

Editoriale

Cose di scienza, Big Data e l'unificazione delle collezioni europee di storia naturale

Things of science, Big Data and unification of the European natural history collections

Vincenzo Vomero

"Cose di scienza", oggetti del patrimonio scientifico materiale, raccolte e collezioni che formano il core business attorno al quale gravita tutto ciò che un museo scientifico offre ai suoi pubblici. Quanta poca attenzione alle cose di scienza dedica l'intelligenza del nostro Paese e quanti significati fondamentali invece queste cose racchiudono in sé. Il contenuto informativo nascosto in un fossile, in una zanzara, in una pietra, in uno strumento scientifico è talmente grande che a raccontarlo non basterebbe un intero museo o una enciclopedia. Noi, gli specialisti delle cose di scienza, siamo appena in grado di metterne in evidenza alcuni, pochi in verità, e sempre collegati alle nostre specializzazioni. Ancora non siamo riusciti, però, a far capire a chi sta fuori del museo che conoscere nei dettagli e comprendere, almeno in parte, l'informazione contenuta in questi oggetti ammassati nei nostri musei è invece cruciale per far evolvere la nostra civiltà, il nostro benessere e la nostra stessa sopravvivenza.

Qualità ambientale, servizi ecosistemici, sostenibilità: è ormai riconosciuto che la biodiversità sia uno degli elementi cardine della valutazione ambientale, e che si misuri in grandissima parte con le specie. Sembra banale ma il messaggio stenta a passare ai decisori politici ed economici del nostro governo. Il dato inconfutabile è che possediamo un immenso patrimonio di cose di scienza. Solo il Progetto CollMap, lanciato e condotto dall'ANMS qualche anno fa, primo e unico in Italia, ha censito un patrimonio di 20 milioni di cose di scienza distribuite in 1705 collezioni conservate in 88 musei scientifici, e lo studio non è finito. Non può e non deve avvenire che per scarsità di risorse umane e finanziarie, quando non si tratti invece di vera e propria sciattezza culturale, singole amministrazioni miopi o ignoranti decretino inconsapevolmente il declino e il deterioramento di questo immenso patrimonio di informazioni scientifiche. Dibattiamo con passione lo stato precario in cui versano le collezioni e la ricerca nei nostri musei, ma non basta solo parlarne. Vanno affrontate le cause che hanno portato all'attuale difficoltà operativa che spesso caratterizza la gestione delle nostre istituzioni. Tutti noi abbiamo detto, scritto e agito per far prendere coscienza del problema, ma non abbiamo mai cercato di aggredirlo alla base. È tempo allora di identificare una road map fattibile e concreta, da condividere con quella parte del governo che ci ha promesso aiuto ed attenzione, che porti a una riconsiderazione complessiva del sistema musei scientifici italiani, della loro governance, gestione, organizzazione della ricerca e delle collezioni, e comunicazione mediante una ottimizzazione delle risorse. Quando avremo ridefinito gli indirizzi per la fruizione e la gestione dei musei scientifici allora potranno essere integrati nel Sistema Museale Nazionale garantendo tutela e valorizzazione del patrimonio materiale e immateriale. Nel frattempo però dobbiamo prestare molta attenzione alle novità che stanno crescendo in Europa. Si stima che soltanto nelle collezioni dei musei scientifici europei siano conservati circa un miliardo e mezzo di esemplari e che l'80% della biodiversità finora descritta sia conservata nelle sole collezioni europee. Si capisce facilmente che per studiare queste immense masse di dati non bastino più sistemi e tecniche usate finora. In un sistema globale di ricerca scientifica è necessario ricalibrare lo stesso modo di pensare alla ricerca, anche dal punto di vista epistemologico.

"Big Data" è un termine relativamente nuovo che dovrà interessare anche i grandi musei scientifici. Quando se ne parla ci si riferisce solo a pratiche già in atto in settori strategici come il telerilevamento, l'andamento dei mercati finanziari o le previsioni meteorologiche mentre l'applicazione allo studio della biodiversità è ancora allo stato embrionario. Tutto lascia prevedere che il nuovo approccio ai Big Data sia trasferibile anche all'analisi dei milioni di dati museali che sono la base dati essenziale per rappresentare la distribuzione degli organismi nel tempo e nello spazio, per comprendere la diacronicità del fenomeno evolutivo e dei popolamenti e per prevedere gli scenari più diversi, dalle modifiche generate dagli attuali cambiamenti climatici fino a shift ecologici, filogenetici, fenologici e biogeografici. Nel campo delle scienze naturali solo i musei e le loro collezioni posseggono quantità di materiali e di informazioni in grado di competere con quelli qui citati per altre branche della conoscenza. Nel caso poi dell'approccio ai Big Data la scienza moderna richiede assolutamente l'accesso digitale ai dati. Statistiche condotte negli ultimi anni ci informano che a livello europeo appena il 10% dei dati dei musei di storia naturale sono digitalizzati. Musei olandesi, inglesi e tedeschi hanno già condotto

importanti catalogazioni digitali mentre la maggioranza degli stati ha appena iniziato. Oltre alla digitalizzazione alcuni centri museali d'eccellenza, pochi in verità, hanno anche iniziato la digitalizzazione delle immagini (non più dell'1% degli esemplari). L'operazione di digitalizzazione, come è facilmente intuibile, non è certo semplice ed economica. I grandi numeri, l'oggettiva complessità del processo, la necessità di materiale correttamente identificato, impongono che l'accesso alla digitalizzazione avvenga non solo con un miglioramento della tecnologia usata ma anche con un più adatto approccio culturale e gestionale.

È quello che abbiamo iniziato ad ipotizzare in Italia qualche anno fa ravvisando la necessità di un Istituto diffuso di tassonomia, una struttura delocalizzata e decentrata che si occupasse di razionalizzare la gestione e l'uso in senso tassonomico delle collezioni italiane presenti su tutto il territorio. Oggi in Europa sta prendendo corpo un grande movimento d'opinione che richiede di operare ad altissimo livello proprio sull'uso dell'enorme patrimonio di collezioni scientifiche dei musei. La sfida complessiva è ambiziosa e di difficile realizzazione ma molto virtuosa e vale assolutamente la pena di collaborare: la parola d'ordine è unificare l'accesso ai dati di raccolta di tutte le collezioni del nostro continente in modo coordinato e integrato. Un primo lotto di nazioni europee sta già indagando gli aspetti tecnici, finanziari, politici e di governo di una nuova infrastruttura dedicata alla creazione di un modello di gestione condivisa dell'intero patrimonio europeo di collezioni in grado di supportare aspetti di unificazione dei servizi open access, piani di incentivazione, pianificazione e priorità di ricerca e sviluppo delle necessarie capacità operative. Se e quando l'infrastruttura vedrà la luce, i musei europei potranno concretamente mettere il patrimonio che conservano a disposizione dell'umanità per affrontare e risolvere alcune delle sfide del mondo attuale (perdita di biodiversità, impatto del cambiamento climatico...). Il grande progetto in via di organizzazione noto come DiSSCo (Distributed System of Scientific Collections) è attualmente in fase di valutazione da parte della commissione ESFRI e coinvolge per ora 21 paesi con 150 istituzioni afferenti. L'Italia c'è, e questo per noi è una cosa di cui andar fieri, ed è presente con un consorzio (coordinato dal Museo di Storia Naturale dell'Università di Firenze) che, naturalmente, comprende anche l'Associazione Nazionale Musei Scientifici. Se il progetto riuscisse a entrare nella road map europea delle grandi Research Infrastructures, avrebbe durata fino al 2030 e porterebbe all'unificazione sovranazionale delle informazioni ora contenute in maniera parcellizzata e dispersa. L'Italia contribuirà a DiSSCo, anche con il valore aggiunto del suo grande patrimonio di expertise tassonomica. E noi tutti sappiamo quanti danni possono arrecare le identificazioni errate o non aggiornate dei materiali oggi oggetto di digitalizzazione (con problemi anche nel campo della metagenomica e non solo). Comunque incrociamo le dita perché il vento cambi in Italia, in Europa e nel mondo.

"Things of science", objects of the material scientific heritage, collections that form the core business underlying all that a scientific museum offers to its public. So little attention is paid to the things of science by the intelligencia of Italy, yet so many fundamental meanings are encompassed in these things. The informational content concealed in a fossil, in a mosquito, in a stone, in a scientific instrument is so great that an entire museum or an encyclopaedia would not be sufficient to recount it all. We, the specialists of things of science, are merely able to highlight some (in truth few) of them and always those related to our specialized fields. However, we have not managed to make it clear to those outside the museum that a detailed knowledge and at least partial understanding of the information contained in the objects amassed in our museums is crucial to the evolution of our civilization, our well-being and our survival.

Environmental quality, ecosystem services, sustainability: it is now well known that biodiversity is one of the key elements of environmental assessment, and it is quantified to a large degree by species. This idea seems banal but the message does not seem to reach the political and economic decision-makers of our government. The irrefutable fact is that we possess a vast scientific heritage. The CollMap Project, launched and conducted by the Italian Association of Scientific Museums (ANMS) a few years ago, the first and only one of its kind in Italy, has censused a patrimony of 20 million things of science distributed among 1705 collections preserved in 88 scientific museums, and the study is still not complete. It is unthinkable that, because of a lack of human and financial resources, if not even true cultural slovenliness, myopic or ignorant administrations might unconsciously sanction the decline and deterioration of this immense wealth of scientific information. Let us passionately discuss the precarious state of the collections and research activities in our museums, but talking about it is not enough. We must deal with the causes that have led to the current operational difficulty that

often characterizes the management of our institutions. All of us have spoken, written and acted to raise awareness of the problem, but we have never tried to attack its basis. Therefore, it is time to identify a practical and concrete roadmap to be shared with that part of the government that has promised us aid and attention. A roadmap that will lead to a comprehensive reconsideration of the Italian system of scientific museums, of their governance, management, organization of research and collections, and communication, through optimization of their resources. When we have redefined the indications for the use and management of scientific museums, they can then be integrated into the National Museum System, guaranteeing protection and enhancement of the material and immaterial heritage. Meanwhile, we must pay close attention to the new developments arising in Europe. It is estimated that the collections of European scientific museums contain about one and a half billion specimens and that 80% of the world's biodiversity described thus far is preserved in the European collections alone. It is easy to comprehend that the systems and techniques used to date are no longer sufficient for the study of these immense masses of data. In a global system of scientific research, it is necessary to recalibrate the very way of thinking about research, also from an epistemological point of view.

"Big Data" is a relatively new term that should be of interest to large scientific museums. When it is talked about, reference is made only to practices already in place in strategic sectors such as remote sensing, financial market trends or weather forecasts. Its application to the study of biodiversity is still in the embryonic stage. Everything suggests that the new approach to Big Data can also be transferred to the analysis of the millions of museum data which are the essential databases for representation of the temporal and spatial distribution of organisms, for an understanding of the diachronicity of the evolutionary process and of populations, and for prediction of the most diverse scenarios, from the modifications caused by current climate changes to ecological, phylogenetic, phenological and biogeographic shifts. In the field of natural sciences, only museums and their collections contain amounts of materials and information able to compete with those cited herein for other fields of knowledge. In the case of the approach to Big Data, modern science absolutely requires digital access to the data. Statistical analyses conducted in recent years tell us that only 10% of the data of European natural history museums are digitized. Dutch, English and German museums have already produced important digital catalogues, while most countries have just begun the process. In addition to digitization, some of the leading museums (only a few in truth) have also started digitizing images (no more than 1% of the specimens). As one can easily understand, the digitization operation is certainly neither simple nor economical. The large numbers, the objective complexity of the process and the need for correctly identified materials require that access to digitization occurs not only with improvement of the adopted technology but also with a more suitable cultural and managerial approach.

This is what we began to hypothesize in Italy a few years ago, recognizing the need for a Diffused Institute of Taxonomy, a delocalized and decentralized structure to streamline the management and taxonomic utilization of the collections present throughout Italy. Today, a broad movement is taking shape in Europe which is demanding the highest level of use of the enormous patrimony of museum collections. Although the overall challenge is ambitious and difficult to face, it is very virtuous and well worth the effort: the key is to unify the access to the data of all the European collections in a coordinated and integrated manner. A first group of European countries is already exploring the technical, financial, political and governmental aspects of a new infrastructure dedicated to the creation of a model of shared management of the entire European patrimony of collections able to support aspects of unification of open access services, incentivization projects, planning, and priorities of research and development of the necessary operational capacities. If and when the infrastructure sees the light, European museums will be able to place the heritage they conserve at the disposal of mankind to address and solve some of the challenges of today's world (loss of biodiversity, impact of climate change...). The large project currently being organized and known as DiSSCo (Distributed System of Scientific Collections) is under evaluation by the European Commission ESFRI and at present involves 21 countries with 150 institutions. Italy is a participant, and this is something for us to be proud of. It is present with a consortium (coordinated by the Museum of Natural History of the University of Florence) which naturally also includes the ANMS. If the project succeeds in becoming part of the European roadmap of large Research Infrastructures, it would last until 2030 and would lead to the supranational unification of the information now preserved in a fragmented and dispersed form. Italy will contribute to DiSSCo, also with the added value of its great wealth of taxonomic expertise. And we all know how much damage can be caused by erroneous or outdated identification of the materials now being digitized (with problems also in the field of metagenomics and others). However, we can cross our fingers in the hope that the wind is changing in Italy, in Europe and in the world.