale di impiego in campo archeologico è stato riconosciuto oltre 30 anni fa. La possibilità di combinare differenti fonti informative, a scala diversa, conferisce al GIS una straordinaria versatilità applicativa che rende questo approccio utile in contesti e obiettivi differenti: il GIS è funzionale alla gestione, integrazione e visualizzazione di complessi archivi georeferenziati e al trattamento dell'informazione nell'ambito dell'analisi spaziale (Allen, Green, Zubrow 1990).

Sia che si tratti di elaborazioni basate su complesse procedure statistiche o di esplorazioni sul campo che mirano alla raccolta di superficie, allo scavo o a ricerche ambientali, i sistemi informativi hanno avuto un forte impatto sulla natura dell'indagine archeologica a livello territoriale. La struttura formale del GIS privilegia una certa rigida prospettiva, spesso geometrica, delle relazioni spaziali spingendo il ricercatore ad abbracciare una visione deterministica dell'ambiente. Tuttavia, non mancano atteggiamenti ispirati all'adozione di una narrativa meno statica e più umanistica che consente di esaminare i dati con uno sguardo diverso.

Il GIS può aiutare il ricercatore nello studio della percezione e della rappresentazione simbolica del paesaggio e a comprendere le modalità

attraverso le quali l'uomo trasforma la natura in immagini culturali e/o religiosi. La potenza di calcolo computazionale non deve sostituirsi ai principi e ai metodi della ricerca archeologica; essa, al contrario può suggerire nuove e originali domande contribuendo con inedite sfide allo sviluppo di particolari orizzonti di studio, particolarmente nel campo dell'Archeologia dei Paesaggi e della Topografia Antica.

L'introduzione dei GIS in archeologia, almeno a partire dalla metà degli anni '80, ha dato un grosso impulso agli studi sull'analisi territoriale, animando un nuovo dibattito metodologico e teorico sullo spazio interrotto dal relativismo della Post-Processual Archaeology. L'ampia diffusione dei GIS, è, oggi, favorita dalla comparsa di efficienti software open source che hanno semplificato la realizzazione di analisi distributive ed inferenziali; l'integrazione delle risorse cartografiche, di natura non direttamente archeologica (carte geologiche, pedologiche, orografiche, etc.), ha contribuito alla nascita di nuovi metodi di ricerca indirizzati principalmente alle analisi multidimensionali con una specifica attenzione alla valutazione della terza dimensione, fino ad allora poco considerata. Attualmente, è difficile immaginare una ricerca territoriale e sul paesaggio antico portata avanti senza l'ausilio di sistemi GIS, divenuti essenziali per qualsiasi tipologia di esplorazione, analisi, interpretazione e, infine, gestione dei dati spaziali.

L'uso dei GIS e la sua accettazione recente presso la stragrande maggioranza degli archeologi, ha suscitato una riflessione metodologica a livello più elevato e astratto e cioè se le potenti funzioni analitiche del software abbiano alla fine sostituito qualsiasi valutazione sul valore della mappa nella ricerca storico-archeologica. La relativa semplicità e velocità con la quale si costruisce una mappa solleva atteggiamenti controversi. La possibilità, pressoché illimitata, di produrre carte tematiche che possono essere tra di loro confrontate, costituisce certamente un vantaggio offerto dalla tecnologia, ma, come è stato da più parti segnalato, il rischio di una tale semplificazione potrebbe risiedere in una minore attenzione dei ricercatori agli aspetti teorici contenuti nel software e nella stessa metodologia archeologica.

J. Conolly e M. Lake (2006), ad esempio, hanno messo in guardia gli studiosi sul rischio che

...many current GIS software packages require just a few mouse clicks to

create an elevation model from a set of contour lines....

G. Lock (2009) ha sottolineato, qualche anno dopo, che

A force within this “routinisation”, and perhaps a major force, is an element of technological determinism – both visibility analysis and movement studies are an integral part of many commercial GIS software packages, almost push-button solutions.

Il rischio è che ogni valutazione sul trattamento dell'informazione a connotazione spaziale sia totalmente lasciata al GIS e alle sue funzioni spaziali con una conseguente sottovalutazione della teoria informatica e archeologica. Le attrattive della tecnologia sono considerate, spesso, superiori a qualsiasi restrizione o svantaggio (Gillings, Hacıgüzeller, Lock 2020).

Il GIS ha contribuito alla riscoperta del potere e della funzione scientifica della mappa nella identificazione di fenomeni che hanno proprietà spaziali, ma questo processo di particolare trattamento del dato solleva importanti interrogativi sulla necessità di possedere una metodologia mista, in parte archeologica, informatica e statistica. Occorre, quindi, sviluppare una maggiore consapevolezza e una reale competenza sul significato dell'analisi spaziali e, specificamente, della funzione dello spazio in archeologia, soprattutto nella prospettiva di una sempre maggiore integrazione di dataset con altri software che elaborano informazioni geo-riferite a livello architettonico, come il BIM (*supra* 31), e modelli simulativi dinamici, come gli ABM (Davies *et al.* 2019). La combinazione di differenti software connessi allo sfruttamento dell'informazione spaziale fornirà in futuro maggiori dati per l'interpretazione e la comprensione dei fenomeni sociali del passato.

#### 4.2. Standard per l'informazione geografica

L'informazione spaziale ha un ruolo importante nella nostra vita quotidiana essendo immersi in un mondo di dati georiferiti generati da una ampia varietà di dispositivi con finalità diverse. Le risorse geografiche hanno una funzione determinante in applicazioni sia di tipo civile che militare; ciò ha spinto al riconoscimento dell'importanza dello scambio di dati e, quindi, della necessità di mettere a punto standard di formato, contenuto e interscambio che facilitino l'accesso, la condivisione ed il riuso dei dataset georiferiti. Un tale atteggiamento ha contribuito ad accrescere l'importanza

dei dati spaziali e, nello stesso tempo, la loro produzione e integrazione.

Allo scopo di gestire e combinare in modo ottimale la varietà delle risorse spaziali disponibili sono state sviluppate specifiche piattaforme che adoperano gli standard geografici proposti negli anni da numerose istituzioni internazionali. Internet ha offerto una base fondamentale per la creazione di SDI finalizzate ad offrire una molteplicità di dati, soluzioni e servizi Web per applicazioni basate su contenuti di tipo geografico.

Tra i vantaggi introdotti dallo sviluppo del Web 2.0 va, senza dubbio, segnalato l'incremento delle applicazioni online che hanno avuto il merito di avvicinare l'utente a risorse spesso remote o inaccessibili. Tra gli oggetti digitali disponibili in rete, una attenzione particolare hanno ricevuto i dati geografici utilizzati per una straordinaria varietà di usi: dalle ricerche demografiche a quelle statistico-commerciali e di servizi e/o di pianificazione e progettazione territoriale. L'informazione a connotazione spaziale è stata veicolata soprattutto attraverso l'implementazione di sistemi cartografici che offrono, con accesso aperto o a pagamento, una ingente quantità di materiali, grezzi o elaborati, prodotti in particolare dalle Pubbliche Amministrazioni e disponibili per una pluralità di usi. I sistemi WebGIS consentono di visualizzare, selezionare ed aggregare dati geografici e, in alcune circostanze, grazie all'impiego di specifici protocolli, di scaricare e/o riutilizzare i contenuti resi disponibili nei portali geografici. L'utente non ha alcuna possibilità di modificare i contenuti predeterminati dall'amministratore del WebGIS o di aggiungerne di nuovi, ma la semplice consultazione consente l'esecuzione di analisi territoriali, talvolta complesse. L'interazione è limitata alla possibilità di accedere ai dati e di visualizzarli combinando differenti strati informativi. Alcuni sistemi mettono a disposizione dell'utente funzioni spaziali elementari come, ad esempio, il filtraggio dei dati alfanumerici, oppure strumenti per la misurazione di distanze, lunghezze e aree o per l'identificazione delle coordinate geografiche delle risorse selezionate.

Il successo dei WebGIS è stato determinato, in modo consistente, dallo sviluppo di una industria dell'informazione geografica fondata sull'acquisizione di dati provenienti da numerose tipologie di sensori, dall'introduzione di molteplici tecnologie spaziali e dalla realizzazione di programmi per la navigazione stradale. Software come OpenStreet-

Map, Google Maps e Google Earth hanno concorso alla espansione di un mercato del dato spaziale e alla costruzione di applicazioni di tipo *mash-up* sviluppate per combinare le risorse geografiche esistenti.

Il settore geo-spaziale ha contribuito in modo efficace all'affermazione di una rete partecipativa e collaborativa, mettendo a disposizione del Web 2.0 un bagaglio di competenze sviluppate nel campo dei sistemi open source con la comparsa, già negli anni '80, di GRASS, il primo software geografico aperto, e, negli anni '90, di MapServer, il primo server geografico per il *Web-mapping*. Dal successo e dalla diffusione di questi strumenti per il trattamento e la condivisione dell'informazione geografica è emersa una comunità di sviluppatori e utenti che ha dato vita a due istituzioni determinanti per il consolidamento e il supporto dell'informazione spaziale: l'OGC nel 1994 e l'OSGeo nel 2006.

L'OGC negli anni ha svolto un importante compito, facilitando il dialogo tra quanti acquisiscono ed elaborano l'informazione geografica attraverso la condivisione di pratiche, la creazione di interfacce, standard e protocolli per garantire l'interoperabilità dei dati; l'obiettivo è rendere l'informazione geografica parte integrante di una struttura informativa più ampia e più generale attraverso il rilascio di standard di codifica del dato spaziale e linee guida per il loro corretto utilizzo. L'OGC ha finora pubblicato oltre 70 standard aperti, liberamente disponibili e senza costi di licenza, relativi alle specifiche di formato per la pubblicazione e la fruizione dei dati; anche gli aspetti tecnici, cioè topografici, connessi all'acquisizione dell'informazione geografica, rientrano tra le attività del consorzio. Tra gli standard rilasciati occorre ricordare:

a) I protocolli per lo scambio dei dati per garantire l'interoperabilità sintattica, tra cui:

- il WMS che definisce una semplice interfaccia HTTP per richiedere immagini di mappe da uno o più server distribuiti in Internet;
- il WFS che, come il servizio WMS richiede oggetti geografici in rete con la differenza che non vengono restituite immagini, bensì le descrizioni degli oggetti spaziali (coordinate spaziali e attributi alfanumerici);
- il WCS che permette richieste complesse che forniscono risposte che includono tutte le informazioni e le caratteristiche come sono all'origi-

- ne (anziché le immagini) in modo da consentire nuove elaborazioni;
- il WMC che specifica come un gruppo di mappe provenienti da uno o più WMS può essere importato su qualsiasi piattaforma.

b) Gli schemi di codifica per assicurare l'interoperabilità semantica tra cui i principali sono GML e il KML.

Grazie al lavoro svolto dall'OGC, un produttore di sensori, di apparecchi di geolocalizzazione e, in generale, di tecnologie spaziali può creare dati georiferiti in un qualsiasi formato, aperto, chiuso o di tipo proprietario, che può essere facilmente mappato su uno dei tanti standard rilasciati dal Consorzio.

Nel settore della standardizzazione dei geo-dati per superare l'eterogeneità e sviluppare architetture collaborative, va ricordata anche l'attività dell'ISO, portata avanti in particolare dalla sezione *Technical Committee 211 Geographic Information/Geomatics*, che ha lavorato in stretta collaborazione con l'OGC. Dal confronto tra i due enti internazionali è nato il gruppo degli standard internazionali ISO19100 che riprende le specifiche dell'OGC e raccoglie oltre 100 standard che forniscono specifiche per i GIS, i metodi, gli strumenti e i servizi per la gestione dei dati spaziali compresa l'acquisizione, l'elaborazione, l'analisi, l'accesso, la presentazione e il trasferimento nonché lo scambio. Lo standard 19101, aggiornato al 2020, definisce il modello di riferimento per la standardizzazione nel campo dell'informazione geografica, descrivendo la nozione di interoperabilità e fornendo i criteri fondamentali attraverso i quali viene realizzata la standardizzazione; sebbene sia strutturato nel contesto delle tecnologie dell'informazione l'ISO 19101 è indipendente da qualsiasi metodo di sviluppo delle applicazioni o approccio all'implementazione di specifici programmi di gestione del dato geografico.

Tra le iniziative a favore di una più ampia integrazione delle risorse geografiche si deve includere anche la direttiva europea INSPIRE entrata in vigore nel 2007 e recepita in Italia nel 2010.

Originariamente il progetto INSPIRE era limitato alle modalità di ricerca di dati geo-riferiti di tipo territoriale allo scopo di contrastare eventuali fenomeni di disastri ambientali determinatisi oltre gli specifici confini nazionali. Successivamente il focus della direttiva si è allargato, compren-

dendo l'esigenza di rendere disponibili in un formato coerente i dataset necessari all'implementazione di politiche comunitarie o di attività nazionali con un particolare impatto sull'ambiente e sul clima. Questi studi si basano su valutazioni che nascono dall'analisi e dalla combinazione di diversi fattori come i dati ambientali, geografici (uso del suolo, confini amministrativi, altitudine, idrologia, reti di trasporto, impianti di produzione, etc.) e geofisici (meteorologia, geologia, etc.). Incrociando i dataset geografici con informazioni socioeconomiche è possibile, ad esempio, sviluppare specifici interventi sulla salute, sull'aria, sulla agricoltura. Allo scopo di gestire e analizzare i dati territoriali, che comprendono tra l'altro misurazioni sulle emissioni, osservazioni sulla biodiversità o dati sulla qualità ambientale, gli Stati Membri hanno progettato e sviluppato infrastrutture di rete mettendo a disposizione documenti geografici, metadati e servizi.

I dati, indicati negli allegati della Direttiva INSPIRE, riguardano i sistemi di coordinate, di griglie geografiche, nomi geografici, unità amministrative, indirizzi, parcelle catastali, reti di trasporto, idrografia, siti protetti (allegato 1); elevazione, copertura del suolo, orto immagini, geologia (allegato 2); unità statistiche, edifici, suolo, utilizzo del territorio e altri dati ambientali (allegato 3). I metadati definiscono le regole per la descrizione dei dati geografici all'interno di una collezione di dati spaziali; l'obiettivo di un buon sistema di metadati è semplificare la ricerca, la localizzazione, la selezione e l'interoperabilità semantica degli archivi. INSPIRE utilizza come standard di riferimento lo schema del Dublin Core.

I servizi di rete disciplinati da INSPIRE seguono le specifiche dell'OGC e dell'ISO TC211 e, in particolare, sono finalizzati all'implementazione dei seguenti servizi:

- ricerche che consentono di identificare i set di dati territoriali e i servizi ad essi relativi in base al contenuto dei metadati;
- consultazione per eseguire operazioni di visualizzazione, navigazione, variazione della scala di visualizzazione, variazione della porzione di territorio inquadrata, sovrapposizione dei set di dati territoriali consultabili, di creazione di legende;
- download dei dati per scaricare copie di set di dati o una parte di essi e, ove fattibile, di accedervi direttamente;
- conversione per trasformare i set di dati territoriali allo scopo di



assicurare lo scambio.

In attuazione della Direttiva, è stata creata nel 2010 una infrastruttura nazionale italiana (INITMA), quale nodo di quella comunitaria, con l'obiettivo di rendere omogenee e condivisibili le informazioni geo-riferite di carattere ambientale. Con lo stesso strumento normativo il vecchio PCN è stato sostituito dal GN, istituito presso il Ministero dell'Ambiente e della Sicurezza Energetica (in precedenza Ministro dell'Ambiente e della Tutela del Territorio e del Mare).

Il GN è il principale punto di accesso nazionale, aperto e accessibile per la visualizzazione e fruizione di ogni informazione territoriale senza limitazioni o barriere sia fisiche che tecnologiche. Il nodo italiano per le finalità della direttiva INSPIRE sostiene l'azione di governo a livello nazionale e territoriale nei processi di pianificazione, monitoraggio e verifica di tutti i procedimenti aventi un impatto a livello territoriale e aggiorna costantemente il patrimonio informativo territoriale ed ambientale a tutti i livelli, incoraggiando le azioni di interazione tra il cittadino, le aziende e la PA. Obiettivo del portale è promuovere la cultura della condivisione del dato territoriale, offrendo servizi accessibili con i protocolli standard WMS, WFS e WCS. GN rende disponibile un software gratuito per l'elaborazione dei dataset spaziali (AdB-Toolbox) e un simulatore tridimensionale del territorio italiano che sovrappone al modello 3D gli strati informativi provenienti dalla banca dati del Geoportale attraverso il protocollo WMS.

Gli obiettivi che hanno ispirato la direttiva europea sulla pubblicazione e condivisione dei dati ambientali non hanno prodotto soltanto l'implementazione del nodo GN. La disponibilità di protocolli, servizi, formati e metadati standard ha stimolato la creazione di numerosi portali regionali e locali che offrono, liberamente o tramite autenticazione, una grande quantità di dati geo-riferiti scaricabili senza particolari conoscenze in un qualsiasi sistema GIS.

La maturità del dominio geo-spaziale è confermata dalla attività di numerosissime istituzioni pubbliche e private, associazioni professionali, settori della ricerca e della formazione che producono e usano dati geografici distribuiti per la maggior parte con licenze aperte e quindi riutilizzabili in una pluralità di applicazioni.

Una ultima annotazione, relativa agli standard utilizzati in ambito

spaziale, deve essere riservata alle convenzioni di recente assegnate alla nomenclatura dei sistemi geografici e cartografici. Le procedure adoperate per la geolocalizzazione sono il risultato di complessi calcoli matematici e geometrici che prendono in considerazione differenti parametri per correggere le deformazioni introdotte dalla proiezione prospettica di una informazione, originariamente posizionata su una superficie curvilinea terrestre irregolare, su un reticolato bidimensionale. Per costruire una carta possono essere adoperati vari sistemi di correzione che prendono il nome di proiezioni geografiche. Questi riferimenti hanno origine in un datum che consente di approssimare la forma ellissoidale e asimmetrica della terra ad una superficie sferica. Allo scopo di ridurre gli errori introdotti dalla proiezione cartografica sono stati messi a punto numerosi sistemi di riferimento generali e locali con l'obiettivo di determinare la localizzazione accurata di un punto in una qualsiasi parte del globo. Esistendo moltissimi sistemi cartografici, la trasformazione delle coordinate tra le varie proiezioni non è un'operazione semplice e del tutto priva di errori, poiché i vari sistemi possono preservare gli angoli, le distanze o le aree reali.

Per districarsi nel mare dei sistemi di coordinate geografiche esistenti in tutto il mondo, è stato messo a punto un servizio online sviluppato inizialmente dall'EPSG e successivamente mantenuto dall'IOGP. Il servizio è costituito da un database che raccoglie 6.293 differenti sistemi di coordinate cartografiche e geografiche locali, regionali e nazionali con i relativi attributi e l'area di riferimento. Ogni sistema ha un identificativo numerico preceduto dal prefisso EPSG che l'utente può cercare nel database partendo dal testo o identificando un punto o un'area su una mappa; il servizio consente anche di convertire le coordinate tra riferimenti cartografici differenti sulla base di specifici algoritmi di trasformazione. L'IOGP ha provveduto all'aggiornamento del modello dati EPSG rendendolo conforme allo standard ISO19111:2019. Numerose applicazioni, in rete o desktop, accettano il riferimento geografico identificato dal codice EPSG facilitando in tal modo l'integrazione e lo scambio dei dataset spaziali appartenenti a sistemi di coordinate differenti.

L'implementazione dei principi FAIR ai dati geo-spaziali non è una pratica tuttora molto comune nel dominio geografico, sebbene il largo uso di standard OGC e ISO con vocabolari condivisi spinga verso il riu-

so, in una forma ancora non ottimale, delle risorse geo-riferite. In generale la comunità di riferimento dei produttori di dati geografici fa largo uso di schemi di metadati comprensibili dall'uomo, ma scarsamente utilizzabili da agenti semantici, mentre la stessa terminologia adoperata è talvolta eterogenea, troppo settoriale e di rado uniforme.

Nel corso degli anni lo sviluppo di standard, protocolli, metadati, servizi e dizionari ha avuto il merito di migliorare le modalità per identificare, accedere e riutilizzare i dati geografici presenti in rete, mentre l'adozione di linee guida ha contribuito ad allargare la comunità dei potenziali utenti facilitando l'accesso anche dei meno esperti al dominio dei dati georiferiti. Numerose iniziative in corso sono indirizzate all'arricchimento del contenuto dell'informazione spaziale e al suo riuso in applicazioni e soluzioni non esclusivamente destinate alla gestione dell'informazione geografica.

#### 4.3. Il rilievo digitale in archeologia

La semplificazione e lo sviluppo delle tecniche di rilievo digitale hanno modificato l'approccio degli archeologi alla documentazione grafica; con minime competenze sul piano del disegno e del rilievo architettonico, i ricercatori possono produrre un'ampia serie di dati multiformi e multidimensionali. La disponibilità di strumentazione di misurazione con basse curve di apprendimento ha facilitato la sostituzione nei cantieri di scavo di architetti, disegnatori e rilevatori professionisti con una generazione di archeologi esperti di nuove tecnologie e rappresentazione grafica.

I vantaggi di una nuova metodologia digitale per il disegno risiedono prevalentemente nella struttura informatica delle apparecchiature che non richiedono una particolare conoscenza delle problematiche connesse ad un corretto rilievo, sia esso topografico o architettonico. Ciò ha reso il disegno digitale una pratica ordinaria, quasi banale, che non necessita di particolari competenze.

La traccia di una minore consapevolezza dell'importanza di possedere una adeguata teoria per utilizzare correttamente strumenti digitali per il rilievo, emerge dalla recente bibliografia, caratterizzata in prevalenza da interventi ai convegni che mostrano le repliche digitali ottenute senza alcuna attenzione al tema della affidabilità geometrica della re-

stituzione. Troppo spesso, si registra l'assenza di una considerazione sul valore delle tecnologie digitali per il rilievo e il loro corretto uso nei contesti di scavo, mentre la necessaria riflessione sulle modalità di acquisizione del dato spaziale è sostituita dalla descrizione della quantità di foto acquisite o di tecnologie impiegate. L'equazione più foto = modello risoluto e, quindi, migliore e più preciso, sintetizza un atteggiamento acritico nel quale la quantità di bit processati supplisce alla mancanza di una adeguata riflessione metodologica sul prodotto finale.

Non è semplice comprendere le cause del rifiuto verso ogni forma di valutazione dell'importanza del metodo, soprattutto quando non si ha conoscenza degli elementi di misura e di processamento che si celano in uno strumento. L'analisi della documentazione grafica ha un valore cruciale nella costruzione del discorso archeologico e nella successiva interpretazione storica e, quindi, ci si attenderebbe una maggiore consapevolezza sull'esigenza di approfondire l'intero processo di esecuzione e restituzione del rilievo digitale; probabilmente le implicazioni metodologiche connesse agli strumenti digitali di misurazione sono derubricate a livello di semplice aggiornamento tecnologico.

L'emergere di una trasformazione disciplinare sostenuta dalle innovazioni digitali impone, anche al mondo della ricerca archeologica, di guardare in modo diverso, più partecipativo e collaborativo, allo studio delle civiltà società passate e, quindi, alle modalità di rappresentazione e mediazione culturale. Rispetto alle tradizionali procedure analogiche, il rilievo digitale è certamente più rapido e, per certi aspetti, più preciso e con maggiori dettagli, ma in una società che si trasforma e che mette al centro della sua riflessione scientifica concetti come credibilità, proprietà, diritti e riuso, è necessario che il ricercatore ripensi ai sistemi di lavoro, alle modalità di interazione con altri studiosi e a nuove forme operative all'interno di un mondo digitale virtuale. L'analisi del metodo necessariamente si intreccia con alcune questioni che più direttamente si connettono all'innovazione, in primo luogo, nel campo della gestione dell'informazioni digitale.

I computer hanno una potenza di calcolo inimmaginabile fino a qualche anno fa. È possibile sviluppare ambienti 3D e renderli facilmente accessibili in rete superando i tradizionali confini della documentazione grafica bidimensionale. Il ricercatore è immerso in un mondo di-

gitale nel quale l'interazione uomo-macchina diventa il punto cruciale per la formulazione di nuove ipotesi, mentre il campo della visualizzazione scientifica dell'informazione estende le facoltà cognitive dello studioso. Nonostante la disponibilità pressoché illimitata di dati digitali rilevati con le più moderne e sofisticate apparecchiature, anche nel campo del rilievo si è imposta una riflessione connessa all'esigenza di condividere dati geometricamente corretti e affidabili e la cui tecnologia di acquisizione e processamento adoperata sia comprensibile per un utente umano come per un agente virtuale.

Analogamente ai dati testuali, il trattamento, la cura e l'archiviazione richiedono una riconsiderazione in funzione della adesione ai principi della Scienza Aperta e, quindi, della accessibilità dell'informazione. Anche per quanto riguarda l'impiego di dati pregressi di tipo geometrico e architettonico occorre una rilettura della documentazione e non una semplice conversione analogica-digitale. Senza una conoscenza, talvolta di base, sui processi di elaborazione di quei documenti grafici, determinanti per la lettura di piccoli e grandi complessi archeologici, il rischio è che si rendano disponibili per il riuso archivi che possono disorientare lo stesso ricercatore.

Gli archeologi considerano la conoscenza archeologica il risultato di un processo il cui obiettivo non è riportare alla luce una realtà fisica sepolta, ma fornire una intellegibile traduzione grafica e/o testuale anche se, spesso parziale e frammentaria, di ciò che viene distrutto durante lo scavo. Il rilievo non costituirebbe la testimonianza visiva di un ragionamento, ma sarebbe il risultato di una forza creativa che rende visibili eventi ed azioni del passato.

Sul ruolo che le tecnologie visive hanno nella costruzione del pensiero archeologico esiste un lungo dibattito (Opghaffter 2021). L'introduzione delle tecniche digitali nella acquisizione e registrazione della documentazione archeologica ha contribuito ad ampliare il confronto tra gli studiosi. Come abbiamo cercato di dimostrare, il perfezionamento in ambito archeologico dei sistemi di rilevazione si è realizzato parallelamente ad una più generale evoluzione tecnologica della società, al punto che possiamo affermare che la stessa metodologia archeologica sia, oggi, il prodotto di trasformazioni tecnologiche, di costanti contatti interdisciplinari con altri domini e, infine, di un profondo e generale muta-

mento dei principi più generali delle scienze.

Lesigenza di produrre analitici rilievi archeologici emerge con i primi scavi del generale Pitt-Rivers alla fine dell'800 al quale si attribuisce la celebre massima «*Describe your illustration dot not illustrate your descriptions*» (Piggott 1965, 174); la comparsa di una componente grafica specializzata si collega alla nascita di una nuova disciplina caratterizzata da un linguaggio visuale e da un testo che ha la funzione di spiegare l'illustrazione e non viceversa. Allo scopo di sottolineare l'importanza di un accurato disegno delle sezioni stratigrafiche, Wheeler (1954, 59) adoperò la singolare analogia di comparare la successione degli strati con una sequenza di lettere non suddivise in parole, in altri termini

...the unmeaning monotony of a sentence spoken without inflection...

Una buona rappresentazione grafica presuppone la comprensione delle forme e della funzione dell'oggetto da riprodurre; il disegno richiede una capacità di vedere e dare una forma, scientificamente corretta, ad un ritrovamento, sebbene, talvolta, rinvenuto in condizioni frammentarie (Medri 2008). Per Bradley (1997) l'attività sul terreno è un processo *in itinere* durante il quale l'archeologo apprende la stratigrafia, identifica gli eventi e ne fornisce una rappresentazione grafica.

L'archeologia da campo diviene una *visual discipline* caratterizzata da una capacità di riconoscere e disegnare tracce di attività antiche che si affina con il tempo. Competenze diagnostiche e grafiche acquisiscono un valore temporale. Senza dubbio, le esperienze professionali condizionano il ricercatore nell'identificare oggetti archeologici e, soprattutto, nel disegnarli accuratamente.

L'introduzione alla metà dell'800 della fotografia, considerata una più oggettiva tecnologia di registrazione dei dati, non ha provocato la rimozione del disegno; al contrario, la fotografia è divenuta un documento visivo complementare al disegno, benché quest'ultimo richieda un intervento umano nella traduzione della materialità dell'oggetto in codifica visuale.

L'innovazione digitale ha modificato il modo di identificare e registrare i dati sul campo. Da un lato l'evoluzione tecnologica mette a disposizione dei ricercatori strumentazioni di facile impiego e trasporto, dall'altro la conoscenza gestita dai dispositivi e sensori per l'acquisizione

dei dati, rende quasi del tutto ininfluyente una competenza sui principi di base del disegno, della fotogrammetria e della topografia. La macchina sostituisce l'uomo nelle attività di registrazione sollevando, nella maggior parte dei casi, il ricercatore dalla esigenza di comprenderne il funzionamento. Dati digitali ad alta risoluzione sono acquisiti da minuscoli dispositivi o da attrezzature più complesse, elaborati e condivisi rapidamente. I laser scanner e i droni stanno sostituendo le cazzuole e le carriole, a livello di utensili di base, per ogni missione archeologica sul campo.

L'ampia diffusione di strumenti digitali per il rilievo non è stata supportata da un incremento nelle competenze dei nuovi rilevatori digitali; l'introduzione delle tecnologie spaziali ha esteso la tipologia dei documenti acquisibili sul campo sollevando, nello stesso tempo, numerose criticità sull'uso e riuso, spesso, acritico dei dati. Al centro del dibattito non è necessario porre l'interrogativo su quale tecnologia produca dati più oggettivi, ma in che termini l'uomo, con le sue conoscenze ed esperienze possa migliorare la qualità della documentazione grafica e trasformarla in realtà comprensibili e condivisibili.

Nonostante sia stato appurato che le strumentazioni che gli archeologi adoperano sul terreno sono di seconda mano, cioè sviluppate per attività non archeologiche, ma rapidamente accettate dalla comunità scientifica in modo del tutto passivo, l'eco di un'onda tecnicista appare ancora presente, in modo particolare, nel settore del rilievo digitale (Beale, Reilly 2017). Questo approccio di tipo neopositivista si manifesta nei termini di una presunta maggiore oggettività del dato spaziale acquisito ed elaborato da una macchina rispetto alle tradizionali tecniche di rilievo.

Le repliche tridimensionali di strutture o stratigrafie sono il risultato dell'impiego di due distinte tecniche, talvolta adoperate in modo complementare (Remondino, Campana 2014; Bosco 2022): la modellazione, che si basa sull'estrusione di elementi architettonici a sezione costante e sul disegno degli elementi strutturali di un edificio, e il *data-capturing* che sfrutta le tecnologie 3D *reality-based* (laser scanner, fotogrammetria aerea e terrestre) che restituiscono nuvole di punti. Un terzo approccio, poco utilizzato, prevede l'impiego di specifici algoritmi che implementano regole procedurali per modellare oggetti solidi in 3D. I tre distinti metodi sono integrabili; nel caso di parti mancanti che non possono essere acquisite dagli strumenti 3D, può intervenire la model-

lazione per eseguire un restauro virtuale.

Il primo metodo presuppone l'esistenza di una mappa misurata o in scala che consenta di ricavare le dimensioni degli oggetti architettonici o strutturali da digitalizzare; il modello può, successivamente, essere migliorato applicando sulla geometria solida una texture che conferisce un aspetto realistico alla replica 3D. La modellazione prevede non soltanto specifiche competenze nel disegno architettonico digitale, ma anche una approfondita conoscenza archeologica e strutturale dell'edificio da restituire graficamente. Nel caso di evidenti lacune, lo studioso può utilizzare confronti con monumenti contermini e contemporanei per proporre tecniche costruttive, soluzioni architettoniche e motivi decorativi. La replica digitale finale implica un consistente intervento umano con una specifica competenza tecnica nel disegno tecnico e, soprattutto, una conoscenza sulle tipologie di monumenti da rilevare e restituire correttamente. I software di modellazione forniscono al ricercatore gli strumenti grafici e di editing per l'implementazione di una precisa replica digitale. L'oggetto tridimensionale finale può essere riutilizzato in software di realtà virtuale per essere contestualizzato in un paesaggio simulato.

A differenza della modellazione, l'acquisizione dei dati geometrici è affidata interamente alla tecnologia prescelta sebbene il ricercatore possa, sulla base di particolari obiettivi di ricerca, variare alcuni parametri per la fase di acquisizione e processamento dei dataset. La scelta della strumentazione e della metodologia di intervento è, spesso, subordinata alla disponibilità delle attrezzature, alle risorse finanziarie e, talvolta, alla concessione di particolari autorizzazioni da parte delle istituzioni locali; di frequente, ad esempio, le autorità governative di numerosi paesi in Africa e Asia tendono a imporre limiti nell'uso dei droni anche per l'esecuzione di rilievi a bassa quota. Al termine della acquisizione ed elaborazione, le tecnologie 3D offrono repliche di grande impatto e di enorme interesse per la ricchezza di dettagli registrati che possono essere esaminati successivamente in laboratorio. Nonostante gli straordinari risultati che questa tecnica consente di raggiungere, sembra che i ricercatori siano più interessati alla quantità di punti e triangoli ottenuti che alla accuratezza della risoluzione geometrica finale. I modelli sono privi delle necessarie informazioni sulle modalità di registrazione e allineamento delle scansioni; la valutazione del prodotto è in genere determi-



nata dalla quantità di foto o scansioni acquisite e dalla tipologia di strumento adoperato. La verifica sull'affidabilità della replica digitale è sostituita dalla illustrazione della capacità di campionamento del laser scanner o dalla risoluzione pixel per le macchine fotografiche. Tuttavia, senza una procedura di controllo e di stima dell'errore, la fedeltà della restituzione digitale finale non è in alcun modo accertabile.

Se gli scenari delineati dall'OS incoraggiano gli studiosi a condividere i loro dati, è necessario che ogni dataset sia auto esplicativo per essere elaborato anche da agenti automatici. Mentre un errore di registrazione del rilievo di pochi centimetri non altera l'interpretazione e la datazione di un oggetto archeologico, la pubblicazione di un dato impreciso può comprometterne il suo corretto riuso. Spesso, l'affidabilità spaziale/geometrica dei dati può essere solo stimata e, a tal proposito, Chapman (2001) ha evidenziato che nessun riuso di dati può essere ammesso in assenza di adeguati metadati. Per Chapman (2001, 19)

The re-usability of archaeological surveys has important implications in relation to the development of digital data management and archiving. The level of abstraction is important to reduce every possible mistake or misunderstanding.

Criteri di trasparenza e registrazione di ogni possibile informazione sulla *provenance* digitale del dato consentono di trasformare una replica in un documento scientifico (D'Andrea 2015a). L'assenza di buoni metadati non può essere compensata da immagini accattivanti, ad alta risoluzione, oppure da modelli interattivi e visualizzabili con specifici dispositivi oculari. Un buon dataset fornisce informazioni significative anche sulla metodologia di scavo e rilievo adottata dal ricercatore. Insufficienti descrizioni sulle strategie di controllo sul terreno del dato geometrico rendono meno comprensibile la replica digitale (D'Andrea *et al.* 2018; 2020). Le informazioni sulle tipologie di misurazione selezionate e uso di metadati appropriati sono elementi che contribuiscono ad una lettura scientifica della risorsa digitale.

Rispetto alla tradizionale documentazione archeologica eseguita con strumenti analogici, il digitale fornisce molteplici strategie e traiettorie di acquisizione e processamento; nonostante gli indubbi vantaggi introdotti

dalle nuove tecniche di rilievo, la transizione digitale è avvenuta in un ambiente acritico dove il tema centrale è circoscritto alla scelta degli strumenti e sulla valutazione dei risultati in termini quantitativi (Huggett 2015).

Mentre nel rilievo analogico e, per certi aspetti, anche nella modellazione digitale, il rilevatore identifica i dettagli e le forme architettoniche e costruttive da misurare e rappresentare graficamente, gli strumenti digitali impiegati nella acquisizione del dato geometrico separano il ricercatore dalla riproduzione del record spaziale. La replica digitale può essere fruita dinamicamente senza alcuna prestabilita prospettiva; la presenza di descrittori relativi all'oggetto fisico digitalizzato abilita la scoperta in rete di ulteriori informazioni arricchendo, così, la percezione del record (Frischer 2008).

L'approccio visuale all'informazione archeologica ha favorito lo sviluppo di una riflessione metodologica e teorica sulla funzione dell'immagine nella interpretazione storica (Moser 2012). Una riproduzione digitale contiene un processo creativo e interpretativo che presuppone competenze tecniche e scientifiche. La documentazione grafica, quindi, non è il duplicato di un oggetto, ma un prodotto scientifico che contribuisce alla stessa formazione della conoscenza archeologica. Nella sfera digitale il record archeologico è codificato in un medium, duplicabile e modificabile, che può essere diffuso velocemente.

L'illustrazione tradizionale racchiudeva un insieme di saperi espressi con convenzioni e simboli grafici. Grazie alla disponibilità di ambienti di visualizzazione scientifica, gli oggetti 3D offrono al ricercatore una esperienza cognitiva nuova il cui obiettivo è fornire strumenti per una esplorazione originale; nella risorsa tridimensionale si integrano e si fondono in forma creativa idee e ipotesi che concorrono alla creazione di molteplici narrative.

L'evoluzione delle tecniche di acquisizione, elaborazione e visualizzazione dell'informazione geometrica sta modificando l'interazione uomo-macchina consentendo di percepire e riproiettare vecchi e nuovi dati per più moderne indagini e ricerche. Mentre la grafica computazionale si limita a creare l'oggetto digitale, la visualizzazione scientifica migliora le nostre capacità cognitive e, quindi, di comprensione della replica 3D. Al ricercatore spetta il ruolo di alimentare la dimensione collaborativa, inclusiva e partecipativa di una scienza che reclama un accesso democratico

e diretto non solo alla cultura, ma anche ai dati; in tale scenario, gli studiosi dovranno prestare una particolare attenzione alla qualità del dato spaziale condiviso, associando i necessari metadati e *paradata* e assicurando, in tal modo, la validazione del dato geometrico/topografico. Soltanto in questo modo sarà possibile riutilizzare l'archivio e integrarlo con altri dataset magari multi-temporali e multidimensionali (Forte 1993).

Vedremo nei successivi paragrafi come le tematiche della condivisione dei dati geometrici (monumenti, statue, reperti mobili, etc.) siano state esaminate nella prospettiva di una Scienza Aperta.

#### 4.4. Oggetti archeologici 3D

Il grande pubblico oramai da tempo si mostra sempre più attento ai contenuti digitali 3D, non soltanto per gli aspetti visuali e spettacolari presenti nelle ricostruzioni, ma anche per quelle informazioni supplementari che descrivono le tecniche impiegate, i modi di realizzazione, la base scientifica e ogni eventuale ulteriore elemento relativo ad ipotesi ricostruttive differenti. In rete si trovano molteplici modelli 3D interattivi - si pensi agli esemplari che riproducono i tanti edifici pubblici o privati di Pompei - e non sempre è facile comprendere quali repliche tridimensionali rispettino i monumenti originari oppure ne siano una fantasiosa rielaborazione o una delle tante versioni alternative scientifiche o para-scientifiche proposte dai ricercatori.

Districarsi tra differenti contenuti 3D non è semplice. Per superare queste limitazioni, una illustrazione completa dell'oggetto culturale deve necessariamente accompagnare le fonti visualizzabili e consultabili; solo in questo modo è possibile garantire il corretto uso e riuso della risorsa online assicurando allo stesso tempo una adeguata conservazione e preservazione della replica digitale nel lungo periodo.

Questo approccio innovativo, basato sulla creazione di modelli 3D altamente scientifici associati a descrizioni di elevata qualità, è stato sperimentato con successo nel progetto europeo 3DICONs. La proposta, ispirata alla strategia Europea 2020, si inseriva nelle tematiche contenute nell'Agenda Europea Digitale, finalizzata allo sviluppo di una economia digitale. In particolare, 3DICONs mirava ad accrescere l'uso ed il riuso delle fonti 3D online attraverso la creazione di una ricca collezione di oggetti tridimensionali di siti, monumenti ed oggetti archeologici ac-

cessibili dal portale EUROPEANA (*supra* 130).

Tra gli obiettivi indicati nel progetto due sono risultati di straordinario interesse per il supporto di procedure di accesso e condivisione dei dati: l'individuazione delle metodologie per garantire la pubblicazione online di repliche 3D, anche molto complesse, e la creazione di uno schema di metadati da collegare alle risorse tridimensionali.

Differenti metodologie e soluzioni tecniche sono state messe a punto per la pubblicazione online degli oggetti 3D, in particolare sul versante delle procedure di ottimizzazione tra quantità di dettagli da visualizzare e complessità del modello accessibile in rete. Nella valutazione del formato da utilizzare alcuni specifici aspetti devono essere esaminati in funzione della potenziale platea di utenti. Qualsiasi opzione doveva essere ispirata ai seguenti criteri di carattere generale:

- servire una ampia varietà di utenti con differenti competenze, obiettivi e strumenti;
- ottimizzare l'esperienza 3D del fruitore finale;
- consentire ad utenti non esperti di accedere in modo semplice ai contenuti 3D;
- realizzare un modello accessibile da una molteplicità di browser e dispositivi evitando installazioni supplementari di software e plug-in;
- supportare il concetto di "risorsa esplorabile" (ad esempio url integrate nel modello).

Alcune importanti raccomandazioni hanno caratterizzato lo sviluppo di 3DICONs e la progettazione delle soluzioni proposte al termine del progetto. In primo luogo, nell'adattare le repliche 3D ai codici del Web, i creatori di contenuti 3D devono valutare se la pubblicazione dell'oggetto richiede particolari formati che garantiscano la corretta fruizione della risorsa. Inoltre, se la complessità della replica digitale non è compatibile con le risorse della rete, una eccessiva semplificazione del modello rischia di compromettere il risultato finale che non risulta utile né per lo specialista né per un pubblico di non addetti ai lavori. 3DICONs si proponeva di fornire soluzioni innovative per affrontare il tema critico della pubblicazione online di complessi archeologici ed architettonici particolarmente vasti garantendo, nello stesso tempo, una alta qualità della

risorsa digitale accessibile.

Per la pubblicazione online degli oggetti 3D, 3DICONs ha sperimentato due approcci principali: il formato 3D PDF o soluzioni basate sull'ambiente WebGL. In un numero limitato di casi sono stati testati anche formati alternativi. Il formato 3D PDF, adoperato con ottimi successi anche da altri progetti europei, tra i quali Linked Heritage, offre la possibilità di integrare negli oggetti 3D anche annotazioni; esso appare un modo rapido ed efficiente per pubblicare e, soprattutto, condividere in rete oggetti 3D. Il documento aperto nel software proprietario di Adobe Acrobat o Reader fornisce un set di strumenti per misurare, annotare, creare sezioni ed animazioni e può essere adoperato offline e online. Particolari viste possono essere predefinite per semplificare l'interazione con l'utente finale, mentre addizionali contenuti come testi, video e immagini possono essere agevolmente integrati nel PDF. I modelli non sono ottimizzati per la visualizzazione online e devono, pertanto, essere necessariamente scaricati; spesso, per semplificare il download, la replica digitale viene decimata con evidenti perdite di dettagli e compressione delle texture. Inoltre, il formato non è utilizzabile per applicazioni su dispositivi mobili, mentre i viewers sviluppati come plug-in dai fornitori dei browser non hanno le funzionalità avanzate di Adobe e non consentono la visualizzazione e fruizione di contenuti 3D. Il formato 3D PDF appare di grande utilità soprattutto per oggetti poco complessi, come statue, vasellame e ambienti di dimensioni ridotte. Tuttavia, il 3D PDF ha una interattività limitata e non è in grado di fornire un ricco senso di immersione. Quando i modelli 3D sono incorporati all'interno di un PDF, di solito non sono collegati dinamicamente ad alcuna tipo di metadati e *paradata*.

Altre soluzioni, basate sul WebGL integrato nei principali browser, sono state sviluppate soprattutto per consentire la visualizzazione online di file particolarmente complessi e con elevati dettagli superando alcuni dei principali svantaggi del formato 3D PDF. L'ambiente WebGL è attualmente stabile e garantisce le funzionalità necessarie non solo alla pubblicazione, ma anche alla ottimizzazione di riproduzioni digitali di architetture e monumenti articolati. Lo sviluppo dello standard WebGL nel 2009 ha rappresentato un significativo passo in avanti per l'integrazione del 3D nel Web. WebGL è una interfaccia di programmazione che consente ai browser di eseguire il rendering di repliche 3D in

modo nativo.

I vantaggi di questi approcci sono in estrema sintesi i seguenti:

- nessuna necessità di installare software o plug-in per visualizzare gli oggetti 3D;
- sono supportati dalla maggior parte dei browser;
- sono multiplatforma e sono visualizzabili sui dispositivi mobili;
- usano in modo trasparente tutte le risorse 3D della scheda grafica.

WebGL usa una architettura di tipo cloud: i modelli risiedono su un server che fornisce il software necessario alla corretta visualizzazione del 3D in una normale pagina html.

Tra gli altri formati adoperati per la visualizzazione online degli oggetti si segnala in particolare l'X3D, successore del VRML; X3D è un linguaggio XML creato per descrivere ambienti virtuali interattivi che possono essere riprodotti e visualizzati in ambiente HTML5. Il formato offre straordinarie capacità di manipolazione e può essere adoperato nel WebGL; tuttavia, il modello finale, basato su differenti file, non può essere referenziato ad un singolo URI. Alcuni oggetti, infine, possono sfruttare le tecnologie adoperate per i video giochi (Unity3D, Unreal, CryEngine) oppure per oggetti pseudo 3D (objectVR).

Nonostante il 3D PDF rimanga un formato molto agevole per la distribuzione di oggetti 3D online, nuove soluzioni sono state sperimentate con l'obiettivo di ampliare ed estendere l'interazione dell'utente con il modello 3D accessibile in rete senza limitare e ridurre, nello stesso tempo, la complessità dell'oggetto originale. Già oggi ambienti come l'HTML5 e il WebGL, sono in grado di rendere gli oggetti 3D in elementi naturali della rete. Numerosi progetti europei, tra cui ARIADNE, hanno sviluppato interfacce grafiche per il caricamento e la visualizzazione online di repliche 3D.

#### 4.4.1. Metadati per il 3D

Considerata la varietà dei processi di digitalizzazione disponibili e la complessità - non soltanto morfologica - degli oggetti culturali da replicare in formato digitale, la produzione in larga scala di record 3D richiede la creazione di un adeguato sistema di gestione per le fasi di acquisizione, elaborazione e pubblicazione dei modelli. In particolare, 3DI-

CONS ha sviluppato uno schema di metadati in grado di registrare l'intero ciclo di vita del record dalla fase iniziale della acquisizione dell'oggetto 3D fino alla conservazione e ai possibili riusi del modello.

Nel campo delle tecnologie per il 3D esistono numerose strumentazioni e metodologie utilizzabili per il *data-capture* e il processing dei dati. Differenti tecniche possono essere integrate in un unico processo oppure essere impiegate in forma alternativa per ragioni di ordine pratico, logistico o di tipo scientifico. Le motivazioni - didattiche e/o comunicative - che ispirano la creazione di un oggetto 3D rendono ancora più complesso il processo di digitalizzazione dell'oggetto culturale poiché di frequente per uno stesso record archeologico abbiamo a disposizione molteplici interpretazioni e ricostruzioni.

I modelli finali, indipendentemente dalla loro complessità e/o risoluzione e dalle tecniche più o meno sofisticate impiegate, non possono essere interpretati e correttamente riutilizzati senza poter disporre delle informazioni relative agli oggetti originari ed al modo ed alle condizioni in cui tali oggetti sono stati digitalizzati. Il ricercatore può annotare tutte le informazioni necessarie ad una completa comprensione del record digitale all'interno di un documento descrittivo in grado di registrare gli eventuali settaggi selezionati (calibrazione, tolleranza, errore strumentale), le condizioni fisiche al momento della ripresa e tutti quegli ulteriori elementi, come le fasi di post-elaborazione, che possono contribuire al miglioramento o perfezionamento del processo.

Per rispondere alla maggiore richiesta di comprensione e trasparenza dei modelli online, 3DICONs ha realizzato uno schema di metadati basato sulla definizione della *provenance*, come richiamato dal W3C *Provenance Incubator Group*, e dei *paradata*, secondo i principi contenuti nella London Charter (*infra* 185), nonché sulla descrizione dettagliata dell'oggetto culturale.

Lo schema adottato dal progetto 3DICONs si basa sull'aggiornamento di un modello dati realizzato dal progetto CARARE, una precedente iniziativa finanziata dalla Commissione Europea. CARARE era stato messo a punto per assicurare l'interoperabilità tra i metadati creati e conservati da differenti istituzioni culturali europee. Lo schema aveva lo scopo di semplificare la mappatura di collezioni di metadati su un tracciato comune. CARARE si basava sulla descrizione dell'oggetto fisi-

co, delle relazioni con le successive riproduzioni digitali e delle diverse attività che coinvolgevano il bene culturale registrato, tra cui, ad esempio, il trasferimento di proprietà e lo spostamento degli oggetti tra differenti musei. L'aggiornamento della versione di CARARE sviluppata da 3DICONs aggiunge al tracciato originario un set di dati che descrive i processi impiegati a livello umano e tecnologico per la digitalizzazione dell'oggetto 3D.

Le informazioni aggiuntive sono state strutturate sulla base dello schema del CRMdig, una estensione del CIDOC-CRM testato positivamente nell'ambito del progetto europeo 3DCOFORM. Il modello appariva particolarmente adatto alla rappresentazione del workflow per il 3D, dall'acquisizione all'elaborazione, dalla decimazione, fino alla pubblicazione e, infine, riuso della risorsa digitale. Attraverso una complessa rete semantica di relazioni connesse ad eventi differenti, lo schema consente di registrare gli strumenti adoperati e le differenti fasi del processo di elaborazione e trattamento del modello 3D; con CRMdig è possibile descrivere correttamente il sistema di organizzazione delle riprese e la metodologia adoperata, ad esempio, per l'allineamento delle scansioni.

Tutti questi aspetti sono stati recepiti nel nuovo schema adottato da 3DICONs; con l'introduzione di alcune specifiche relazioni (*general\_purpose*; *had\_specific\_purpose*) è stata semplificata l'associazione tra le diverse fasi di digitalizzazione e processing dei dati con gli obiettivi generali e specifici delle singole ricostruzioni digitali 3D. Leggendo i metadati è quindi possibile comprendere se la bassa risoluzione di un oggetto sia stata prescelta per aumentare l'interazione e l'interattività online dell'oggetto per finalità divulgative oppure se i livelli di dettaglio riprodotti abbiano una affidabilità scientifica o siano discrezionali, arbitrari o una delle possibili ricostruzioni. I metadati contengono, inoltre, informazioni sull'accesso, la licenza e le possibilità di riuso dei dati e di qualsiasi altra informazione digitale associata al modello.

Poiché l'obiettivo di 3DICONs era la creazione di modelli 3D da fruire all'interno di EUROPEANA, lo schema è stato successivamente mappato sull'EDM (*supra* 131), sebbene quest'ultimo soltanto in minima parte può essere allineato al CRMdig. Pertanto, molte delle informazioni registrate con lo schema di 3DICONs non sono mappabili sull'EDM e



non sono visualizzabili in EUROPEANA.

Il trasferimento dei metadati verso l'aggregatore europeo è stato eseguito utilizzando differenti strategie. Mentre alcuni partner del progetto 3DICONs disponevano, già prima dell'avvio delle attività, di dati strutturati relativi agli oggetti culturali da replicare in 3D, altri soggetti non avevano alcuna banca dati di tipo descrittivo o testuale; se nel primo caso era necessario mappare i record locali sullo schema di 3DICONs consentendo, eventualmente, l'aggiunta delle voci mancanti per la *provenance* e i *paradata*, nel secondo, invece, si doveva sviluppare uno strumento in grado di semplificare la compilazione dei metadati.

Per i dati formalizzati preesistenti, come ad esempio i record dei database di collezioni museali, è stato riutilizzato uno strumento online sviluppato nel progetto CARARE; l'applicazione MINT, adeguata ai nuovi obiettivi di 3DICONs, forniva le funzioni necessarie al caricamento dei dati e al successivo *mapping*. L'utente può decidere come eseguire l'aggregazione tra i due schemi e quali voci aggiungere nella trasformazione. MINT verifica la correttezza della procedura, la conformità allo schema, la cardinalità delle voci, l'univocità dei campi indicizzati. Al termine dei controlli viene eseguita la conversione nel nuovo tracciato elaborato da 3DICONs e, se richiesto, anche sullo schema EDM adottato da EUROPEANA.

Diverso è stato l'approccio messo a punto per i partner senza precedenti dati descrittivi e con poche competenze nella produzione di metadati. Per facilitare la creazione di record è stato sviluppato un altro applicativo, *Metadata Editing Tool*, che fornisce le funzioni base per la compilazione online di nuovi metadati; il sistema si basa su semplici template o blocchi di informazioni che possono essere creati, modificati e riutilizzati per la redazione delle schede. Allo scopo di favorire l'omogeneità dei record e, soprattutto, per promuovere l'uso di strumenti semantici, gli utenti hanno utilizzato nella creazione dei metadati alcune voci selezionate dai valori di vocabolari associati ad URI persistenti; in particolare per gli attori, i concetti principali e i luoghi si è fatto riferimento rispettivamente a risorse DBpedia, Getty Thesaurus, Geonames. La scelta di creare i metadati parallelamente alla acquisizione ed elaborazione degli oggetti 3D, ha assicurato uno sviluppo lineare per l'intero processo di digitalizzazione e pubblicazione delle risorse 3D.

Alcune delle soluzioni proposte da 3DICONs sono, oggi, sostituite

da applicazioni che sfruttano in modo ottimizzato le tecnologie semantiche. Tuttavia, alcune delle idee innovative messe a punto nel progetto soprattutto per la definizione di un workflow, rimangono ancora fondamentali (3D content in *Europeana* 2020).

#### 4.4.2. IPR e 3D

Strettamente connesso al tema della divulgazione online del contenuto 3D è il riconoscimento della proprietà intellettuale. In 3DICONs questo aspetto è stato esaminato considerando in primo luogo come le differenti legislazioni nazionali analizzano e proteggono le varie fasi di lavoro. Lo schema di IPR proposto è finalizzato a salvaguardare i diritti individuali sui metadati, sui modelli e sulle eventuali derivazioni (foto e video) oltre che sull'integrazione dell'oggetto 3D in nuove risorse multimediali.

Nella creazione di un contenuto 3D gli attori coinvolti sono numerosi: dal custode del bene fino al soggetto che, dopo l'acquisizione ed il trattamento dei dati grezzi, distribuisce ed eventualmente commercializza il modello finale. Un corretto modello di IPR deve, quindi, prevedere ogni aspetto della produzione di un oggetto 3D proteggendo le competenze e gli investimenti dei differenti soggetti interessati.

Se il proprietario del bene, pur impegnando poche risorse finanziarie ed umane, dispone di un grande potere di controllo sull'intero processo, il soggetto che realizza il 3D impiega attrezzature costose e un know-how altamente specialistico che tenderà a tutelare soprattutto nei confronti di un eventuale soggetto terzo interessato a produrre repliche dall'originario contenuto 3D.

Lo schema di IPR, selezionato da 3DICONs, integra in un unico modello tutte le fasi del processo cercando di evidenziare il ruolo di ciascun attore nella creazione delle risorse digitali; in tal modo l'accordo si rivela particolarmente utile nel caso di riuso, a fini commerciali, di oggetti 3D consentendo di allocare correttamente i diritti relativi ai diversi soggetti coinvolti nei differenti stadi dell'intera procedura.

Ogni istituzione, in base alle proprie politiche e alle legislazioni nazionali, può liberamente decidere come tutelare i propri diritti. Sebbene siano disponibili licenze restrittive, che possono prevedere il pagamento per il solo accesso al modello, o accordi liberali, che al contrario offrono

l'accesso senza alcuna limitazione, l'adozione di una licenza del tipo CC-BY-NC-ND consente la distribuzione non commerciale del contenuto ponendo l'obbligo, in caso di riuso, che sia sempre segnalata l'organizzazione e/o i soggetti che hanno creato in origine la risorsa digitale.

Per quanto riguarda i metadati, 3DICONs ha adottato la strategia seguita da EUROPEANA che, per incrementare la visibilità dei propri dati ed il riuso delle risorse, rilascia liberamente i record con licenza Creative Commons.

3DICONs ha fornito ad EUROPEANA, oltre 3.000 metadati relativi a circa 5.000 modelli 3D, più di 17.000 immagini ad alta risoluzione e 287 video; tutte le risorse con i relativi metadati sono scaricabili gratuitamente dal sito di EUROPEANA o dai differenti portali messi a disposizione dai partner. Per un numero limitato di oggetti, realizzati ad alta risoluzione, è necessario richiedere al proprietario del contenuto una specifica autorizzazione per il download.

All'interno di EUROPEANA divengono accessibili in 3D complessi archeologici e architettonici che testimoniano la ricchezza e la varietà del patrimonio culturale europeo; spesso si tratta di siti localizzati in aree remote non facilmente raggiungibili oppure di oggetti fragili e di non semplice comprensione da parte di utenti non specialisti. Molte ricostruzioni sono state realizzate con tecniche differenti e con diversa risoluzione per soddisfare una ampia serie di potenziali utenti. In alcuni casi i modelli sono basati su ipotesi ricostruttive alternative, mentre in altri oggetti dispersi in differenti collezioni e musei sono stati virtualmente ricollocati nel loro contesto originario.

Per raggiungere gli obiettivi definiti nella proposta progettuale, 3DICONs ha integrato strumenti e metodi per il 3D, dalla digitalizzazione e acquisizione con differenti tecnologie, fino al rendering finale; il protocollo al termine del progetto ha fornito numerosi approfondimenti sui formati 3D per il Web, sui metadati e, infine, suggerimenti sulla tutela della proprietà intellettuale.

L'esperienza, che emerge alla conclusione delle attività, rafforza il convincimento che nella produzione di un qualsiasi oggetto 3D un passaggio obbligato sia rappresentato dalla creazione di metadati di elevata qualità che favoriscano un più efficiente uso e riuso delle fonti migliorando l'accesso al contenuto 3D in rete. 3DICONs ha raggiunto questo obiettivo sviluppando uno schema di metadati in grado di semplificare

l'archiviazione e la pubblicazione online dei modelli riprodotti.

Per sfruttare completamente le potenzialità del modello dati implementato, 3DICONs ha realizzato uno specifico portale, ideato dapprima come strumento di verifica dello stato di avanzamento della procedura di *harvesting*. Il sito è stato successivamente trasformato in un repository per la presentazione degli oggetti 3D con i relativi metadati e *paradata*. La piattaforma è consultabile attraverso un motore di ricerca che consente di selezionare i modelli partendo dalla localizzazione dell'oggetto su una mappa oppure da una lista di valori. L'utente può visualizzare l'oggetto e, grazie ai metadati associati, ottenere informazioni complete sui processi, sulle attrezzature e sulle eventuali licenze per il riuso della fonte.

#### 4.4.3. Principi di base per il 3D

Un recente documento, redatto da un gruppo di esperti sul patrimonio culturale e di EUROPEANA su incarico della Commissione Europea, ha fornito nuove linee guida sui beni culturali 3D (*Basic principles* 2020). Il report contiene 10 principi di base ed una serie di suggerimenti rivolti a professionisti, istituzioni e autorità responsabili della gestione del patrimonio culturale europeo, nonché un aggiornamento sui formati per i documenti 3D online e su come migliorare la progettazione di infrastrutture. Il documento elenca una serie di raccomandazioni elementari indirizzate ad utenti non esperti, ma non mancano importanti approfondimenti sul tema delle licenze e, soprattutto, della formazione; a quest'ultimo obiettivo è dedicata una breve rassegna di recenti contributi sulla digitalizzazione 3D.

Il report sottolinea che le tecnologie 3D non devono in alcun modo sostituire il bene culturale fisico, sebbene, in taluni casi, soltanto il digitale consenta specifici approfondimenti. Ad esempio, una replica 3D fisica può essere impiegata per attività didattiche rivolte a persone con disabilità visive; l'accesso al patrimonio culturale digitale offre un'esperienza tattile che può trasmettere sensazioni sulla forma, il volume e la consistenza dei materiali.

Nel processo di progettazione di una acquisizione 3D una significativa importanza deve essere assegnata al gruppo di utenti target del contenuto 3D e al modo attraverso il quale l'oggetto 3D verrebbe utilizzato. Se si tratta di conservatori, professionisti, archeologi, gestori di un edifi-

cio storico o di un sito archeologico, visitatori di musei o studenti, l'aggregatore di risorse 3D deve soddisfare le esigenze dei vari potenziali utenti rispettando le singole competenze necessarie per una piena fruizione dei contenuti digitali in rete.

Chi si occupa di 3D deve essere un professionista e non un operatore inesperto poiché è determinante avere sia la conoscenza delle attrezzature e delle fasi di elaborazione, e nello stesso tempo la capacità di comprendere i limiti delle diverse tecniche 3D e di analizzarne e giudicarne i risultati. Un'altra importante raccomandazione riguarda standard e best practice. Per massimizzare l'utilità, l'accessibilità e il potenziale a lungo termine dei contenuti 3D, occorre scegliere formati aperti e possibilmente licenze chiare che favoriscano il riuso. Avendo tempo e spazio a disposizione per la conservazione dei dati, le repliche 3D dovrebbe essere codificate in almeno due distinti formati, di cui almeno uno aperto. Il report fornisce una lista, non esaustiva, di formati aperti comunemente utilizzati per la generazione di modelli 3D: glTF, X3D, STL, OBJ, DAE, PLY, WRL, DICOM o IFC. I dati di alta qualità e ad alta risoluzione possono essere memorizzati nei formati \*.OBJ o \*.DAE. I formati più comuni per la visualizzazione della realtà virtuale e aumentata sono .USDZ e .glTF. Per la stampa 3D sono ampiamente utilizzati i formati .STL e per la visualizzazione sul web .X3D e .GLB.

#### 4.5. La visualizzazione del dato geometrico 3D

Il numero e la qualità delle informazioni online cresce rapidamente anche come conseguenza dell'affermazione di un modello collaborativo di ricerca e di linguaggi Web che contribuiscono a migliorare l'interazione uomo-macchina sul versante della ricerca e dello sfruttamento del dato in rete. Nonostante ciò, la creazione sul Web di modelli 3D complessi, composti da milioni di punti, è ancora un obiettivo difficile da raggiungere senza una specifica competenza; inoltre, sono ancora di difficile applicazione le soluzioni che consentono in tempo reale di convertire i modelli 3D e renderli pubblicabili sul Web.

La presenza di siti web specializzati, che pubblicano collezioni di oggetti culturali in 3D, suscita l'interesse di un vasto pubblico che chiede strumenti più raffinati per la navigazione, l'interrogazione e l'esplorazione delle repliche digitali 3D. Motivazioni scientifiche, esigenze didattiche

che e intenti formativi hanno favorito la progettazione e realizzazione di numerosi aggregatori e piattaforme Web 3D che offrono al ricercatore la possibilità di rendere accessibili i loro modelli digitali. Anche la recente epidemia legata alla diffusione del Covid19 ha contribuito ad una maggiore consapevolezza sull'uso dei dati 3D, accelerando la virtualizzazione della cultura e delle forme della sua valorizzazione.

Numerose istituzioni culturali utilizzano strumentazioni 3D per documentare e rendere fruibile il loro patrimonio spesso inaccessibile o con ingresso limitato oppure a rischio di distruzione naturale o antropica. La disponibilità di archivi di dati 3D relativi a monumenti, oggetti o semplici scavi archeologici ha sollevato interessanti quesiti su come le collezioni 3D possano coinvolgere un più vasto pubblico in quella intersezione sottile tra mondo digitale ed esperienza soggettiva. La disponibilità di ricche collezioni di dati 3D favorisce la creazione di mostre online che sfruttano le potenzialità della realtà immersiva per trasformare la visita museale in una avventura coinvolgente.

Differenti progetti di ricerca hanno sviluppato iniziative per definire il processo di digitalizzazione 3D (3DICONs), per creare servizi e repository Web (ARIADNE), per promuovere la conservazione dei tridimensionali (3DCOFORM e CARARE), per manipolare e riassemblare anche automaticamente frammenti 3D (ArchAIDE, GRAVITATE). Alcune ricerche sono invece dedicate allo studio dell'impatto che l'accesso alle collezioni 3D in rete produce nella ricerca e nel pubblico. Ciò che emerge da queste indagini, legate allo studio degli aspetti pedagogici sottesi alla pubblicazione di collezioni 3D in rete, è che le repliche digitali possono avere il ruolo di normale complemento all'insegnamento e possono servire quando l'accesso ai dati reali sia limitato; nonostante ciò, la mancanza di esperienza sugli oggetti fisici confonde l'utente che non è, sempre, in grado di identificare correttamente la risorsa 3D esaminata (Ekengren *et al.* 2021). Le immagini e i modelli appaiono, così, troppo astratti per costituire la base di una comprensione generale dei manufatti; mancano informazioni o impressioni sul peso dell'oggetto o dettagli tattili sulla eventuale decorazione e su materiale che sono essenziali per il riconoscimento e la classificazione. Gli studenti però possono avere accesso agli oggetti in qualsiasi orario anche da casa, possono ingrandire anche minuscoli dettagli evidenziandoli anche attraverso la

modifica dei parametri di illuminazione; inoltre, possono aggiungere delle annotazioni condividendole con altri allievi. Infine, l'accesso alla banca dati, con i relativi metadati, consente di operare confronti e aggregazioni di manufatti tipologicamente affini.

Il tema del confronto tra oggetti 3D e campioni reali nella identificazione delle specie, in questo caso zoo-archeologiche, è alla base della creazione della piattaforma pilota Bonify sviluppata con il framework 3DHOP (Nobles, Çakırla, Svetachov 2019). Le collezioni di riferimento zoo-archeologico risultano di straordinaria utilità scientifica per la ricerca e la formazione dei giovani professionisti; i ricercatori possono confrontare ossa antiche con scheletri recenti allo scopo di identificare la specie. Quando due animali, come la pecora e la capra, sono imparentati, identificare le distinzioni morfologiche può diventare un compito difficile; tuttavia, l'incremento degli scheletri a disposizione per le comparazioni e studi analitici sui dettagli delle ossa consentono di rimuovere le differenze. Le limitazioni nell'accesso alle collezioni, rende il 3D un insostituibile strumento per l'esecuzione delle ricerche. Pertanto, sono stati digitalizzati alcuni campioni di ossa estratti dalle collezioni custodite presso il *Groningen Institute of Archaeology*. Successivamente, allo scopo di comprendere il livello di interazione e percezione con gli oggetti digitali presenti sulla piattaforma Web, sono stati intervistati 22 utenti tra professionisti e studenti che avevano avuto almeno una precedente esperienza con le collezioni reali. Le risposte fornite confermano che la piattaforma non sostituisce il confronto fisico, ma può certamente essere complementare ed è di ausilio nella ricerca e nell'insegnamento. Secondo gli autori della ricerca (Nobles, Çakırla, Svetachov 2019) i risultati del questionario mostrano che, allo stato, nessun processo di digitalizzazione, per quanto corretto geometricamente, può sostituire la visione autoptica dei campioni ossei. Ma, il patrimonio archeo-zoologico potrebbe raggiungere in rete una massa critica e una quantità e qualità tale di repliche da essere difficilmente gestibile con i tradizionali approcci. Inoltre, il tema della sostenibilità finanziaria e, in conseguenza di ciò, della preservazione degli archivi digitali 3D, impone una riflessione sull'intera catena dalla digitalizzazione alla presentazione online della replica.

La svolta digitale dell'archeologia e di alcuni specifici settori di indagine ha prodotto, tra l'altro, la creazione di numerose piattaforme, anche a paga-

mento, implementate da aziende private (ad esempio, Google, Sketchfab, Autodesk). Sketchfab utilizza un efficiente visualizzatore web 3D, ma è limitato alla navigazione di un singolo oggetto impedendo ogni confronto interattivo tra le riproduzioni digitali di differenti manufatti. Pertanto, la piattaforma è utilizzata, soprattutto, per la visualizzazione di record 3D e fornisce un supporto inadeguato per compiti più avanzati, mentre i metadati associati all'oggetto registrano pochissime informazioni.

Per contrastare la tendenza ad una privatizzazione nella creazione e, quindi, nell'accesso ai documenti 3D e allo scopo di offrire specifici strumenti e servizi orientati allo studio delle collezioni sono state sviluppate alcune infrastrutture di ricerca. ARIADNE, ad esempio, mette a disposizione degli utenti un servizio per caricare e pubblicare risorse visive spaziali e geometriche anche complesse. L'infrastruttura consente il caricamento automatico di file multimediale e li trasforma in una pagina Web che supporta la visualizzazione interattiva delle risorse. Il servizio non è limitato alla pubblicazione dei dati 3D; è possibile anche caricare immagini ad altissima risoluzione e dataset composti da risorse che si riferiscono ad uno specifico oggetto. ARIADNE si basa su WebGL e 3DHOP assicurando, così, una piena compatibilità con tutti i browser. Il sistema sfrutta il framework multi-risoluzione Nexus che converte la replica digitale in poche migliaia di triangoli a diversa risoluzione che sono successivamente riassemblati in base al punto di vista selezionato dall'utente, riducendo, in questo modo, la quantità di dati che devono essere trasmessi al Web. La sola limitazione consiste nella tipologia di formato e nella grandezza del file; il servizio di caricamento accetta modelli codificati in .ply fino a 500Mb.

L'importanza che ha ricevuto la creazione e pubblicazione del dato 3D in rete ha spinto alla formazione di numerosi gruppi di ricerca impegnati nell'approfondimento delle tematiche relative, soprattutto, agli aspetti della condivisione e integrazione delle risorse digitali 3D. CS3DP è una comunità costituita in maggioranza da studiosi americani, tra cui bibliotecari, che ha di recente pubblicato un importante volume (Moore, Rountrey, Kettler 2022) che raccoglie i contributi emersi nel corso di due forum organizzati nel 2018 e di cinque gruppi di lavoro indirizzati alla analisi degli archivi 3D da differenti prospettive. CS3DP intende esplorare l'area della conservazione e delle best-practices allo scopo di supporta-



re una gestione più responsabile dei dati tridimensionali. La comunità, inoltre, ha l'obiettivo di garantire il progressivo sviluppo di standard per il 3D e migliorare le pratiche di cura e protezione degli archivi digitali.

Altri gruppi che affrontano tematiche connesse al coordinamento delle attività di digitalizzazione e pubblicazione di archivi 3D sono il IIIF 3D Group e NFDI4Culture. Entrambe le comunità non hanno l'obiettivo di sviluppare nuove specifiche tecniche in termini di formato e/o linguaggi, ma più semplicemente di incrementare l'interoperabilità dei dataset attraverso una strategia che coinvolga utenti, bibliotecari e organizzazioni culturali nel miglioramento dell'uso degli standard Web. I due gruppi puntano alla creazione di una rete di esperti in grado di incoraggiare il confronto e lo scambio di buone pratiche tra produttori e utenti di dati tridimensionali con l'obiettivo di offrire un costante aggiornamento sulla varietà crescente di tecnologie e tipi di dati 3D. Altri progetti forniscono, come 3DICONs, suggerimenti per l'ottimizzazione dei flussi di lavoro 3D relativi alla digitalizzazione, ai metadati e alle tecniche per la pubblicazione di dati di ricerca tridimensionali. Share3D consente il riuso di contenuti caricati su Sketchfab in applicazioni per il turismo o la ricerca, mentre GLAM3D mette a disposizione degli utenti una guida completa per l'implementazione di dati aperti 3D.

Le biblioteche hanno di recente investito nella progettazione di infrastrutture, servizi e persone in grado di supportare in modo significativo non solo la ricerca e la didattica, ma anche le applicazioni di modellazione e acquisizione 3D. Al pari delle carte, dei libri e dei manoscritti, i documenti digitali 3D possono essere acquisiti e gestiti dalle biblioteche utilizzando una varietà di metodi e formati nonché modalità di visualizzazione e stampa dei modelli tridimensionali.

Molti progetti di ricerca e pubblicazioni esaminano l'utilità dei formati .OBJ, .3Ds, .u3D, .o3D, .x3D e .DAE, ma esistono anche altri formati di file, come ad esempio .UNITY, .SKP, .DXF e .BLEND, popolari tra le comunità di computer grafica 3D (Champion, Rahaman 2020). Uno studio di McHenry e Bajcsy (2008) registra oltre 15 anni fa oltre 140 formati di file per modelli 3D, mentre una recente indagine di Sketchfab (2019) ha stabilito che i formati .glTF e .ply sono sempre più utilizzati dalle comunità culturali.

Le infrastrutture di ricerca per il patrimonio culturale stanno modifi-

cando il tradizionale approccio del pubblico all'esperienza della visita museale, di mostre e siti archeologici attraverso la creazione di collezioni digitali 3D esplorabili in forma interattiva. La digitalizzazione dei beni culturali ispira nuovi interventi di conservazione, documentazione e divulgazione, mentre software, sempre più potenti e accessibili, contribuiscono alla realizzazione di una varietà di oggetti 3D, disponibili in formati flessibili e con collegamenti a piattaforme di social media. Tutto ciò apre nuove frontiere nella rappresentazione e diffusione del patrimonio culturale, ma nello stesso tempo solleva importanti questioni sulla affidabilità dei formati e degli standard, sulle forme di garanzia del diritto d'autore (D'Andrea *et al.* 2022) e, soprattutto, sulla conservazione nel lungo termine dei dati assicurandone l'accesso e l'uso per differenti motivi. Poiché le repliche 3D giocano un ruolo fondamentale nella comprensione e nell'interpretazione di contesti e siti archeologici, autenticità e certificazione sono aspetti fondamentali da garantire nella conservazione di un dataset tridimensionale. Le Carta di Londra e Siviglia (*supra* 00) sottolineano l'importanza di associare ai dati 3D metadati di alta qualità per descrivere in modo trasparente il processo di digitalizzazione 3D.

Nonostante il numero crescente di record culturali 3D (Münster 2018, 2022), e la presenza di oltre 50 archivi commerciali disponibili per la condivisione e lo scambio di modelli 3D, Champion e Rahaman (2019; 2020) sostengono che dati affidabili, formati di file efficienti, metadati e *paradata* di qualità e licenze sul diritto d'autore sono problematiche che ostacolano, ancora oggi, l'archiviazione e rallentano la diffusione del patrimonio digitale 3D. Per i due studiosi, nonostante la crescente domanda di modelli digitali 3D e di funzionalità correlate, mancano infrastrutture istituzionali o commerciali in grado di garantire un approccio scientifico. Per invertire la tendenza gli autori suggeriscono la creazione di piattaforme Web che prevedano le seguenti funzionalità:

- la misurazione degli oggetti e l'estrazione di piante e sezioni trasversali;
- l'assegnazione e la gestione di DOI ai modelli;
- la realizzazione di collegamenti a documenti d'archivio;
- l'associazione con metadati affidabili;
- e, infine, l'uso di sistemi di visualizzazione che consentano il con-

fronto tra differenti repliche digitali 3D come nei tradizionali metodi di indagine archeologica.

## Capitolo 5

### CONCLUSIONI

#### 5.1. Organismi internazionali per la documentazione archeologica

L'archeologia è una scienza che documenta oggetti e processi materiali e immateriali che devono essere tutelati e adeguatamente salvaguardati, con le stesse modalità con le quali si assicura la conservazione e valorizzazione del bene fisico descritto. L'importanza assunta dalla documentazione archeologica nello studio dei contesti e dei singoli reperti ha favorito lo sviluppo di iniziative finalizzate a garantire la corretta acquisizione e conservazione di report e prodotti grafici e fotografici. La realizzazione di una corretta registrazione delle attività di ricerca condotte sul terreno ha un ruolo centrale nella pratica archeologica; per tali motivi, numerose istituzioni sovranazionali, impegnate nella tutela e nella valorizzazione dei beni culturali, hanno approvato e rilasciato specifiche linee guida per l'esecuzione e la protezione della documentazione archeologica.

L'attivismo di alcuni organismi internazionali negli anni '50 è, probabilmente, da collegare all'esigenza di fornire, in primo luogo alla comunità scientifica degli archeologi e dei conservatori, strumenti utili e aggiornati per seguire il dinamismo con il quale procedeva la ricostruzione post-bellica in Europa. La rapida costruzione di nuovi quartieri e impianti industriali dava impulso ad una nuova e più intensa stagione di esplorazioni archeologiche sul terreno. In assenza di una visione strategica e senza un coordinamento scientifico delle diverse attività, vennero stabiliti alcuni principi generali il cui obiettivo era quello di evitare la dispersione della conoscenza.

Le prime importanti raccomandazioni vennero introdotte in una convenzione adottata dall'UNESCO nel 1956. Oltre ad inviti generici alla costituzione di istituzioni nazionali per la tutela dei beni archeologici (art. 11, comma a) e alla conservazione della documentazione degli oggetti mobili ed immobili, la convenzione fissava l'obbligo di pubblicare i risultati degli scavi (art. 24, comma b), stabilendo che la documen-

tazione dovesse essere resa disponibile alla comunità scientifica fatti salvi i diritti degli scavatori (art. 25).

Successivamente, nel 1965 l'ICOMOS approvò la Carta di Venezia per il restauro e la conservazione di monumenti e siti. L'art. 16 della convenzione disponeva la produzione di report analitici e critici illustrati con disegni e fotografie a corredo delle attività di restauro e scavo. Nel 1990 l'ICOMOS adottò un nuovo strumento, la *Charter for the protection and management of the archaeological heritage*, che forniva linee guida ancora una volta indirizzate alle forme di gestione dello scavo e dei risultati scientifici. L'atto invitava le autorità pubbliche nazionali ad emanare provvedimenti legislativi che riconoscessero il valore della produzione, raccolta e archiviazione della documentazione archeologica.

All'incirca negli stessi anni, a La Valletta, il Consiglio Europeo approvava la Convenzione (1992) per la protezione del patrimonio archeologico che conteneva generiche indicazioni sulle responsabilità dei singoli Stati nella conservazione del patrimonio culturale nazionale. In precedenza, il Consiglio Europeo (1969) aveva ratificato una convenzione finalizzata al contrasto del traffico internazionale di oggetti d'arte e alla promozione della cooperazione e dello scambio in tema di politiche culturali sui beni archeologici.

Con il tempo le raccomandazioni disposte per la salvaguardia del patrimonio sulla terraferma furono estese alla protezione e gestione dei beni culturali sommersi. Nel 1996 l'ICOMOS rilasciò la *Charter on the protection and management of underwater cultural heritage* che stabiliva all'art. 1 la necessità di garantire una adeguata documentazione degli interventi in quanto parte integrante della progettazione (art. 2) e della metodologia (art. 5); l'art. 8, dedicato alla documentazione, precisava che

All investigation must be thoroughly documented in accordance with current professional standards of archaeological documentation.

Per la prima volta, in un documento approvato da un organismo sovranazionale si trova un esplicito riferimento alla produzione di una documentazione conforme a standard internazionali.

Sulla scia della carta dell'ICOMOS, anche l'UNESCO approvò, nel 2001, un documento specifico per la protezione del patrimonio cultura-

le sommerso. In questa convenzione la responsabilità e la cura di questa parte rilevante dell'eredità culturale diventa un tema centrale in quanto parte integrante delle tracce dell'attività umana. L'art. 2, comma 10, incoraggia la produzione della documentazione e del suo accesso, mentre l'art. 19, comma 1, invita gli Stati a cooperare nella esplorazione e descrizione dei resti sommersi. Nell'allegato relativo alle "Regole concernenti gli interventi sul patrimonio culturale subacqueo" viene precisato che ogni progetto (regola 10) deve prevedere uno specifico programma di documentazione. Le regole 26, 27 e 31 forniscono indicazioni sulle modalità di registrazione dei dati che, come l'analoga convenzione dell'ICOMOS, devono essere coerenti con gli standard professionali. La Carta sottolinea l'importanza delle fotografie, dei disegni, delle planimetrie, di note esplicative e di un inventario accurato dei ritrovamenti, mentre fornisce indicazioni (regola 33) sulle modalità di conservazione della documentazione relativa al progetto (regola 34) «in conformità con gli standard professionali internazionali».

L'ultimo strumento attuato per favorire la protezione e la promozione del patrimonio culturale europeo è la convenzione di Faro adottata dal Consiglio Europeo nel 2005, ma entrata in vigore soltanto nel 2011 e non ancora ratificata da tutti gli Stati membri. Il documento si focalizza, soprattutto, su concetti quali inclusione e partecipazione per garantire l'accesso diretto ai beni culturali. L'art. 14 è dedicato a promuovere un impegno per lo sviluppo e l'uso di strumenti digitali in funzione dell'accrescimento delle opportunità di fruizione del patrimonio culturale preservando diversità di linguaggio e culturali; in particolare, il comma b) raccomanda l'uso di standard internazionali per lo studio, la valorizzazione e la protezione del patrimonio, mentre il comma d) riconosce che la creazione di contenuti digitali non deve pregiudicare la conservazione del patrimonio esistente. Il richiamo alle nozioni di partecipazione, inclusione e comunità di eredità contiene implicitamente un sostegno alla promozione di strumenti accessibili e condivisibili che si ispirano ai principi dell'OS.

La crescente importanza delle tecnologie digitali nella riproduzione e valorizzazione del patrimonio culturale era stata già sancita nel 2003 dall'UNESCO che aveva adottato la carta per la preservazione del Digital Heritage. La convenzione non fornisce alcuna raccomandazione spe-

cifica sulla conservazione dei dati digitali, ma chiarisce che l'intero patrimonio, in qualsiasi forma esso si concretizzi (testi, immagini, database, video, etc.), deve essere considerato un capitale culturale condiviso e, quindi, accessibile al pubblico. L'UNESCO intendeva, in tal modo, incoraggiare utenti e istituzioni culturali nello sviluppo di strategie indirizzate all'identificazione di quali dati digitali conservare e, soprattutto, in che forma.

La progressiva attenzione al tema della registrazione di una adeguata documentazione, sia dei beni rinvenuti sulla terraferma che nelle realtà sottomarine, emerge distintamente dalla breve rassegna sulle attività che vari organismi internazionali hanno portato avanti in coerenza con gli obiettivi di tutela e salvaguardia del patrimonio culturale mondiale. Nello stesso tempo, si afferma un crescente interesse verso l'adozione di pratiche standard professionali, fino ad una esplicita menzione, nella convenzione di Faro, di strumenti digitali per migliorare la conoscenza. La centralità di una documentazione digitale dei beni culturali è presente anche nella convenzione UNESCO del 2003 a conferma dell'ampia diffusione raggiunta dagli strumenti computazionali per lo studio, la valorizzazione e la tutela del patrimonio culturale.

## 5.2. L'archeografia digitale

Nel volume sono state indicate le tappe più significative dell'emergere di una nuova disciplina denominata Archeologia Digitale. La definizione sostituisce la precedente espressione Archeologia Computazionale che appariva ricalcata dalla pratica di altre scienze. Per Djindjian (2019) il termine Archeologia Digitale sembra riflettere il successo che la digitalizzazione ha avuto nella accettazione dell'uso dei metodi informatici anche in archeologia. La comparsa di nuove riviste scientifiche e di congressi sul tema del rapporto tra archeologia e informatica suggeriscono l'affermazione di una nuova disciplina con un suo distinto status scientifico con la complicità significativa della diffusione delle tecnologie per il 3D che hanno modificato in larga misura l'approccio dello studioso alla creazione e valutazione della documentazione grafica.

Il riconoscimento di un settore di studi autonomo e separato dal più ampio raggruppamento delle metodologie della ricerca archeologica impone una riflessione sui principali temi di intervento dell'Archeologia

Digitale, sulle traiettorie di ricerca e sui prossimi obiettivi a medio e lungo termine. Come si è tentato di dimostrare, qualsiasi considerazione sul futuro della disciplina si lega in modo indissolubile ai progressi delle tecniche computazionali e, soprattutto, alle trasformazioni di una società sempre più pervasa dall'informatica e dalle connessioni digitali.

Due importanti prodotti editoriali, pubblicati tra il 2006 e il 2009, evidenziano i profondi cambiamenti che l'informatica ha prodotto nella economia, nella cultura e nella società globale. Nel 2006 D. Tapscott e A. D. Williams presentano il volume *Wikinomics: How Mass Collaboration Changes Everything*, mentre nel 2009 T. Hey, S. Tansley e K. Tolle curano una raccolta di articoli in un contributo dal titolo *The Fourth Paradigm: Data-Intensive scientific discovery*. Il primo lavoro illustra i benefici derivanti, nel mondo dell'impresa, dall'avvio di una collaborazione di massa costruita sulla apertura, scambio e condivisione di dati ed esperienze. Il secondo, invece, identifica nella trasformazione della rete in contenitore di documenti digitali, una straordinaria occasione per progettare un modello di *data-science* orientato alla esplorazione e manipolazione di grandi quantità di informazioni; mentre nell'ambito dei normali paradigmi scientifici occorre definire dei modelli teorici prima di poter esaminare dati sperimentali, oggi, grazie allo sviluppo di sofisticati algoritmi di *machine learning* e di AI, è possibile individuare regolarità e correlazioni nei dati presenti in rete, rendendo meno vantaggioso un approccio scientifico basato sulla definizione di teorie aprioristiche.

I due contributi, ben presto divenuti molto popolari e di riferimento per diverse comunità scientifiche, evidenziano il ruolo assunto progressivamente dalla rete che si trasforma in un contenitore di dati scientifici, economici, culturali che devono poter essere strutturati e messi in relazione alla ricerca di un ordine. Per ottenere risultati significativi e corretti è necessario che i dati siano preparati e tra loro confrontabili e, pertanto, che in un qualche modo siano standardizzati o standardizzabili.

La conoscenza non deriva, quindi, da modelli o teorie che sono elaborate dagli scienziati per spiegare fenomeni ed eventi del mondo fisico, ma da aggregazioni statistiche eseguite da particolari software; al ricercatore spetta il compito di scoprire il significato delle regolarità e delle correlazioni. Campioni di dataset sono utilizzabili per addestrare, per esempio attraverso l'uso di reti neurali, le macchine che apprendono e



che possono utilizzare la conoscenza acquisita per analizzare archivi ben più complessi.

L'estensione al mondo dell'archeologia di un approccio *data-science* solleva inevitabilmente alcune criticità connesse, in primo luogo, alla natura parziale, incompleta e ambigua della fonte archeologica e, successivamente, alla constatazione che i dati primari sono in un qualche modo strutturati in base ad un linguaggio e/o a una teoria. Il ricorso a tecniche di *data mining* o di algoritmi derivati dall'AI non può avvenire senza alcun riferimento ad un mondo reale; tuttavia, le analisi sui BD possono contribuire a comprendere al meglio correlazioni, tendenze e modelli con l'aiuto di tecniche di simulazione e modellazione.

Nella ricerca archeologica non sono stati ancora ben delineati gli scenari di un possibile impiego dell'AI o dei BD. Ad ogni modo, la semplice convergenza dati e algoritmi non è sufficiente a produrre conoscenza e per Djindjian (2020) questa prospettiva è del tutto illusoria poiché il solo accumulo di dati, per quanto ampio, non è capace di fornire conoscenze.

Per poter sfruttare gli avanzamenti delle tecniche informatiche è necessario preparare i dati, cioè formalizzarli; il processo di strutturazione dei dati rappresenta ancora oggi, anche alla luce dell'affermazione dei BD, un passaggio cruciale non solo in termini computazionali, ma anche di analisi delle metodologie di documentazione.

L'archeografia digitale avrà dunque nel futuro della disciplina un ruolo sempre più centrale, contribuendo alla progettazione di infrastrutture e repository in grado di favorire l'integrazione dei dati *digital born, legacy* e aperti. Nel corso degli ultimi anni, l'archeologia si è trasformata in una disciplina collaborativa e caratterizzata da una costante interazione con studiosi provenienti da altri settori in una prospettiva realmente interdisciplinare.

Se i dati, con il loro contenuto informativo, resistono al tempo, la trasformazione dei metodi e delle tecniche consente di guardare in modo diverso agli archivi pregressi. L'archeologia si trova oggi di fronte ad un bivio: conservare un approccio legato a forme di micro-avanzamento della conoscenza ricostruita attraverso la successione delle US, oppure provare a far emergere un nuovo modo di fare ricerca pienamente supportato dalle tecnologie computazionali, come delineato dalla prospettiva della *digital transformation*.

Questo approccio non potrà che partire da competenze maggiormente interdisciplinari, attente all'analisi dell'intero ciclo della conoscenza archeologica anche in rapporto alla progettazione di forme di comunicazione adeguate alle innovazioni imposte da una società pienamente digitale. Grazie alla rete, alla condivisione e distribuzione dei dati, emergeranno nuovi modi di organizzare, interpretare e pubblicare i record archeologici con tutta la loro criticità e parzialità senza per questo ridurne il potenziale informativo.

Tale processo, che mette al centro della disciplina architetture, metodi e teorie *data-centered* anche oltre i confini epistemologici della ricerca archeologica, spingerà a riconsiderare il ruolo che assume l'informazione alla luce dei nuovi media, tra forme ristrette di comunicazione, limitate alla comunità di riferimento, e modalità partecipative destinate ad un più ampio pubblico (Niccolucci 2020). L'uso razionale della rete e dei dati condurrà, probabilmente nel breve periodo, se non ad un cambio di paradigma a nuovi approcci. Ciò avverrà nel quadro di una rivalutazione delle tecnologie informatiche, dei nuovi saperi e, soprattutto, delle competenze professionali orientate al trattamento digitale delle informazioni. Gli archeologi saranno chiamati a riposizionare il loro ruolo nella creazione della conoscenza, passando da attori centrali ad agenti di una più vasta rete di metodi e tecnologie digitali.

Buone pratiche, policy per la creazione di archivi conformi ai FAIR data, uso intensivo di metadata, *paradata* e terminologie standard aiutano, già oggi, a mantenere e rendere accessibili gli archivi digitali, ma non incoraggiano automaticamente il loro riuso (Kansa, Kansa 2022); probabilmente è necessario aumentare gli sforzi per sviluppare migliori metodi per modellare, creare e pulire gli archivi digitali e, indipendentemente dai linguaggi e dalle tecnologie disponibili, coinvolgere con maggiore slancio la comunità scientifica nella identificazione delle pratiche più convenienti nella strutturazione dei dati di ricerca.

La diffusione delle filosofie open e di un ambito di ricerca proiettato sui dati e sulla loro pubblicazione in rete, ha contribuito ad una migliore identificazione di un nuovo approccio disciplinare, totalmente digitale, che oscilla tra una visione legata tradizionalmente all'interpretazione storica del dato e un'altra che sembra, invece, più ancorata ad un modello archeografico finalizzato alla descrizione e al trattamento del dato.

Come abbiamo visto, l'archeografia si configura come una disciplina il cui obiettivo è fornire una buona descrizione dei dati di partenza che supporti lo studioso in analisi comparative. Per Mannoni (2000) l'archeografia è il braccio destro dell'archeologia. Tuttavia, per lo studioso genovese l'archeografia è connessa, soprattutto, all'archeometria, cioè all'esito di un approccio che derivando dalle scienze dure, conferisce alla descrizione della risorsa archeologica un valore oggettivo.

La definizione di un metodo di classificazione finalizzato alla costruzione di una base comune per specialisti provenienti da vari domini e con differenti prospettive di ricerca si ritrova anche negli scopi dell'archeografia digitale. Tuttavia, mentre l'interesse di Mannoni (2002) si traduce in una visione positivista della tecnologia, l'obiettivo dei metodi digitali non si concretizza nella scelta di un sistema unico di rappresentazione e catalogazione dell'informazione, ma, piuttosto, nella adozione di una modalità flessibile indirizzata, in ultima analisi, alla integrazione e alla manipolazione di fonti eterogenee. L'informatica non pretende di imporre soluzioni unificanti, ma può facilitare la costruzione di percorsi flessibili, condivisi e realmente interdisciplinari senza che un dominio prevalga sull'altro.

L'Archeologia Digitale si propone, principalmente, di offrire ai ricercatori metodi per favorire l'integrazione e la gestione dei propri archivi, indipendentemente dalle forme di strutturazione delle banche dati.

Il computer è stato a lungo considerato un elettrodomestico utile e pratico, ma pur sempre uno strumento utilizzato dall'archeologo per risparmiare tempo e ottimizzare la gestione dei dati e non certamente per orientare nuove ricerche o ipotizzare innovative traiettorie di indagine. Oggi, al contrario, il ruolo dell'informatica è cruciale nella promozione di quella interazione tra saperi tecnico-scientifici e discipline umanistiche che contraddistingue, come abbiamo visto, la *digital transformation*.

Gli sviluppi dell'archeografia digitale, centrata sulla creazione, gestione e manipolazione del dato (testuale, geo-riferito, 3D, etc.), sollecitano una analisi approfondita delle tecniche informatiche, specialmente di quelle funzionali alla produzione di una maggiore accuratezza dei dati nella prospettiva della long-term preservation, della standardizzazione degli archivi, dell'accesso ai *legacy data* e della usabilità delle risorse online.

L'archeografia digitale contribuisce alla diffusione e al consolidamento di uno scenario che privilegia l'amministrazione, la gestione e la

trasparenza dei record (Marvick *et al.* 2017; Moscati 2017) e nel quale accesso, dati e metodi aperti alimentano una ricerca collaborativa e condivisa. Il rischio, presente nella trasformazione digitale, è che l'obiettivo di una ricerca completamente aperta e accessibile, limiti ogni riflessione sui metodi e sulla loro efficacia e, quindi, privilegiando gli aspetti di condivisione del record, sacrifichi l'adozione di più rigorosi approcci di indagine. Questa disciplina deve, dunque, candidarsi a trasformarsi in un settore di ricerca sui metodi, sulle metodologie e sulle teorie digitali approfondendo, principalmente, quel legame che unisce la semplice registrazione digitale di una informazione alla più ambiziosa ricostruzione dei processi sociali, culturali ed economici di una società del passato. Oggetto delle indagini saranno proprio quelle tecnologie digitali che hanno fortemente contribuito, negli ultimi anni, allo sviluppo di nuovi metodi di ricerca e nuovi orizzonti di indagine.

Termini come GIS, database, multimedialità, simulazione e virtualità entrano nella pratica quotidiana dei ricercatori e dei professionisti, favorendo evoluzioni normative che investono la produzione e la cura degli archivi digitali. Anche la legislazione italiana si è adeguata alle trasformazioni prevedendo che, nelle procedure di verifica preventiva dell'interesse archeologico, siano previste forme di documentazione digitale dei dati raccolti in funzione della divulgazione dei risultati, di edizioni scientifiche e didattiche e, infine, delle eventuali ricostruzioni virtuali (D'Andrea 2021).

L'informatica trasforma gli esiti materiali della ricerca in bit che, spesso, appaiono manipolabili in funzione della produzione, forma, ricezione e uso dei media. Il digitale muta l'archeologia che, da scienza distruttiva, diviene una disciplina creativa nella quale il record storico-archeologico può assumere molteplici forme comunicative (Roosvelt *et al.* 2015). L'informatica ottimizza i processi di registrazione e gestione del dato testuale e grafico, indipendentemente dai software utilizzati e dagli obiettivi di ricerca dello studioso. Tuttavia, il tema principale di una Archeologia Digitale con un suo status disciplinare è quello della comprensione della formazione del record archeologico e, quindi, della sua descrizione, documentazione e contestualizzazione. Per tali ragioni, di recente, alla tradizionale visione efficientista del computer si è sostituito un approccio più orientato ad una riflessione metodologica sulla

qualità scientifica del dato digitale piuttosto che sulla sua velocità di acquisizione ed elaborazione (Caraher 2019).

I dati restano al cuore dell'attività dell'Archeologia Digitale e dello specifico settore applicativo che chiamiamo archeografia digitale. L'osservazione, la registrazione del record e, soprattutto, la sua contestualizzazione sono aspetti cruciali che illustrano l'emergere di una nuova area di ricerca, realmente interdisciplinare, in grado di utilizzare in modo vantaggioso apporti da numerosi domini. La rete e le tecnologie del Semantic Web contribuiscono ad una innovazione disciplinare nella quale al centro del metodo resta la riflessione sulla parzialità, incompletezza e ambiguità del documento digitale.

La semplificazione dei sistemi di acquisizione, elaborazione, e gestione del dato, favorisce il passaggio da una archeologia dei modelli ad una archeologia dei dati, focalizzata sulla identificazione dei processi di formalizzazione e trattamento del record archeologico (D'Andrea 2016; 2021). Per raggiungere questo obiettivo i ricercatori dovranno abbracciare un modello di classificazione e di catalogazione basato sull'uso di metadati standard e thesauri secondo modalità conformi ai principi FAIR, sfruttando gli avanzamenti prodotti dalla tecnologia dei LOD per la creazione di connessioni e interazioni (Kansa, Kansa, Watrall 2011). Tutto ciò semplificherà l'adozione di approcci legati all'AI applicabili a grandi volumi di dati liberamente accessibili ed interrogabili.

### 5.3. I dati territoriali

Tra i settori di ricerca che hanno tratto maggiori benefici dalla diffusione dei dati digitali possiamo, senza dubbio, inserire le analisi territoriali. La recente diffusione degli Open Data cartografici, prodotti dalle Pubbliche Amministrazioni, incoraggia un approccio innovativo al dato geografico ed alla sua importazione in sistemi che gestiscono informazioni archeologiche; questa ampia disponibilità di dati territoriali modifica, in particolare, il rapporto tra le istituzioni pubbliche e la ricerca storica, favorendo una nuova cultura del territorio nella quale l'archeologia trova una sua specifica dimensione in una ottica di partecipazione e collaborazione.

La proliferazione in rete di dati territoriali e ambientali è connessa a due condizioni innovative; da un lato abbiamo avuto un incremento dei dati satellitari a grandissima risoluzione (pixel < 1 m) oppure declassifi-

cati, come i CORONA, acquisibili gratuitamente o a costi ridotti; dall'altro la produzione di dati spaziali da parte di istituzioni pubbliche o private, per scopi non direttamente legati alla ricerca, ma utilizzabili anche in ambito archeologico. Numerosi dataset cartografici sono, ad esempio, di particolare utilità nella costruzione di applicazioni legate alla ricostruzione delle forme del popolamento (modelli induttivi, deduttivi e predittivi) o per la gestione del paesaggio culturale in termini amministrativi e/o di pianificazione urbanistica.

La creazione di infrastrutture di rete e lo sviluppo di standard di condivisione delle informazioni a connotazione spaziale hanno favorito l'impiego di tematismi territoriali per lo studio del paesaggio antico. I protocolli WMS e WFM sono, oramai, standard che garantiscono l'interoperabilità e la corretta visualizzazione e distribuzione di risorse georiferite ed organizzate in strati informativi.

Tutti i software GIS possono visualizzare dati forniti da terze parti grazie ai protocolli WMS e WFM implementati in Web-GIS o *cloud GIS*, assicurando accesso, condivisione e navigazione tra i dataset spaziali pubblicati. In questo modo, ad esempio, è possibile caricare nel proprio sistema una serie di punti geodetici e scaricarne le relative monografie, oppure tematismi quali l'idrografia o l'orografia o la copertura del suolo.

La progettazione e l'implementazione di un GIS per analisi territoriali richiedeva, fino a pochi anni fa, un grosso dispendio di energie e costi connessi alla digitalizzazione della cartografia tradizionale a stampa. Oggi, al contrario, l'inserimento di informazioni geografiche di una area campione risulta una attività ordinaria semplificata grazie alla disponibilità, in larga misura gratuita, di dati in rete visualizzabili tramite protocolli di connessione.

Una ulteriore spinta al riuso dei dati geografici online viene dagli Open Data resi pubblici da numerosi Enti Locali e visualizzabili non soltanto con i protocolli WMS, ma anche come dataset scaricabili. Gli archivi possono essere aggiunti ad un progetto GIS implementato, ad esempio, per esaminare l'organizzazione dello spazio urbano e le sopravvivenze dei tracciati in epoca moderna, oppure per censire e catalogare beni culturali che sono collocati in una determinata area. I ricercatori hanno a disposizione un archivio di dati territoriali praticamente illimitato e riutilizzabile grazie a formati e protocolli standard.

La recente disponibilità di dati geografici commissionati da istituzioni pubbliche o private, insieme alla diffusione delle foto satellitari e le immagini da drone, ha accresciuto le potenzialità di ricerca nel campo dell'analisi spaziale e territoriale. Nonostante ciò, permane un atteggiamento di insufficiente attenzione ai dati presenti in rete, un atteggiamento forse connesso alla scarsa attitudine degli archeologici alla condivisione dei dati di ricerca. Resistenze, oramai antistoriche e antiscientifiche, stanno frenando la diffusione della cultura dell'OS in archeologia in Italia; parallelamente resta aperto il nodo della formazione specialistica nel campo dell'informazione spaziale che non appare ancora al centro dell'agenda accademica.

Eppure, in base alle più recenti norme sulla Archeologia Preventiva o agli atti del MIC sulle concessioni di scavo, si può facilmente dedurre come le conoscenze cartografiche siano necessarie per la realizzazione di specifici prodotti richiesti dalle normative citate. Ad esempio, per quanto riguarda la documentazione di scavo, la circolare MIC prescrive il conferimento dei dati al GNA, con i metadati relativi al dataset nella sua interezza, comprendente il report di scavo (*collection-level*) e i singoli allegati che lo compongono (*file-level*). La circolare del MIC dispone che la strutturazione dei campi sia conforme agli standard sviluppati dal progetto ARIADNE che, a loro volta, sono conformi allo standard ISO19115, richiesto dalla normativa europea INSPIRE per i files geo-spaziali. Gli archivi conferiti al GNA sono gestiti dall'ICA del Ministero attraverso una piattaforma dedicata alla raccolta, conservazione e condivisione dei dati.

Nelle intenzioni del MIC il sistema informativo dovrebbe conservare le informazioni acquisite con l'attività di tutela (assistenza alla costruzione di opere pubbliche e/o private, procedure di archeologia preventiva, rinvenimenti fortuiti) e quelle provenienti dalle ricerche universitarie e di altri istituti. Il geoportale dovrebbe, quindi, fungere da *hub* per la raccolta delle informazioni inedite attraverso standard di formato e di contenuto che garantirebbero l'interoperabilità del portale con gli altri sistemi appartenenti allo stesso Ministero, quali SigecWeb e VIR. I dati del geoportale sarebbero destinati a confluire nel RNDT, istituito presso l'autorità nazionale preposta alla conservazione dei dati (AGID). Per assicurare la corretta trasmissione ed integrazione dei dataset spaziali, i

file del geoportale sono in formato vettoriale (.shp, .dxf o .gpkg) e proiettati nel sistema di riferimento WGS84 (EPSG 4326).

L'obiettivo del Ministero è duplice: da un lato costruire un sistema standard per la raccolta e pubblicazione Open Access dei dati a connotazione spaziale e, dall'altro, rendere immediatamente disponibili e fruibili gli archivi garantendone, nello stesso tempo, la proprietà intellettuale. Rileggendo le circolari del Ministero e le norme, ci si dovrebbe interrogare sulle competenze degli archeologi, giovani e meno giovani, e comprendere quanto i professionisti e i ricercatori siano in grado di assolvere senza difficoltà agli obblighi previsti dai regolamenti. Una tale considerazione deve necessariamente innescare una profonda riflessione sul tema più generale della preparazione degli archeologi in settori oramai diffusi dell'Archeologia Digitale, in primo luogo quello delle analisi spaziali e dei GIS.

Il quesito è stato, come abbiamo segnalato in precedenza (*supra* 41), sollevato da Carandini (2008) che ha proposto la costruzione di sistemi informativi archeologici unitari, statali, regionali e universitari, capaci di condividere requisiti e procedure essenziali. Purtroppo, alle giuste osservazioni di Carandini non è seguita alcuna innovazione né a livello di formazione di base e professionale dei giovani archeologi, né sul piano della progettazione di sistemi territoriali condivisi. Il GNA, che resta l'unico progetto finora elaborato, è ancora in fase di sviluppo.

Tranne poche eccezioni, si deve registrare, almeno per l'ambito italiano, una preparazione professionale basata in larga misura sulle tradizionali competenze storico-artistiche, con una marginalizzazione dei percorsi professionalizzanti e tecnici. Il divario tra una conoscenza classica ed una tecnica potrà in parte essere colmato attraverso lo sviluppo di settori della *digital innovation* grazie anche ai significativi investimenti previsti dal recente PNRR.

#### 5.4. I dati 3D

Un altro settore che produce molteplici tipologie di dati, che andranno conservati e mantenuti nel lungo periodo, è quello del 3D. L'impiego di tecniche di laser scanning, di fotogrammetria terrestre e aerea ha trasformato il consueto approccio alla documentazione, spesso, acquisita direttamente sul terreno dopo l'identificazione degli elementi da rap-



presentare graficamente (Bosco, D'Andrea 2019; Bosco 2022). Oggi, la fase di raccolta dei dati grafici solleva nuovi problemi e criticità connessi al processo di discretizzazione, cioè alla estrazione di forme e concetti definiti all'interno del *continuum* di coordinate spaziali registrate in una nuvola di punti.

Le metodologie 3D *reality-based* sono state accolte in ambito archeologico nell'ottica multidisciplinare di un miglioramento della fase di documentazione finalizzata a registrare e monitorare lo stato dell'arte di siti, monumenti ed oggetti, nonché a supportare l'interpretazione scientifica, il restauro o l'analisi strutturale dei beni archeologici (Limoncelli 2012; Bosco, D'Andrea 2019). Il rilevatore sul campo non deve vedere e comprendere l'oggetto che disegna e rappresenta, bensì deve possedere le competenze per assicurare l'accuratezza e la precisione di una acquisizione digitale che garantisca la possibilità successiva di una corretta elaborazione grafica dalla quale estrarre oggetti archeologici finiti.

Il rilievo digitale non risponde ad una precisa domanda scientifica, né documenta uno specifico contesto di studio; esso viene adoperato perché, con la sua ricchezza di particolari e dettagli, contribuisce ad una riflessione post-scavo che consente una lettura e rilettura dell'oggetto/contesto archeologico in tutte le sue dimensioni e non soltanto in quelle proiettate in 2D, come le tradizionali rappresentazioni grafiche e le fotografie. Tuttavia, l'estrema semplicità di impiego degli strumenti di acquisizione e processamento dei dati non sempre favorisce lo studioso perché la rapida elaborazione di un oggetto tridimensionale è, spesso, realizzata a svantaggio dei necessari controlli geometrici e topografici che risultano fondamentali per la creazione di una corretta replica digitale; l'assenza di informazioni sui parametri come risoluzione e verifica sul terreno rende l'elaborato finale una bella immagine che abbellisce le stratigrafie (verticali e/o orizzontali), ma non può essere considerato un vero documento scientifico.

Una inadeguata valutazione degli aspetti più direttamente connessi alla precisione ed alla accuratezza del rilievo, può indebolire la correttezza metrica della restituzione finale fornendo, nella migliore delle ipotesi, dati non verificati con misure eseguite con idonei sistemi topografici (D'Andrea *et al.* 2018, 2020). Spesso, il concetto di accuratezza di un disegno è confuso con la nozione di dettaglio; mentre l'accuratezza mi-

sura il grado di corrispondenza del dato teorico con il dato misurato, il dettaglio equivale alla quantità di punti rilevati.

Generalmente la fedeltà geometrica di un rilievo digitale è determinata, in modo improprio, dalla quantità di foto o scansioni acquisite, mentre, al contrario, l'affidabilità di un disegno dovrebbe basarsi su misure di controllo. I rilevatori basano l'accuratezza di un rilievo sulle caratteristiche di campionamento del laser scanner o sulla risoluzione pixel della macchina fotografica; tuttavia, per validare la geometria di un disegno è necessaria una fase di controllo eseguita con il rilevamento di punti a terra misurati con la stazione totale. La correttezza di sezioni, piante e prospetti estratti dalla nuvola di punti può essere compromessa dalla mancata verifica della geometria o dalla errata roto-traslazione del rilievo.

L'assenza di informazioni sul processo di acquisizione, elaborazione e verifica geometrica conclusiva del rilievo rischia di pregiudicare il corretto riutilizzo del dato. Come abbiamo menzionato in precedenza occorre definire un sistema di metadati e *paradata* che conservino informazioni precise sulla *provenance*, ovvero sul processo tecnico della digitalizzazione, e sulle motivazioni alla base della realizzazione della replica digitale, incluse le finalità (scientifiche, didattiche o comunicative) e le ipotesi ricostruttive.

Poiché un oggetto digitale tridimensionale può essere il risultato della integrazione di differenti tecniche di acquisizione, la registrazione dei *paradata* è importante per evidenziare quali elementi possano aver condizionato la creazione della risorsa 3D e la sua interpretazione e ricostruzione.

Le repliche tridimensionali, quindi, indipendentemente dalla complessità delle strumentazioni impiegate, non possono essere esaminate e correttamente riutilizzate senza poter disporre delle informazioni relative all'intero processo di digitalizzazione. Il ricercatore deve, quindi, annotare tutte le informazioni necessarie ad una completa interpretazione del record digitale. Per dare una risposta all'esigenza di una maggiore comprensione e trasparenza dei modelli 3D online, il progetto europeo 3DICONs (*supra* 158) ha proposto uno schema di metadati basato sulla definizione della *provenance*, come richiamato dal W3C *Provenance Incubator Group*, e dei *paradata*, secondo i principi contenuti nella London Charter (D'Andrea, Fernie 2013; D'Andrea 2016).

### 5.5. Archeologia e *data-science*

I dati, in conclusione, qualsiasi sia la loro finalità e produzione, costituiscono una risorsa insostituibile della ricerca archeologica passata, presente e futura. In un discorso presentato nel 2006 alla conferenza annuale dell'*Association of National Advertisers*, il matematico inglese Clive Humby sottolineò che “*Data is the new oil*”.

Analogamente al petrolio, i dati devono essere raffinati, scomposti e analizzati per avere valore, altrimenti risultano inutilizzabili. Perciò, anche se i dati sono abbondanti, non sono sempre della stessa qualità: possono essere incompleti, parziali, inesatti oppure celare pregiudizi o tesi bizzarre. Mentre il petrolio ha contribuito allo sviluppo socioeconomico mondiale tra il XIX ed il XX secolo, sono le connessioni, le tecnologie ed i dati a svolgere, oggi, questa fondamentale funzione di crescita. I dati promuovono il progresso sociale, creando, nello stesso tempo, nuovi modelli economici e rinnovando i tradizionali approcci professionali.

Negli ultimi anni numerose e differenti motivazioni hanno spinto gli studiosi ad intraprendere nuovi percorsi per l'elaborazione, la gestione e la cura dei dataset archeologici. La crescita delle indagini archeologiche sul campo, a cui si sono aggiunte le attività della archeologia commerciale, ha generato un aumento esponenziale dei dati, richiedendo nuovi approcci e sistemi per archiviare, salvaguardare e, infine, sfruttare questa fonte di informazioni apparentemente infinita. I risultati delle esplorazioni devono essere divulgati e diffusi in modo appropriato anche secondo i principi della Archeologia Pubblica che estendono i tradizionali confini del pubblico di ricercatori, professionisti e appassionati.

Lo sviluppo delle indagini archeologiche e la loro divulgazione, principalmente attraverso i nuovi media e i social (Goodall, Roberts 2019), ha determinato un ampliamento della comunità scientifica. Mentre il secolo scorso era caratterizzato dalla presenza di una ristretta comunità archeologica, limitata grosso modo all'Occidente, negli anni più recenti, con gli avanzamenti tecnologici e professionali di gran parte dei Paesi in via di sviluppo, la ricerca archeologica ha conosciuto una nuova stagione di iniziative e scavi condotti da archeologi locali e autorità nazionali. L'emergere di una comunità scientifica mondiale solleva numerose criticità, tra le quali come gestire il potenziale informativo contenuto nei

documenti archeologici restituiti dalle ultime ricerche. Nello stesso tempo, emerge l'esigenza di una rivalutazione della documentazione pregressa che spinge i ricercatori verso un nuovo modo di rileggere gli archivi archeologici più antichi.

Il mondo digitale è apparso, da subito, uno strumento affidabile e potente nel quale integrare tutta la conoscenza disponibile indipendentemente dalle diversità culturali, linguistiche e, per certi aspetti, metodologiche e teoretiche ingabbiate nella documentazione utilizzabile.

Negli anni '80 dello scorso secolo alcune importanti iniziative si sono concentrate sulla conversione digitale dei dati dando origine ad un successivo fenomeno, potremo dire, quasi di massa, caratterizzato dalla produzione digitale di ogni archivio archeologico. All'inizio, questo processo è stato portato avanti da gruppi di ricerca e autorità nazionali che hanno selezionato un proprio specifico approccio con software e formati proprietari e una rappresentazione dei dati generata sulla base di specifiche linee di ricerca.

L'assenza di un coordinamento tra i vari progetti ha prodotto una sorta di torre di Babele digitale dove le applicazioni, in larga misura, non comunicano tra loro e i dati restano circoscritti ad uno specifico gruppo di ricerca. Nonostante ciò, la conversione analogica-digitale degli archivi e l'emergere di processi tecnici esclusivamente digitali hanno stimolato nuove riflessioni metodologiche, soprattutto in funzione della necessaria condivisione dei dati.

Le problematiche relative alla mancanza di integrazione tra i diversi archivi digitali, connesse all'esigenza di superare la frammentazione nella ricerca archeologica, sono state di recente riesaminate soprattutto alla luce dei progressi del Web e del settore delle tecnologie legate alla rappresentazione ed integrazione dell'informazione digitale. L'introduzione di standard tecnici ha supportato una prima, ma utile, integrazione sintattica limitata ai formati dei file e ai protocolli di comunicazione. Successivamente la ricerca ha esteso le forme della interoperabilità includendo elementi e concetti semantici.

L'obiettivo della integrazione fisica e semantica tra gli archivi non è stato raggiunto progettando un super-standard o un super-database, che avrebbe certamente incontrato l'opposizione tenace dei ricercatori, ma attraverso la definizione di nuovi strumenti concettuali flessibili in

grado di garantire l'unione dei documenti digitali senza, però, modificare le strutture degli archivi digitali originali.

Lo sviluppo del Web e delle tecnologie semantiche ha modificato le forme di ricerca e visualizzazione dei dati culturali producendo un miglioramento delle tecniche di *data* e *knowledge mining*. Nello stesso tempo, la definizione di nuovi linguaggi per strutturare la conoscenza (ontologie) e i dati (metadati, *paradata* e thesauri) ha contribuito allo sviluppo di un nuovo modo di fare scienza basato su dati aperti, accessibili e riutilizzabili. Lo scenario generato da un cambiamento, che è insieme tecnico e metodologico, è stato diversamente interpretato; l'affermazione più appropriata coniata per descrivere questo mutamento è "quarto paradigma" (Hey, Tansle, Tolle 2009).

Anche l'archeologia è entrata nel mondo del *data science*, una nuova scienza che analizza i principi metodologici e multidisciplinare relativi all'estrazione ed alla interpretazione dei dati provenienti da varie fonti. Ma, per poter sfruttare interamente i vantaggi dei nuovi paradigmi scientifici, occorre che il ricercatore sia consapevole del proprio ruolo di produttore e consumatore di informazioni. Dati digitali, cartacei e *legacy* contribuiscono ad una nuova fase di accumulazione non capitalistica, ma di risorse informative. Lo scenario dell'OS, lo sviluppo delle tecnologie del Semantic Web, dell'AI e delle infrastrutture di ricerca, non sono una visione futuribile, ma una realtà attuale e concreta che dovrà in un futuro non troppo lontano incoraggiare la crescita di una nuova archeologia dei dati.

La raccolta automatica di dati digitali può diventare un compito facile e veloce che supera le nostre capacità di convalidare, analizzare e visualizzare informazioni e archivi. La ricerca appare ottimizzata da un pensiero computazionale che potenzia la rete trasformandola in un grande amplificatore intellettuale universale. Nonostante il forte ottimismo nei confronti nei progressi della scienza, non è ancora chiaro come questo neopositivismo rinnoverà il processo di scoperta scientifica. Senza dubbio, una scienza ad alta intensità di dati produrrà un sostanziale cambiamento nel modo di fare scienza, contribuendo alla definizione di uno scenario nel quale risorse accademiche, testi, banche dati e altro materiale associato saranno navigabili e integrabili senza soluzione di continuità.

Il tema stimolerà un maggiore l'interesse per il *data-sharing* anche nel mondo dell'archeologia, non soltanto quella digitale, favorendo una riflessione sui metodi e sulle pratiche di ricerca incluse le procedure di migrazione degli archivi digital-born e *legacy data*.

Il focus di questa trasformazione, che non potrà essere rallentata da resistenze anacronistiche, è la comprensione della conoscenza implicita che si cela nei processi archeografici. Questo sarà il principale terreno di studio dell'Archeologia Digitale intesa come disciplina che esplora l'uso pratico del computer e delle tecniche informatiche nella ricerca storico-archeologica. Gli strumenti digitali non sono semplici elettrodomestici che sostituiscono l'attività umana in lavori meccanici e ripetitivi, ma rappresentano innovazioni che estendono le facoltà corporee e sensoriali dei ricercatori. L'analisi sulle pratiche di scavo e sulle modalità di registrazione della documentazione si sposterà sempre di più dall'idea di una ricerca archeologica sul campo come risultato di processo distruttivo della realtà fisica e materiale a quella più creativa della produzione di dati e dei relativi media digitali.



## ABBREVIAZIONI

AAT	<i>Art and Architecture Thesaurus</i>
ABIM/ArcheoBIM	<i>Archaeological Building Information Modelling</i>
ABM	<i>Agent-Based Modelling</i>
ACDH-CH	<i>Austrian Centre for Digital Humanities and Cultural Heritage</i>
AGID	<i>Agenzia per l'Italia Digitale</i>
AI	<i>Artificial Intelligence</i>
ADS	<i>Archaeology Data Service</i>
AGNES	<i>Archaeological Grey-Literature Named Entity Search</i>
BARTOC	<i>Basic Register of Thesauri, Ontologies and Classification</i>
BBT	<i>BackBone Thesaurus</i>
BD	<i>Big Data</i>
BIM	<i>Building Information Modelling</i>
CAD	<i>Computer-Aided Design</i>
CC-BY	<i>Creative Commons Attribution</i>
CC-BY-SA	<i>Creative Commons Attribution ShareAlike</i>
CC-BY-NC-ND	<i>Creative Commons Attribution-Non-Commercial-No-Derivative</i>
CIDOC	<i>International Committee for Documentation</i>
CIDOC-CRM	<i>International Committee for Documentation – Conceptual Reference Model</i>
CLARIN	<i>Common Language Resources and Technology Initiative</i>
CONA	<i>Cultural Objects Name Authority</i>
DAI	<i>Deutsches Archäologisches Institut</i>
DANS	<i>Data Archiving and Networked Services</i>
DARIAH	<i>Digital Research Infrastructure for the Arts and Humanities</i>
DOI	<i>Digital Object Identifier</i>
EAC	<i>Europae Archaeologiae Consilium</i>
EDLnet	<i>Europeana Digital Library</i>
EDM	<i>European Data Model</i>
EOSC	<i>European Open Science Cloud</i>
EPI-DOC	<i>Epigraphic Documents</i>
EPSG	<i>European Petroleum Survey Group</i>



ESE	<i>Europeana Semantic Elements</i>
EXIF	<i>EXchangeable Image File</i>
e-RIHS	<i>European Research Infrastructures for Heritage Science</i>
FAIMS	<i>Field Acquired Information Management Systems</i>
FAIR	<i>Findable, Accessible, Interoperable, Reusable</i>
FISH	<i>Forum on Information Standards in Heritage</i>
FORCE11	<i>the Future of Research Communication and E-Scholarship</i>
GIS	<i>Geographic Information System</i>
GML	<i>Geographic Markup Language</i>
GN	<i>Geoportale Nazionale</i>
GNA	<i>Geoportale Nazionale per l'Archeologia</i>
GNSS	<i>Global Navigation Satellite System</i>
GPS	<i>Global Positioning System</i>
HBIM	<i>Historical Building Information Modelling</i>
HIDA-MIDAS	<i>Hierarchischer Dokument Administrator - Marburger Inventarisations, Dokumentations und Administrationssystem</i>
HTML	<i>HyperText Markup Language</i>
HTTP	<i>HyperText Transfer Protocol</i>
KML	<i>Key Markup Language</i>
KOS	<i>Knowledge Organization Systems</i>
IA	<i>Iconography Authority</i>
ICA	<i>Istituto Centrale per l'Archeologia</i>
ICCD	<i>Istituto Centrale per il Catalogo e la Documentazione</i>
ICOM	<i>International Council of Museums</i>
ICOMOS	<i>International Council on Monuments and Sites</i>
IFC	<i>Industry Foundation Classes</i>
IIIF	<i>International Interoperability Image Framework</i>
INITMA	<i>Informazione Territoriale e Monitoraggio Ambientale</i>
INSPIRE	<i>Infrastructure for Spatial Information in Europe</i>
IOGP	<i>International Association of Oil and Gas Producers</i>
IoT	<i>Internet of Things</i>
IGO	<i>Intergovernmental Organizations</i>
IPR	<i>Intellectual Property Rights</i>
ISNI	<i>International Standard Name Identifier</i>
ISO	<i>International Organization for Standardization</i>
LCP	<i>Levantine Ceramics Project</i>
LIMC	<i>Lexicon Iconographicum Mythologiae Classicae</i>

LOD	<i>Linked Open Data</i>
MAPPA	Metodologie Applicate alla Predittività del potenziale Archeologico
MASA	<i>Mémoires des Archéologiques et des Sites Archéologiques</i>
MIC	Ministero della Cultura
MET	<i>Multilingual Egyptian Thesaurus</i>
MOD	MAPPA Open Data
NGA/NIMA	<i>National Geospatial-Intelligence Agency / National Imagery and Mapping Agency</i>
NHB	<i>National Heritage Board - Swedish National Heritage Archive</i>
NISO	<i>National Information Standards Organization</i>
NRHE	<i>National Record of the Historic Environment</i>
OAI-PMH	<i>Open Archive Initiative – Protocol for Metadata Harvesting</i>
ODBL	<i>Open Database License</i>
ODC-By	<i>Open Data Commons Attribution License</i>
OGC	<i>Open Geospatial Consortium</i>
OS	<i>Open Science</i>
OSIG	<i>American Archaeology Open Science Interest Group</i>
OSGeo	<i>Open-Source Geospatial Foundation</i>
OWL	<i>Ontology Web Language</i>
PAS	<i>Portable Antiquities Scheme</i>
PCN	Portale Cartografico Nazionale
PI/PID	<i>Persistent Identifier</i>
PNRR	Piano Nazionale di Ripresa e Resilienza
RCE	<i>Cultural Heritage Agency</i>
RDF	<i>Resource Description Framework</i>
RDFS	<i>Resource Description Framework Schema</i>
RNDT	Repertorio Nazionale dei Dati Territoriali
SDI	<i>Spatial Digital Infrastructures</i>
SKOS	<i>Simple Knowledge Organization System</i>
SIGEC	Sistema Informativo Generale del Catalogo
SND	<i>Swedish National Data Service</i>
TGN	<i>Thesaurus of Geographic Names</i>
TEI	<i>Text Encoding Initiative</i>
THOT	<i>Thesauri and Ontology</i>
ULAN	<i>Union List of Artist Names</i>
URI	<i>Uniform Resource Identifier</i>

URL	<i>Uniform Resource Locator</i>
USGS-GNIS	<i>United States Geological Survey – Geographic Names Information System</i>
XML	<i>eXtensible Markup Language</i>
VIAF	<i>Virtual International Authority File</i>
VIR	<i>Vincoli in Rete</i>
W <sub>3</sub> C	<i>World Wide Web Consortium</i>
WGS84	<i>World Geodetic System 1984</i>
WCS	<i>Web Coverage Server</i>
WFS	<i>Web Feature Server</i>
WMC	<i>Web Map Content</i>
WMS	<i>Web Mapping Server</i>

## BIBLIOGRAFIA

### *3D Content in Europeana 2020*

*3D Content in Europeana*. Task force Europeana Network Association Members Council. [<https://pro.europeana.eu/project/3d-content-in-europeana>].

### Aloia *et al.* 2017

N. Aloia, F. Debole, A. Felicetti, I. Galluccio, M. Theodoridou, Mapping the ARIADNE Catalog Data Model to CIDOC CRM: Bridging Resource Discovery and Item-Level Access, *SCientific RESearch and Information Technology* 7(1), 1-8. [[dx.doi.org/10.2423/i22394303v7n1p1](https://doi.org/10.2423/i22394303v7n1p1)].

### Allen, Green, Zubrow 1990

K.M. Allen, S. Green, E.B.W. Zubrow, *Interpreting Space: GIS and Archaeology*. London.

### Allepuz *et al.* 2021

T. Allepuz, P. Gordo, S. Medina, P. del Fresno Bernal, J. Vicens Tarré, A. Mauri Martí, Towards an Ontology-Driven Information System for Archaeological Pottery Studies: The GREYWARE Experience, *Applied Sciences* 11, 7989. [[doi.org/10.3390/app11177989](https://doi.org/10.3390/app11177989)].

### Aspöck, Geser 2014

E. Aspöck, G. Geser, What is an Archaeological Research *Infrastructure* and Why Do We Need it? Aims and Challenges of ARIADNE. In *Proceedings of the 18th International Conference on Cultural Heritage and New Technologies*, 1-10. Vienna. [[https://archiv.chnt.at/wp-content/uploads/Aspoeck\\_Geser\\_2014.pdf](https://archiv.chnt.at/wp-content/uploads/Aspoeck_Geser_2014.pdf)].

### Baca 2016

M. Baca (ed.), *Introduction to Metadata*. Los Angeles (3<sup>rd</sup> ed.).

### Barbuti 2021

N. Barbuti, Thinking Digital Libraries for Preservation as Digital Cultural Heritage: by R to R4 Facet of FAIR Principles, *International Journal of Digital Libraries* 22, 309-318. [[doi.org/10.1007/s00799-020-00291-7](https://doi.org/10.1007/s00799-020-00291-7)].

*Basic Principles* 2020

*Basic Principles and Tips for 3D Digitisation of Tangible Cultural Heritage for Cultural Heritage Professionals and Institutions and Other Custodians of Cultural Heritage* (Expert Group on Digital Cultural Heritage and Europeana). [digital-strategy.ec.europa.eu/en/library/basic-principles-and-tips-3d-digitisation-cultural-heritage].

## Beale, Reilly 2017

G. Beale, P. Reilly, Digital Practice as Meaning Making in Archaeology, *Internet Archaeology* 44. [doi.org/10.11141/ia.44.13].

## Bentkowska-Kafel, Denard, Baker 2012

A. Bentkowska-Kafel, H. Denard, D. Baker (eds.), *Paradata and Transparency in Virtual Heritage*. Surrey.

## Bevan 2015

A. Bevan, The Data Deluge, *Antiquity* 89, 1473-1484.

Borgo *et al.* 2022

S. Borgo, R. Ferrario, A. Gangemi, N. Guarino, C. Masolo, D. Porello, E.M. Sanfilippo, L. Vieu, DOLCE: A Descriptive Ontology for Linguistic and Cognitive Engineering, *Applied Ontology* 17(1), 45-69. [doi.org/10.3233/AO-210259].

Börjesson *et al.* 2022

L. Börjesson, O. Skö, Z. Friberg, D. Löwenborg, G. Palsson, I. Huvila, Re-purposing Excavation Database Content as Paradata. An Explorative Analysis of Paradata Identification Challenges and Opportunities, *KULA: Knowledge Creation, Dissemination, and Preservation Studies* 6(3). [doi.org/10.18357/kula.221].

## Bosco 2022

A. Bosco, *3D Surveying Methods and Digital Information Management for Archaeological Heritage* (BAR International Series 3091). Oxford.

## Bosco, D'Andrea 2019

A. Bosco, A. D'Andrea, La documentazione grafica per lo studio dei beni archeologici attraverso la replica digitale 3D. In P. Belardi (ed.), *Riflessioni/Reflections: l'arte del disegno/il disegno dell'arte* (Atti del 41° convegno internazionale dei Docenti delle discipline della Rappresentazione, Perugia), 425-430. Roma. [doi.org/10.36165/1144].

Bosco *et al.* 2018a

A. Bosco, L. Carpentiero, A. D'Andrea, E. Minucci, F. Pesando, R. Valentini, Nuove indagini nell'isolato 4-6 di Paestum, *Newsletter di Archeologia CISA* 9, 165-192.

Bosco *et al.* 2018b

A. Bosco, A. D'Andrea, M. Nuzzolo, R. Pirelli, P. Zanfaglia, A virtual reconstruction of the Sun Temple of Niuserra: From Scans to ABIM. In M. Matsumoto, E. Uleberg (eds.), *CAA2016: Oceans of Data (Proceedings of the 44<sup>th</sup> Conference on Computer Applications and Quantitative Methods in Archaeology)*, 377-388. Oxford.

Bosco *et al.* 2021

A. Bosco, L. Carpentiero, A. D'Andrea, E. Minucci, R. Valentini, Developing an ABIM System: A New Prospective for Archaeological Data Management, *Archeologia e Calcolatori* 32(2) 167-176. [doi: 10.19282/ac.32.2.2021.15].

Bradley 1997

R. Bradley, To See is to Have Seen. Craft Traditions in British Field Archaeology. In B.L. Molyneaux (ed.), *The Cultural Life of Images: Visual Representation in Archaeology*, 62-72. London-New York.

Brandhorst 1993

J.P.J. Brandhorst, Quantifiability in Iconography, *Knowledge Organization* 20(1), 12-19.

Brandsen, Lippok 2021

A. Brandsen, F. Lippok, A Burning Question: Using an Intelligent Grey Literature Search Engine to Change our Views on Early Medieval Burial Practices in the Netherlands, *Journal of Archaeological Science* 133. [doi:10.1016/j.jas.2021.105456].

Brogiolo 1997

G.P. Brogiolo, Archeologia e Istituzioni: statalismo o policentrismo?, *Archeologia Medievale* 24, 7-30.

Brogiolo 2012

G.P. Brogiolo, Archeologia pubblica in Italia: quale futuro?, *Post-Classical Archaeologies* 2, 269-278.

Calandra *et al.* 2021

E. Calandra, V. Boi, A. Falcone, V. Acconcia, S. Di Giorgio, S. Massara, P. Ronzino, Policy and Practice for Digital Archaeological Archiving in Italy, *Internet Archaeology* 58. [doi.org/10.11141/ia.58.27].

Carandini 2000

A. Carandini, *Storia dalla terra. Manuale di scavo archeologico*. Torino.

Carandini 2008

A. Carandini, *Archeologia classica. Vedere il tempo antico con gli occhi del 2000*. Torino.

Caraher 2019

W. Caraher, Slow Archaeology, Punk Archaeology, and the 'Archaeology of Care', *European Journal of Archaeology* 22(3), 372-385. [doi:10.1017/ea.2019.15].

Casarotto 2022

A. Casarotto, Digitising Legacy Field Survey Data: A Methodological Approach Based on Student Internships, *Digital* 2(4), 422-443.

Castelli, Felicetti 2019

L. Castelli, A. Felicetti, Modelli e strumenti semantici per la documentazione dell'Heritage Science. In *Atti della Conferenza GARR 2018* (Cagliari 3-5 ottobre 2018). Roma. [doi.org//10.26314/GARR-Conf18-proceedings-10].

Champion, Rahaman 2019

E. Champion, H. Rahaman, 3D Digital Heritage Models as Sustainable Scholarly Resources, *Sustainability* 11(8), 1-8. [doi.org/10.3390/su11082425].

Champion, Rahaman 2020

E. Champion, H. Rahaman, Survey of 3D Digital Heritage Repositories and Platforms, *Virtual Archaeology Review* 11(23), 1-15. [doi.org/10.4995/var.2020.13226].

Chapman 2001

H. Chapman, Understanding and Using Archaeological Topographic Surveys: The 'Error Conspiracy'. In Z. Stančič, T. Veljanovski (eds.), *Computing Archaeology for Understanding the Past, CAA 2000, Computer Applications and Quantitative Methods in Archaeology*, 19-24. Oxford.

Chenhall 1967

R.G. Chenhall, The Impact of Computers on Archaeological Theory: An Appraisal and Projection, *Computers and the Humanities* 3, 15-24.

Clark 1968

D.L. Clark, *Analytical Archaeology*. London.

Clark 1972

D.L. Clark, *Models in Archaeology*. London.

Conolly, Lake 2006

J. Conolly, M. Lake, *Geographical Information Systems in Archaeology*. Cambridge.

Consiglio Europeo 1969

Consiglio Europeo 1969, *Convenzione europea per la protezione del patrimonio archeologico*. Londra. [<https://rm.coe.int/1680072318>].

Consiglio Europeo 1992

Consiglio Europeo 1992, *Convenzione europea per la protezione del patrimonio archeologico*. La Valletta. [<https://rm.coe.int/168007bd45>].

Consiglio Europeo 2005

Consiglio Europeo 2005, *Convenzione di Faro*. Faro. [<https://rm.coe.int/la-convenzione-di-faro-la-via-da-seguire-per-il-patrimonio-culturale/1680a11087>].

Cowgill 1968

G. Cowgill, Computer Applications in Archaeology, *Computers and the Humanities* 2(1), 17-23.

D'Andrea 2003

A. D'Andrea, *Documentazione Archeologica, Standard e trattamento digitale*. Budapest.

D'Andrea 2015a

A. D'Andrea, 3DICONs: Oggetti digitali archeologici 3D. In P. Ronzino (a c.), *L'integrazione dei dati archeologici digitali. Esperienze e prospettive in Italia* (Atti del Workshop, Lecce, Italia, 1-2 ottobre), 63-70. [<https://ceur-ws.org/Vol-1634/paper8.pdf>].



## D'Andrea 2015b

A. D'Andrea, Dati digitali e metodologia della ricerca archeologica. In P. Ronzino (a c.), *L'integrazione dei dati archeologici digitali. Esperienze e prospettive in Italia* (Atti del Workshop, Lecce, Italia, 1-2 ottobre), 10-17. [<https://ceur-ws.org/Vol-1634/paper2.pdf>].

## D'Andrea 2016

A. D'Andrea, Dall'archeologia dei modelli all'archeologia dei Dati, *AION Annali di Archeologia e Storia Antica* 19-20, 303-318.

## D'Andrea 2021

A. D'Andrea, Tra archeologia globale e archeologia 2.0: Il ruolo dell'informatica nella ricerca archeologica. In Tiziano Mannoni. *Attualità e sviluppi di metodi e idee*, vol. 2, 686-692. Sesto Fiorentino (FI).

## D'Andrea 2022a

A. D'Andrea, Dal CAD al GIS: criticità nel riuso di *legacy data* archeologici, *Archeologie tra Oriente e Occidente* 1, 35-47.

## D'Andrea 2022b

A. D'Andrea, Chi ha paura dei metodi informatici in archeologia, *Il Capitale Culturale* 25, 269-288.

## D'Andrea 2022c

A. D'Andrea, L'archeografia digitale: dalla illustrazione alla visualizzazione scientifica, *AION Annali di Archeologia e Storia Antica* 29, 199-213.

## D'Andrea, Fernie 2013

A. D'Andrea, K. Fernie, CARARE 2.0: A Metadata Schema for 3D Cultural Objects. In *Digital Heritage International Congress* Vol. 2, 137-143. Marsiglia. [<https://ieeexplore.ieee.org/document/6744745>].

## D'Andrea, Niccolucci 2001

A. D'Andrea, F. Niccolucci, L'informatica dell'archeologo. Alcune istruzioni per l'uso, *Archeologia e Calcolatori* 12, 199-220.

## D'Andrea et al. 2018

A. D'Andrea, A. Coralini, A. Bosco, A. Fiorini, R. Valentini, A 3D Topographic Network for the Study and Maintenance of the Insula III of Herculaneum. In *IEEE International Workshop on Metrology for Archaeology and Cultural Heritage (MetroArchaeo)*. [[doi.org/10.1109/MetroArchaeo43810.2018.9089809](https://doi.org/10.1109/MetroArchaeo43810.2018.9089809)].

D'Andrea *et al.* 2020

A. D'Andrea, A. Coralini, A. Bosco, A. Fiorini, R. Valentini, Ercolano. Documentare e studiare siti archeologici storici: il contributo delle tecnologie digitali, *FOLD&R*, 1-28.

D'Andrea *et al.* 2022

A. D'Andrea, M. Conyers, K.K. Courtney, E. Finch, M. Levine, Copyright and Legal Issues Surrounding 3D Data. In Moore *et al.* 2022, 205-257.

David, Spence 2003

P. David, M. Spence, *Towards Institutional Infrastructures for eScience: The Scope of the Challenges* (Oxford Internet Institute Report 2). Oxford. [dx.doi.org/10.2139/ssrn.1325240].

Davies *et al.* 2019

B. Davies, I. Romanowska, K. Harris, S.A. Crabtree, Combining Geographic Information Systems and Agent-Based Models in Archaeology: Part 2 of 3, *Advances in Archaeological Practice* 7(2), 185-193. [doi.org/10.1017/aap.2019.5].

Dell'Unto, Landeschi 2022

N. Dell'Unto, G. Landeschi, *Archaeological 3D GIS*. London–New York.

Di Giorgio, Ronzino 2019

S. Di Giorgio, P. Ronzino, Un modello per la compilazione di Data Management Plan per l'archeologia. In *Atti della Conferenza GARR 2018*, (Cagliari 3-5 ottobre 2018). Roma. [doi.org/10.26314/GARR-Confi18-proceedings-07].

Digital Science in Horizon 2020

*Digital Science in Horizon* ec.europa.eu/newsroom/dae/document.cfm?doc\_id=2124.

Djindjian 1996

F. Djindjian, Méthode archéologique assistée par ordinateur, *Archeologia e Calcolatori* 7, 1259-1266.

Djindjian 2019

F. Djindjian, Archaeology and Computers: A Long Story in the Making of Modern Archaeology, *Archeologia e Calcolatori* 30, 13-20.

Djindjian 2020

F. Djindjian, Big Data and Archaeology, *Humanities Arts and Society Magazine* 1, 208-217.

Doerr 2003

M. Doerr, The CIDOC CRM: An Ontological Approach to Semantic Interoperability of Metadata, *AI Magazine* 24(3), 75-92.

Dore, Murphy 2012

C. Dore, M. Murphy, Integration of Historic Building Information Modeling (HBIM) and 3D GIS for Recording and Managing Cultural Heritage Sites. In *Proceedings of the 18th International Conference on Virtual Systems and Multimedia (VSMM 2012: Virtual Systems in the Information Society, Milan, Italy, 2-5 September 2012)*, 369-376. Milano.

Dunning, Sansone, Teperek 2019

A. Dunning, S.A., Sansone, M. Teperek, The Layered Cake of FAIR Coordination: How Many is too Many?, *Scientific Data blog* 22. [blogs.nature.com/scientificdata/2019/10/22/the-layered-cake/].

Ekengren *et al.* 2021

F. Ekengren, M. Callieri, D. Dinunno, Å. Berggren, S. MacHeridis, N. Dell'Unto, Dynamic Collections: A 3D Web Infrastructure for Artifact Engagement, *Open Archaeology* 7(1), 337-352.

Evans, Daly 2006

T.L. Evans, P. Daly (eds.), *Digital Archaeology. Bridging Method and Theory*. London-New York.

Faniel *et al.* 2021

I. Faniel, A. Austin, S.W. Kansa, E. Kansa, J. Jacobs, P. France, Identifying Opportunities for Collective Curation During Archaeological Excavations, *International Journal of Digital Curation* 16(1), 1-17. [doi.org/10.2218/ijdc.v16i1.742].

Figuera 2021

M. Figuera, A Fuzzy Approach to Evaluate the Attributions Reliability in the Archaeological Sources, *International Journal of Digital Libraries* 22, 289-296. [doi.org/10.1007/s00799-020-00284-6].

Felicetti, Lorenzini 2011

A. Felicetti, M. Lorenzini, Metadata and Tools for Integration and Preservation of Cultural Heritage 3d Information, *Geoinformatics FCE CTU*, 6, 118-124. [<https://doi.org/10.14311/gi.6.16>].

Felicetti *et al.* 2015

A. Felicetti, I. Galluccio, C. Luddi, M.L. Mancinelli, T.y Scarselli, A.D. Madonna, Integrating Terminological Tools and Semantic Archaeological Information: The ICCD RA Schema and Thesaurus. In P. Ronzino (a c.), *Proceedings of the Workshop on Extending, Mapping and Focusing the CRM* co-located with 19<sup>th</sup> International Conference on Theory and Practice of Digital Libraries (Poznan, Poland, September 17, CEUR Workshop Proceedings 1656), 28-43.

Forte 1993

M. Forte, Un esperimento di visualizzazione scientifica per l'archeologia del paesaggio: la navigazione nel paesaggio virtuale, *Archeologia e Calcolatori* 4, 137-152.

Francovich 1999

R. Francovich, Archeologia medievale e informatica: dieci anni dopo, *Archeologia e Calcolatori* 10, 45-61.

Frischer 2008

B. Frischer, Introduction. From Digital Illustration to Digital Heuristics. In B. Frischer, A. Dakouri-Hild (eds.), *Beyond Illustration: 2D and 3D Digital Technologies as Tools for Discovery in Archaeology* (BAR International Series S1805), v-xxiv, Oxford.

Garagnani 2021

S. Garagnani, A. Gaucci, P. Moscati, M. Gaiani, *ArchaeoBIM. Theory, Processes and Digital Methodologies for the Lost Heritage*. Bologna.

Gardin 1959

J.-C. Gardin, Four Codes for the Description of Artefacts: An Essay in Archaeological Technique and Theory, *American Anthropologist* 60, 335-357.

Gardin 1971

J.-C. Gardin, Archaeology and Computers: New Perspectives, *International Social Sciences Journal* 23(2), 189-203.

Gardin 1999

J.-C. Gardin, Calcul et narrativité dans les publications archéologiques, *Archeologia e Calcolatori* 10, 63-78.

Geser *et al.* 2022

G. Geser, J.D. Richards, F. Massara, H. Wright, Data Management Policies and Practices of Digital Archaeological Repositories, *Internet Archaeology* 59. [doi.org/10.11141/ia.59.2].

Gillings, Hacıgüzeller, Lock 2020

M. Gillings, P. Hacıgüzeller, G. Lock, Archaeology and Spatial Analysis. In M. Gillings, P. Hacıgüzeller, G. Lock (eds.), *Archaeological Spatial Analysis: A Methodological Guide*, 1-16. New York.

Goodall, Roberts 2019

M. Goodall, B. Roberts (eds.), *New Media Archaeologies*. Amsterdam. [doi.org/10.1017/9789048532094].

Guarino 1998

N. Guarino, Formal Ontology and Information Systems. In N. Guarino (ed.), *Proceedings of the 1st International Conference on Formal Ontologies in Information Systems*, FOIS'98 (Trento, Italy 6-8 June), 3-15. Amsterdam.

Guarino, Giarretta 1995

N. Guarino, P. Giarretta, Ontologies and Knowledge Bases: Towards a Terminological Clarification. In N. Mars (ed.), *Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing*, 25-32. Amsterdam.

Guermendi 1996

M.P. Guermendi, L'informatica come risorsa decisiva nella gestione del patrimonio archeologico: le attività dell'Istituto Beni Culturali della Regione Emilia-Romagna, *Archeologia e Calcolatori* 7, 837-848.

Gruber 1995

T.R. Gruber, Toward Principles for the Design of Ontologies Used for Knowledge Sharing?, *International Journal of Human-Computer Studies* 43(5-6), 907-928. [doi.org/10.1006/ijhc.1995.1081].

Hacıgüzeller *et al.* 2021

P. Hacıgüzeller, K. van Daele, F. Carpentier, R. Ribbens, Digital Archiving of Archaeological Resources in Flanders (Belgium): A Brief Review, *Internet Archaeology* 58. [doi.org/10.11141/ia.58.7].

Harris 1983

E.C. Harris, *Principi di stratigrafia archeologica*. Roma.

Hasnain, Rebholz-Schuhmann 2018

A. Hasnain, D. Rebholz-Schuhmann, Assessing FAIR Data Principles against the 5-Stars Open Data Principles. In *The Semantic Web: ESWC 2018 Satellite Events. ESWC 2018* (Lecture Notes in Computer Science, vol. 11155). [doi.org/10.1007/978-3-319-98192-5\_60].

Heersmink 2012

R. Heersmink, Mind and Artifact: A Multidimensional Matrix for Exploring Cognition-Artifact Relations. In J. Bishop, Y. Erden (eds.), *The 5th AISB Symposium on Computing and Philosophy: Computing, Philosophy and the Question of Bio-Machine Hybrids, AISB/IACAP World Congress 2012*, Birmingham (UK): *Society for the Study of Artificial Intelligence and Simulation of Behaviour*, 48-55. Birmingham.

Hermon, Niccolucci 2021

S. Hermon, F. Niccolucci, FAIR Data and Cultural Heritage Special Issue Editorial Note, *International Journal on Digital Libraries* 22, 251-255. [doi.org/10.1007/s00799-021-00309-8].

Hermon *et al.* 2022

S. Hermon, F. Niccolucci, N. Bakirtzis, S. Gasanova, Digital Twins in Cultural Heritage. [dx.doi.org/10.2139/ssrn.4274457].

Hey, Tansley, Tolle 2009

T. Hey, S. Tansley, K. Tolle (eds.), *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Washington.

Hodder 2012

I. Hodder, *Entangled: An Archaeology of the Relationships between Humans and Things*. Oxford.

Hodder, Orton 1976

I. Hodder, C. Orton, *Spatial Analysis in Archaeology* (New Studies in Archaeology 1). Cambridge.

Hollander 2021

H. Hollander, Digital Dutch Archaeology: Future perspectives, *Internet Archaeology* 58. [doi.org/10.11141/ia.58.28].

Huggett 2004

J. Huggett, Archaeology and the New Technological Fetishism, *Archeologia e Calcolatori* 15, 81-92.

Huggett 2012

J. Huggett, What Lies Beneath: Lifting the Lid on Archaeological Computing. In A. Chrysanthi, P. Murrieta Flores, C. Papadopoulos (eds.), *Thinking Beyond the Tool: Archaeological Computing and the Interpretative Process*, 204-214. Oxford.

Huggett 2015

J. Huggett, A Manifesto for an Introspective Digital Archaeology, *Open Archaeology* 1, 86-95.

Huggett 2017

J. Huggett, The Apparatus of Digital Archaeology, *Internet Archaeology* 44. [doi.org/10.11141/ia.44.7].

Huggett 2022

J. Huggett, Data Legacies, Epistemic Anxieties, and Digital Imaginaries in Archaeology, *Digital* 2(2), 267-295. [doi.org/10.3390/digital2020016].

Huvila 2022

I. Huvila, Improving the Usefulness of Research Data with Better Paradata, *Open Information Science* 6(1), 28-48. [doi.org/10.1515/opis-2022-0129].

ICOMOS 1965

ICOMOS 1965, *International Charter for the Conservation and Restoration of Monuments and Sites*. Venezia. [https://www.icomos.org/images/DOCUMENTS/Charters/venice\_e.pdf].

ICOMOS 1990

ICOMOS 1990, *Charter for the Protection and Management of the Archaeological Heritage*. Lausanne. [https://www.icomos.org/images/DOCUMENTS/Charters/arch\_e.pdf].

ICOMOS 1996

ICOMOS 1996, *Charter on the Protection and Management of Underwater Cultural Heritage*. Sofia. [[https://www.icomos.org/images/DOCUMENTS/Charters/underwater\\_e.pdf](https://www.icomos.org/images/DOCUMENTS/Charters/underwater_e.pdf)].

Kansa, Kansa 2022

E. Kansa, S.W. Kansa, Promoting Data Quality and Reuse in Archaeology through Collaborative Identifier Practices, *Proceedings National Academy Science USA* 119. [doi: 10.1073/pnas.2109313118].

Kansa, Kansa, Watrall 2011

E. Kansa, S.W. Kansa, E. Watrall (eds.), *Archaeology 2.0: New Approaches to Communication and Collaboration* (Cotsen Digital Archaeology Series 1). Los Angeles.

Kansa et al. 2020

S.W. Kansa, L. Atici, E. Kansa, R. Meadow, Archaeological Analysis in the Information Age: Guidelines for Maximizing the Reach, Comprehensiveness, and Longevity of Data, *Advances in Archaeological Practice* 8(1), 40-52. [doi:10.1017/aap.2019.36].

Kintigh et al. 2014

K. Kintigh, J. Altschul, M. Beaudry, R. Drennan, A. Kinzig, T. Kohler, M. Zeder, Grand Challenges for Archaeology, *American Antiquity* 79(1), 5-24.

Kuhn 1962

T. Kuhn, *The Structure of Scientific Revolutions*. Chicago.

Jakobsson et al. 2021

U. Jakobsson, D. Novák, J.D. Richards, B. Štular, H. Wright, Digital Archiving in Archaeology: The State of the Art, *Internet Archaeology* 58. [doi.org/10.11141/ia.58.23].

Lake 2012

M. Lake, Open Archaeology, *World Archaeology* 44(4), 471-478. [doi.org/10.1080/00438243.2012.748521].

Limoncelli 2012

M. Limoncelli, *Il restauro virtuale in Archeologia*. Roma.

Lock 2003

G. Lock, *Using Computers in Archaeology: Towards Virtual Pasts*. London.



Lock 2009

G. Lock, Archaeological Computing then and now: Theory and Practice, Intentions and Tensions, *Archeologia e Calcolatori* 20, 75-84.

Lombardo *et al.* 2023

V. Lombardo, T. Karatas, M. Gulmini, L. Guidorzi, L.D. Angelici, Transdisciplinary Approach to Archaeological Investigations in a Semantic Web Perspective, *Semantic Web* 14(2), 361-383. [doi.org/10.3233/SW-223016].

Löwenborg 2018

D. Löwenborg, Knowledge Production with Data from Archaeological Excavations. In I. Huvila (ed.), *Archaeology and Archaeological Information in the Digital Society*, 37-53. London.

Löwenborg *et al.* 2021

D. Löwenborg, M. Jonsson, A. Larsson, J. Nordinge, A Turn Towards the Digital. An Overview of Swedish Heritage Information Management Today, *Internet Archaeology* 58. [doi.org/10.11141/ia.58.19].

Mannoni 2000

T. Mannoni, Alcune considerazioni metodologiche. In G.P. Brogiolo, G. Olcese (a c.), *Produzione ceramica in area padana tra il II secolo a.C. e il VIII secolo d.C.: nuovi dati e prospettive di ricerca* (Desenzano del Garda, 8-10 aprile 1999, Mantova), 217-219. Firenze.

Mannoni 2002

T. Mannoni, Dall'Archeologia globale del territorio alla Carte del patrimonio archeologico, architettonico e storico-ambientale. In *Archeologia dei Centri Storici: analisi, conoscenze e conservazione* (Atti del seminario di Archeologia dell'Urbanistica, 14/21 novembre 1998, Trento), 21-27. Trento.

Markos, Kalacy, Sarris 2022

K. Markos, T. Kalacy, A. Sarris, Bridging Digital Approaches and Legacy in Archaeology, *Digital* 2, 538-545.

Marlet, Rodier 2019

O. Marlet, X. Rodier, A Way to Express the Reliability of Archaeological Data: Data Traceability at the *Laboratoire Archéologie et Territoires* (Tours, France), *International Journal on Digital Libraries* 22, 251-255. [doi.org/10.1007/s00799-019-00274-3].

Marwick 2017

B. Marwick, Computational Reproducibility in Archaeological Research: Basic Principles and a Case Study of their Implementation, *Journal of Archaeological Method and Theory* 24(2), 424-450.

May 2017

K. May, Digital Archaeological Heritage: an introduction, *Internet Archaeology* 43. [doi.org/10.11141/ia.43.1].

McHenry, Bajcsy 2008

K. McHenry, P. Bajcsy, *An Overview of 3d Data Content, File Formats and Viewers* (Research Reports: National Center for Supercomputing Applications, NCSA, University of Illinois at Urbana-Champaign, Urbana IL). [www.archives.gov/files/applied-research/ncsa/8-an-overview-of-3d-data-content-file-formats-and-viewers.pdf].

McKeague 2021

P. McKeague, A Snapshot in Time: A Review of Current Approaches to Archaeological Archiving in Scotland, *Internet Archaeology* 58. [doi.org/10.11141/ia.58.14].

Medri 2008

M. Medri, *Manuale di rilievo archeologico*. Roma.

Myers, Dalgity, Avramides 2016

D. Myers, A. Dalgity, I. Avramides, The Arches Heritage Inventory and Management System: A Platform for the Heritage Field, *Journal of Cultural Heritage Management and Sustainable Development* 6(2), 213-224.

Moberg 1981

C.A. Moberg, *Introduzione all'archeologia*. Milano.

Mons et al. 2017

B. Mons, N. Cameron, J. Velterop, M. Dumontier, L.O.B. Da Silva Santos, M.D. Wilkinson, Cloudy, Increasingly FAIR. Revisiting the FAIR Data Guiding Principles for the European Open Science Cloud, *Information Services & Use* 37(1), 49-56. [doi.org/10.3233/ISU-170824].

Moore, Rountrey, Kettler 2022

J. Moore, A. Rountrey, H.S. Kettler (eds.), *3D Data Creation and Curation: Community Standards for 3D Data Preservation*. Chicago, USA.

Morgan, Eve 2012

C. Morgan, S. Eve, DIY and Digital Archaeology: Hhat are you Doing to Participate?, *World Archaeology* 44(4), 521-537.

Moscati 2013

P. Moscati, Jean-Claude Gardin (Parigi 1925-2013). Dalla meccanografia all'informatica archeologica, *Archeologia e Calcolatori* 24, 7-24.

Moscati 2017

P. Moscati, Open Data, Open Knowledge, Open Science: quali prospettive, *Archeologia e Calcolatori* supplemento 9, 137-140.

Moscati 2019

P. Moscati, Informatica archeologica e Archeologia digitale. Le risposte dalla rete, *Archeologia e Calcolatori* 30, 21-38.

Moser 2011

S. Moser, Archaeological Visualization. Early Artifact Illustration and the Birth of the Archaeological Image. In I. Hodder (ed.), *Archaeological Theory Today*, 292-322. Chichester.

Münster 2018

S. Münster, Digital 3D Modelling in the Humanities. In 13<sup>th</sup> *Annual International Conference of the Alliance of Digital Humanities, Digital Heritage 2018*, 627-629. Mexico City.

Münster 2022

S. Münster, Digital 3D Technologies for Humanities Research and Education: An Overview, *Applied Science* 12, 2426. [doi.org/10.3390/app12052426].

Niccolucci 2017

F. Niccolucci, Ariadne e gli Open Data: come trasformare i dati archeologici da Open a "FAIR", *Archeologia e calcolatori* supplemento 9, 141-150.

Niccolucci 2020

F. Niccolucci, From Digital Archaeology to Data-Centric Archaeological Research, *Magazén* 1(1), 35-53. [doi.org/10.30687/mag/2020/01/002].

Niccolucci, Richards 2013

F. Niccolucci, J.D. Richards, ARIADNE: Advanced Research Infrastructures for Archaeological Dataset Networking in Europe. A New Project to Foster

and Support Archaeological Data Sharing, *The European Archaeologist* 39. [e-a-a.org/TEA/archive/TEA\_39\_SUMMER\_2013/rep1\_39.pdf].

Niccolucci, Felicetti, Hermon 2022

F. Niccolucci, A. Felicetti, S. Hermon, Populating the Data Space for Cultural Heritage with Heritage Digital Twins, *Data* 7, 105. [doi.org/10.3390/data7080105].

Niccolucci, Hermon, Doerr 2015

F. Niccolucci, S. Hermon, M. Doerr, The Formal Logical Foundations of Archaeological Ontologies. In J. Barcelo, I. Bogdanovic (eds.), *Mathematics and Archaeology*, 86-99. Boca Raton.

Niccolucci *et al.* 2018

F. Niccolucci, C. Meghini, A. Felicetti, L. Pezzati, Condivisione dei dati sui beni culturali: DIGILAB, l'esperienza di ARIADNE e di E-RIHS. In *Atti della Conferenza GARR 2017* (Venezia 15-17 novembre 2017), 129-132. Roma. [doi.org/10.26314/GARR-Confi7-proceedings-24].

Nobles, Çakirlar, Svetachov 2019

G. Nobles, C. Çakirlar, P. Svetachov, Bonify 1.0: Evaluating Virtual Reference Collections in Teaching and Research, *Archaeological and Anthropological Sciences* 11(10), 5705-5716. [doi.org/10.1007/s12520-019-00898-1].

Olsen *et al.* 2012

B. Olsen, M. Shanks, T. Webmoor, C. Witmore, *Archaeology. The Discipline of Things*. California.

Opgenhaffen 2021

L. Opgenhaffen, Visualizing Archeologists: A Reflexive History of Visualization Practice in Archaeology, *Open Archaeology* 7, 353-377.

Opgenhaffen 2022

L. Opgenhaffen, Archives in Action. The Impact of Digital Technology on Archaeological Recording Strategies and Ensuing Open Research Archive, *Digital Applications in Archaeology and Cultural Heritage* 27. [doi.org/10.1016/j.daach.2022.e00231].

Orlandi 1997

T. Orlandi, Informatica, formalizzazione e discipline umanistiche. In T. Orlandi (a c.) *Il problema della formalizzazione* (Ciclo di seminari, febbraio-giugno 1994, Accademia Nazionale dei Lincei). Roma.

Orlandi 2004

T. Orlandi, Archeologia teorica e informatica archeologica. Un rapporto difficile, *Archeologia e Calcolatori* 15, 41-50.

Piggott 1965

S. Piggott, Archaeological Draughtsmanship: Principles and Practice Part I: Principles and Retrospect, *Antiquity* 39, 165-176.

Al-Rawashdeh Balqies Sadoun, Al Fukara 2012

S. Al-Rawashdeh Balqies Sadoun, A. Al Fukara, CAD File Conversion to GIS Layers: Issues and Solutions. In *International Conference on Computer, Information and Telecommunication Systems (CITS)*, 1-6.

Raccomandazione 1970

Commissione Europea, Raccomandazione 1970. [eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32021H1970&from=EN].

Remondino, Campana 2014

F. Remondino, S. Campana (eds.), *3d Recording and Modelling in Archaeology and Cultural Heritage: Theory and Best Practices* (BAR, International Series 2598). Oxford.

Renfrew, Bahn 2006

C. Renfrew, P. Bahn, *Archeologia. Teoria, metodi, pratica*. Bologna.

Richards 2012

J.D. Richards, Digital Infrastructures for Archaeological Research: A European Perspective, *The CSA Newsletter* 25(2). [<https://csanet.org/newsletter/fall12/nlf1202.html>].

Richards 2017

J.D. Richards, Twenty Years Preserving Data: A View from the United Kingdom, *Advances in Archaeological Practice* 5(3), 227-237.

Richards 2021

J.D. Richards, Archiving Archaeological Data in the United Kingdom, *Internet Archaeology* 58. [doi.org/10.11141/ia.58.21].

Richards, Niccolucci 2019

J.D. Richards, F. Niccolucci (eds.) *The Ariadne Impact*. Budapest.

Richards *et al.* 2021

J.D. Richards, U. Jakobsson, D. Novák, B. Štular, H. Wright, Digital Archiving in Archaeology: The State of the Art. Introduction, *Internet Archaeology* 58. [doi.org/10.11141/ia.58.23].

Roosevelt *et al.* 2015

C.H. Roosevelt, P. Cobb, E. Moss, B.R. Olson, U. Sinan, Excavation is Destruction Digitization: Advances in Archaeological Practice, *Journal of Field Archaeology* 40(3), 325-346.

Ross *et al.* 2022

S. Ross, B. Ballsun-Stanton, S. Cassidy, P. Crook, J. Klump, A. Sobotkova, FAIRer Data Through Digital Recording: The FAIMS Mobile Experience, *Journal of Computer Applications in Archaeology* 5(1), 271-285. [doi.org/10.5334/jcaa.96].

SAA 2018

SAA Task Force. *Guidelines for Promotion and Tenure for Archaeologists in Diverse Academic Roles*. [https://documents.saa.org/container/docs/default-source/doc-careerpractice/promotion-and-tenure-white-paper-02-12-19-web.pdf?sfvrsn=8a5815cd\_4].

Sahoo *et al.* 2009

S.S. Sahoo, W. Halb, S. Hellmann, K. Idehen, J. Thibodeau, S. Auer, J. Sequeda, A. Ezzat, *A Survey of Current Approaches for Mapping Relational Databases to RDF*. In *W3C Incubator Group*. 2009 [www.w3.org/2005/Incubator/rdb2rdf/RDB2RDF\_SurveyReport.pdf].

Schnapp 1996

A. Schnapp, *The Discovery of the Past: The Origins of Archaeology*. London.

Sketchfab 2019

Sketchfab, *Cultural Heritage Survey*. [https://docs.google.com/presentation/d/bc1qk55vk7wjzg3pmxlh59rv5dlgewd9jem5nrt4wlhw79CIHT5A/edit#slide=id.g5dae1f95c1\_o\_0].

Sobotkova 2018

A. Sobotkova, Sociotechnical Obstacles to Archaeological Data Reuse, *Advances in Archaeological Practice* 6(2), 117-24. [doi.org/10.1017/aap.2017.37].

Tan *et al.* 2022

J. Tan, J. Leng, X. Zeng, D. Feng, P. Yu, Digital Twin for Xiegong's Architectural Archaeological Research: A Case Study of Xuanluo Hall, Sichuan, China, *Buildings* 12, 1053. [doi.org/10.3390/buildings12071053].

Tapscot, Williams 2007

D. Tapscot, A.D. Williams, *La collaborazione di massa che sta cambiando il mondo*. Milano.

Tudhope, Binding, May 2008

D. Tudhope, C. Binding, K. May, Semantic Interoperability Issues from a Case Study in Archaeology. Semantic Interoperability in the European Digital Library. In S. Kollias, J. Cousins (eds.), *Proceedings of the First International Workshop SIEDL*, 88-99. Tenerife.

Ullah 2015

I. Ullah, Integrating Older Survey Data into Modern Research Paradigms: Identifying and Correcting Spatial Error in "Legacy" Datasets, *Advances in Archaeological Practice* 3(4), 331-350. [doi:10.7183/2326-3768.3.4.331].

UNESCO 1956

UNESCO 1956, *Recommendation on International Principles Applicable to Archaeological Excavations*, New Delhi. [https://www.unesco.org/en/legal-affairs/recommendation-international-principles-applicable-archaeological-excavations].

UNESCO 2001

UNESCO 2001, *Convention on the Protection of the Underwater Cultural Heritage*, Paris. [https://unesdoc.unesco.org/ark:/48223/pf0000126065].

UNESCO 2003

UNESCO 2003, *Charter on the Preservation of the Digital Heritage*, Paris. [https://unesdoc.unesco.org/ark:/48223/pf0000179529?posInSet=1&queryId=9940778f-c1b4-4260-8e70-1c90d4da3cba].

van Helden, Hong, Allison 2018

D. van Helden, Y. Hong, P. Allison, Building an Ontology of Tablewares using "Legacy Data", *Internet Archaeology* 50. [doi.org/10.11141/ia.50.13].

van Ruymbeke *et al.* 2018

M. van Ruymbeke, P. Hallot, G.A. Nys, R. Billen, Implementation of Multiple Interpretation Data Model Concepts in CIDOC CRM and Compatible

Models, *Virtual Archaeology Review* 9(19), 50-65. [doi.org/10.4995/var.2018.8884].

W3C 2000

W3C Recommendation 6 October 2020. Extensible Markup Language (XML) 1.0 (Second Edition). [http://www.w3c.org/TR/2000/REC-xml-20001006].

Whallon 1972

R. Whallon, The Computer in Archaeology: A Critical Survey, *Computers and the Humanities* 7(1), 29-45.

Wheeler 1954

M. Wheeler, *Archaeology from the Earth*. Baltimore, Maryland.

Wilkinson *et al.* 2016

M.D. Wilkinson, M. Dumontier, I. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.W. Boiten, L. B. da Silva Santos, P.E. Bourne, J. Bouwman, A.J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C.T. Evelo, r. Finkers, A. Gonzalez-Beltran, J.G.A. Gray, P. Groth, C. Goble, J.S. Grethe, J. Heringa, P.A.C. Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S.J. Lusher, M.E. Martone, A. Mons, A.L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M.A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, B. Mons, The FAIR Guiding Principles for Scientific Data Management and Stewardship, *Scientific Data* 3. [doi.org/10.1038/sdata.2016.18].

Wilkinson *et al.* 2019

M.D. Wilkinson, M. Dumontier, I. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.W. Boiten, L. B. da Silva Santos, P.E. Bourne, J. Bouwman, A.J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C.T. Evelo, r. Finkers, A. Gonzalez-Beltran, J.G.A. Gray, P. Groth, C. Goble, J.S. Grethe, J. Heringa, P.A.C. Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S.J. Lusher, M.E. Martone, A. Mons, A.L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M.A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, B. Mons, Addendum: FAIR Guiding Principles for Scientific Data Management and Stewardship, *Scientific Data* 6. [doi.org/10.1038/s41597-019-0009-6].



## Witcher 2008

R.E. Witcher, (Re)surveying Mediterranean Rural Landscapes: GIS and Legacy Survey Data, *Internet Archaeology* 24. [doi.org/10.11141/ia.24.2].

## Wright, Richards 2018

H. Wright, J.D. Richards, Reflections on Collaborative: Archaeology and Large-Scale Online Research Infrastructures, *Journal of Field Archaeology* 43, suppl.1, 60-67. [doi.org/10.1080/00934690.2018.1511960].

## Zubrow 2006

E.B.W. Zubrow, Digital Archaeology: A Historical Context. In T.L. Evans, P. Daly (eds.), *Digital Archaeology. Bridging Method and Theory*, 8-26. London, New York.

## Zanini 2021

E. Zanini, Riflettendo ancora sul matrix di Harris: una vita (professionale) dopo. In Tiziano Mannoni. *Attualità e sviluppi di metodi e idee*, vol. 2, 665-671.

## SITOGRAFIA

(ultimo accesso 28 aprile 2023)

3DCOFORM

[www.3d-coform.eu](http://www.3d-coform.eu).

3DHOP

[3dhop.net](http://3dhop.net).

3DICONs

[www.3dicons-project.eu](http://www.3dicons-project.eu).

3DICONs, Portal

[3dicons.ceti.gr](http://3dicons.ceti.gr).

ACDH-CH

[vocabs.dariah.eu/en/about](http://vocabs.dariah.eu/en/about).

ADS, FAIR DATA

[archaeologydataservice.ac.uk/help-guidance/data-reuse/fair-data](http://archaeologydataservice.ac.uk/help-guidance/data-reuse/fair-data).

ADS, metadata

[archaeologydataservice.ac.uk/help-guidance/guides-to-good-practice/the-project-lifecycle/project-metadata](http://archaeologydataservice.ac.uk/help-guidance/guides-to-good-practice/the-project-lifecycle/project-metadata).

ARACHNE

[arachne.dainst.org](http://arachne.dainst.org).

ARCHAIDE

[www.archaide.eu](http://www.archaide.eu).

Archeologia e Calcolatori

[www.ispc.cnr.it/it\\_it/2020/07/08/la-rivista-archeologia-e-calcolatori-su-europeana/](http://www.ispc.cnr.it/it_it/2020/07/08/la-rivista-archeologia-e-calcolatori-su-europeana/).

ARCHES

[www.archesproject.org](http://www.archesproject.org).

ArCO

[wit.istc.cnr.it/arco](http://wit.istc.cnr.it/arco).

ARKYES

[www.arkyves.org](http://www.arkyves.org).

ARIADNE

[ariadne-infrastructure.eu](http://ariadne-infrastructure.eu).

[visual.ariadne-infrastructure.eu](http://visual.ariadne-infrastructure.eu).

BARTOC

[bartoc.org](http://bartoc.org).

BBT

[www.backbonethesaurus.eu](http://www.backbonethesaurus.eu).

BONIFY

[www.digitalbones.eu/home](http://www.digitalbones.eu/home).

CARARE

[pro.carare.eu/en/introduction-carare-aggregation-services/carare-metadata-schema](http://pro.carare.eu/en/introduction-carare-aggregation-services/carare-metadata-schema).

[mint.image.ece.ntua.gr/redmine/projects/mint/wiki/Introduction\\_to\\_MINT](http://mint.image.ece.ntua.gr/redmine/projects/mint/wiki/Introduction_to_MINT).

Catalogo Ministero della Cultura

[dati.beniculturali.it/descrizione\\_dataset](http://dati.beniculturali.it/descrizione_dataset).

Catalogo Nazionale dei Dati Aperti

[www.dati.gov.it](http://www.dati.gov.it).

CS<sub>3</sub>DP

[cs3dp.org](http://cs3dp.org).

CIDOC-CRM

[www.cidoc-crm.org](http://www.cidoc-crm.org).

[www.cidoc-crm.org/collaborations](http://www.cidoc-crm.org/collaborations);

[www.progettomagia.unifi.it/vp-98-cidoc-crmtex.html](http://www.progettomagia.unifi.it/vp-98-cidoc-crmtex.html).

CIDOC-CRMdig

[www.cidoc-crm.org/crmDIG](http://www.cidoc-crm.org/crmDIG).

CLARIAH

[clariah.at](http://clariah.at).

CLARIN

[www.clarin.eu](http://www.clarin.eu).

Collections Trust, Historic England  
[www.heritage-standards.org.uk/wp-content/uploads/2020/02/ObjType\\_class.pdf](http://www.heritage-standards.org.uk/wp-content/uploads/2020/02/ObjType_class.pdf).  
[www.heritagedata.org/blog/vocabularies-provided](http://www.heritagedata.org/blog/vocabularies-provided).

CORONA  
[earthexplorer.usgs.gov](http://earthexplorer.usgs.gov).

CulturalOn  
[dati.beniculturali.it/cis](http://dati.beniculturali.it/cis).

DARIAH  
[www.dariah.eu](http://www.dariah.eu).

Dati Europei, portale ufficiale  
[data.europa.eu/en](http://data.europa.eu/en).

DBPedia  
[www.dbpedia.org](http://www.dbpedia.org).

Digital Topographical Bibliography  
[topbib.griffith.ox.ac.uk//index.html](http://topbib.griffith.ox.ac.uk//index.html).

Dublin Core  
[www.dublincore.org](http://www.dublincore.org).

EAGLE  
[www.eagle-network.eu](http://www.eagle-network.eu).  
[www.eagle-network.eu/virtual-exhibition](http://www.eagle-network.eu/virtual-exhibition).  
[webgl-eagle.d4science.org](http://webgl-eagle.d4science.org).  
[www.eagle-network.eu/resources/flagship-mobile-app/](http://www.eagle-network.eu/resources/flagship-mobile-app/).

EDM  
[pro.europeana.eu/page/edm-documentation](http://pro.europeana.eu/page/edm-documentation).

Ename Charter  
[www.enamecenter.org](http://www.enamecenter.org).

EPSG  
[epsg.io](http://epsg.io).

e-RIHS  
[www.e-rihs.eu](http://www.e-rihs.eu).

ESFRI

[www.esfri.eu/roadmap-2016](http://www.esfri.eu/roadmap-2016).

EUROPEANA

[www.europeana.eu/portal](http://www.europeana.eu/portal).

[www.europeana.eu/en/themes/archaeology](http://www.europeana.eu/en/themes/archaeology).

[pro.europeana.eu/project/europeana-archaeology](http://pro.europeana.eu/project/europeana-archaeology); [europeanaarchaeology.carare.eu](http://europeanaarchaeology.carare.eu).

[www.europeana.eu/en/exhibitions/uncovering-hidden-stories](http://www.europeana.eu/en/exhibitions/uncovering-hidden-stories).

[app-share3d.imsi.athenarc.gr/mappings/home](http://app-share3d.imsi.athenarc.gr/mappings/home).

FAIMS

[faims.edu.au/about](http://faims.edu.au/about).

FAIRsFAIR

[www.fairsfair.eu](http://www.fairsfair.eu).

Fasti online

[www.fastionline.org](http://www.fastionline.org).

Geoportale Nazionale Archeologia

[gna.d4science.org/home](http://gna.d4science.org/home).

Geonames

[www.geonames.org/](http://www.geonames.org/)

Getty Museum Thesauri

[www.getty.edu/research/tools/vocabularies/index.html](http://www.getty.edu/research/tools/vocabularies/index.html).

GLAM3D

[glam3d.org](http://glam3d.org).

GO-FAIR

[www.go-fair.org](http://www.go-fair.org).

GN

[www.pcn.minambiente.it](http://www.pcn.minambiente.it).

GRAVITATE

[cordis.europa.eu/project/id/665155](http://cordis.europa.eu/project/id/665155).

Historic Environment Scotland

[canmore.org.uk](http://canmore.org.uk).

IANUS

<http://datenportal.ianus-fdz.de>.

ICCD –FAIR Data

[www.iccd.beniculturali.it/getFile.php?id=7045](http://www.iccd.beniculturali.it/getFile.php?id=7045).

[vast-lab.org/thesaurus/ra/vocab](http://vast-lab.org/thesaurus/ra/vocab).

[github.com/ICCD-MiBACT/Standard-catalografici/tree/master/strumenti-terminologici](https://github.com/ICCD-MiBACT/Standard-catalografici/tree/master/strumenti-terminologici).

ICONCLASS

[iconclass.org](http://iconclass.org); [iconclass.org/help/aboutc#webpisys](http://iconclass.org/help/aboutc#webpisys)

[iconclass.org/testset/ICONCLASS\\_and\\_AI.pdf](http://iconclass.org/testset/ICONCLASS_and_AI.pdf).

iDAI

<https://idai.world>.

IdRef

[www.idref.fr](http://www.idref.fr).

IIIF 3D

[iiif.io](http://iiif.io).

INSPIRE

[inspire.ec.europa.eu](http://inspire.ec.europa.eu).

IOGP

[www.iogp.org](http://www.iogp.org).

ISO TC211

[www.iso.org/committee/54904.html](http://www.iso.org/committee/54904.html); [committee.iso.org/home/tc211](http://committee.iso.org/home/tc211).

ISO 19101

[www.iso.org/standard/59164.html](http://www.iso.org/standard/59164.html).

ISO 19111

[www.iso.org/standard/74039.html](http://www.iso.org/standard/74039.html).

ISO 19115

[www.iso.org/standard/53798.html](http://www.iso.org/standard/53798.html).

ISO 25964

[www.iso.org/standard/53657.html](http://www.iso.org/standard/53657.html).

Levantine Ceramics Project

[www.levantineceramics.org](http://www.levantineceramics.org).

LIMC

[weblimc.org](http://weblimc.org).

LIMC-EAGLE

[www.eagle-network.eu/wp-content/uploads/2014/01/EAGLE-2014-20140929-SZABADOS-notxt-PDF.pdf](http://www.eagle-network.eu/wp-content/uploads/2014/01/EAGLE-2014-20140929-SZABADOS-notxt-PDF.pdf)

Linked Heritage

[www.linkedheritage.eu/](http://www.linkedheritage.eu/).

London Charter

[www.londoncharter.org](http://www.londoncharter.org).

MAPPA – MOD

[mappaproject.arch.unipi.it/mod/Index.php](http://mappaproject.arch.unipi.it/mod/Index.php).

MET

<https://bartoc.org/en/node/978>.

NFDI4Culture

[nfdi4culture.de](http://nfdi4culture.de).

NGA/NIMA

[geonames.nga.mil/geonames/GeographicNamesSearch](http://geonames.nga.mil/geonames/GeographicNamesSearch).

NEXUS

[vcg.isti.cnr.it/nexus](http://vcg.isti.cnr.it/nexus).

OAI-PMH

[www.openarchives.org/pmh](http://www.openarchives.org/pmh).

OGC

[www.ogc.org/standards](http://www.ogc.org/standards).

Open Definition

[opendefinition.org](http://opendefinition.org).

Open Science Framework

[osf.io](http://osf.io).

OntopiA

[github.com/italia/daf-ontologie-vocabolari-controllati/wiki](https://github.com/italia/daf-ontologie-vocabolari-controllati/wiki).

Ontology Design Patterns

[ontologydesignpatterns.org/wiki/Main\\_Page](http://ontologydesignpatterns.org/wiki/Main_Page).

OSGeo

[www.osgeo.org](http://www.osgeo.org).

OWL-Time

[www.w3.org/TR/owl-time](http://www.w3.org/TR/owl-time).

Pactol

[pactols.frantiq.fr/opentheso](http://pactols.frantiq.fr/opentheso).

PARTHENOS

[www.parthenos-project.eu](http://www.parthenos-project.eu).

PAS

<https://finds.org.uk/database>.

PeriodO

[perio.do/en/.client.perio.do/?page=open-backend](http://perio.do/en/.client.perio.do/?page=open-backend).

Pleiades

[pleiades.stoa.org/vocabularies/place-types](http://pleiades.stoa.org/vocabularies/place-types).

ORCID

[orcid.org](http://orcid.org).

ResearchSpace

[researchspace.org](http://researchspace.org).

RESILIENCE

[www.resilience-ri.eu](http://www.resilience-ri.eu).

Sevilla Principles

[www.sevilleprinciples.com](http://www.sevilleprinciples.com).

Share3D

[www.share3d.com](http://www.share3d.com).

SIGECWEB

[www.sigecweb.beniculturali.it/it.iccd.sigec.axweb.Main/](http://www.sigecweb.beniculturali.it/it.iccd.sigec.axweb.Main/)

SITAR

[www.archeositarproject.it/wp-content/uploads/2020/10/Mapping-SITAR-towards-CIDOC-CRM-final\\_12.10.2016-1-1.pdf](http://www.archeositarproject.it/wp-content/uploads/2020/10/Mapping-SITAR-towards-CIDOC-CRM-final_12.10.2016-1-1.pdf)

Sketchfab

[sketchfab.com](http://sketchfab.com).



Smithsonian 3D Digitization  
[3d.si.edu](http://3d.si.edu).

SND  
[snd.gu.se/en](http://snd.gu.se/en).

Sustainable Archaeology  
[sustarc.mcmaster.ca](http://sustarc.mcmaster.ca).

tDAR  
[www.tdar.org](http://www.tdar.org).  
[www.tdar.org/using-tdar/creating-and-editing-resources/datasets-create-and-edit](http://www.tdar.org/using-tdar/creating-and-editing-resources/datasets-create-and-edit).

TGN  
[www.getty.edu/research/tools/vocabularies/tgn](http://www.getty.edu/research/tools/vocabularies/tgn).

THOT  
[thot.philo.ulg.ac.be](http://thot.philo.ulg.ac.be).

TOPOI  
[repository.edition-topoi.org](http://repository.edition-topoi.org).

UNESCO thesaurus  
[vocabularies.unesco.org](http://vocabularies.unesco.org).

USGS-GNIS  
[geonames.usgs.gov/index.html](http://geonames.usgs.gov/index.html).

VIR  
[www.sigecweb.beniculturali.it/it.iccd.sigec.axweb.Main](http://www.sigecweb.beniculturali.it/it.iccd.sigec.axweb.Main).

W3C Provenance Incubator Group  
[www.w3.org/2005/Incubator/prov/wiki/What\\_Is\\_Provenance](http://www.w3.org/2005/Incubator/prov/wiki/What_Is_Provenance).

Wikidata  
[www.wikidata.org/wiki/Wikidata:Main\\_Page](http://www.wikidata.org/wiki/Wikidata:Main_Page).

WissKi  
[wiss-ki.eu](http://wiss-ki.eu).





IL TORCOLIERE • Officine Grafico-Editoriali d'Ateneo

Università di Napoli L'Orientale  
stampato nel mese di maggio 2023

Lo studio della relazione tra i dati digitali e la metodologia archeologica costituisce uno dei principali campi di interesse dell'Archeologia Digitale, una disciplina che si è affermata recentemente, in parallelo alla trasformazione digitale della società. L'introduzione di strumenti e metodi digitali nel campo della ricerca storica e archeologica ha cambiato l'approccio degli studiosi alla ricostruzione del mondo antico. Il volume si propone di indagare la composizione della nuova cassetta degli attrezzi dell'archeologo che ha affiancato alla tradizionale cazzuola sofisticate attrezzature che producono dati digitali di alta qualità. Il libro delinea gli scenari inaugurati dai principi dell'Open Science, introducendo il lettore alle più recenti sfide che coinvolgono il settore delle scienze archeologiche.

ANDREA D'ANDREA è archeologo. Dal 1993 ha preso parte a numerose esplorazioni archeologiche in Italia e all'estero, occupandosi soprattutto dei rilievi topografici e digitali. Dal 2010 al 2020 ha diretto la missione archeologica ad Abou Ghurab, in Egitto, finanziata dal Ministero degli Affari Esteri e dal 2020 è codirettore delle attività condotte nel sito UNESCO di al-Balid (Sultanato dell'Oman). Dal 2014 al 2021 è stato Direttore Tecnico del Centro Interdipartimentale di Servizi di Archeologia dell'Università di Napoli L'Orientale. Insegna *Archeologia Digitale*, presso il corso di Laurea Magistrale in Archeologia dell'Università di Napoli L'Orientale, e *Cartografia e Topografia* presso la Scuola Interateneo di Specializzazione in Beni Archeologici, Università di Salerno e Università di Napoli L'Orientale.

È Direttore della rivista internazionale online *Archeologie tra Oriente e Occidente* (ISSN 2974-7651) edita da Unior Press.