

SILVIA ORLANDI

## DIGITAL EPIGRAPHY: TRA AUTOMAZIONE E SINGOLARIZZAZIONE

**ABSTRACT:** Recent trends in scholarly publications, public debate on newspapers, and digital projects in the humanities seem to privilege the quantitative and computational approach to historical studies, as it is considered more objective, and therefore more correct. But some examples driven from direct experience in the field of digital epigraphy show how important the human component of any digital project still is for a more complete and correct comprehension of the traces of the past.

**KEYWORDS:** Digital Epigraphy; Big Data; Machine Reading; Visual Recognition

### *Premessa*

Il 25 aprile 2019 Andrea Camilleri, Andrea Giardina e Liliana Segre hanno lanciato, dalle pagine di *Repubblica*, un appello per la difesa della storia (“La storia è un bene comune, salviamola!”):

[https://www.repubblica.it/robinson/2019/04/25/news/la\\_storia\\_e\\_un\\_bene\\_comune\\_salviamola-224857998/](https://www.repubblica.it/robinson/2019/04/25/news/la_storia_e_un_bene_comune_salviamola-224857998/)),<sup>1</sup> che ha subito riscosso un grande successo e conquistato decine di migliaia di sostenitori, sia tra i privati cittadini che tra le istituzioni. Ma, in mezzo a un coro di consensi, non sono mancate voci discordi, come quella di Gilberto Corbellini che, sul *Sole 24 ore* (<https://www.ilsole24ore.com/art/questa-storia-e-davvero-molto-falsa->

---

<sup>1</sup> Tutti i link e le banche dati citati nel testo e nelle note sono stati consultati il 25 aprile 2020.

ABAXajuB),<sup>2</sup> ha polemicamente osservato, a partire dal libro di Alex Rosenberg, *How History Gets Things Wrong* (MIT Press, 2018) come in realtà molti studi storici si basino sulla ‘teoria della mente’, cioè sul quel tratto innato per cui attribuiamo credenze, aspettative, scopi agli altri, e sono quindi falsi, *a parte quelli quantitativi o che usano approcci controllabili* (il corsivo è mio). Sembrerebbe, dunque, che solo una metodologia simile a quella delle scienze esatte possa garantire oggettività e quindi validità agli studi storici. In questa direzione, effettivamente, sembrano muoversi molti studi recenti anche nel campo della storia antica, che valorizzano efficacemente l’analisi scientifica dei cambiamenti climatici o dei resti paleoantropologici e paleoambientali per interpretare fenomeni storici come, ad esempio, la crisi dell’Impero Romano nel III secolo d. C.<sup>3</sup>

Parallelamente, tra le iniziative nel campo delle *Digital Humanities* fortemente sponsorizzate da ‘Europeana’, il portale europeo della cultura, spicca *Time Machine* (<https://www.timemachine.eu/>), un ambizioso progetto che si propone di dare una svolta agli studi umanistici estendendo a questo campo l’analisi dei *Big Data* già applicata con successo ad altri ambiti della cultura scientifica, e sviluppando enormemente, grazie a questa, le possibilità, ad esempio, di lettura e analisi automatica delle fonti storiche digitalizzate. La disponibilità di una grande mole di dati, acquisiti senza una selezione preliminare e con tecniche che abbattano tempi e costi della digitalizzazione e consentano di aumentarne drasticamente la quantità,<sup>4</sup> aprirebbe alla ricerca storica scenari e temi di studio che la mole stessa dei dati potrebbe suggerire.

---

<sup>2</sup> Tra le risposte apparse in seguito alla pubblicazione di questo articolo si veda, per esempio, quella di un gruppo di ricercatori del CNR pubblicata su alfabet2:

<https://www.alfabet2.it/2019/05/26/al-cnr-la-storia-e-una-scienza-una-risposta-allintervento-di-gilberto-corbellini/>

<sup>3</sup> Tra i volumi che hanno avuto più rilevanza nel dibattito su questo tema vanno citati quelli di K. Harper, *Il destino di Roma. Clima, epidemie e la fine di un impero*, Torino, Einaudi, 2017 e W. Scheidel (ed.), *The Science of Roman History. Biology, Climate and the Future of the Past*, Princeton, Princeton University Press, 2018.

<sup>4</sup> Ma senza tenere conto dell’impatto negativo legato alla presenza di dati di scarsa qualità, che aumenta con l’aumentare della mole dei dati, come si ricava dallo studio di P. Woodall, A. Borek, J. Gao, M. Oberhofer, A. Koronios, *An Investigation of How Data Quality is affected by Dataset Size in the Context of Big Data Analytics*, in *19th International Conference on Information Quality (ICIQ)*, Xi’an 2014, p. 24-33.

È dunque di questo che lo studio della storia ha bisogno, di un processo di automazione ‘spinto’ che lo preservi dalla mancanza di oggettività delle intelligenze non artificiali che a questo studio si dedicano?

La questione ha troppe implicazioni – filosofiche, epistemologiche, politiche, culturali – per poter essere esaurita nel breve spazio di una nota. Questo contributo, dunque, ha lo scopo di proporre una riflessione su questo tema a partire dalla personale esperienza di chi scrive nel campo della *Digital Epigraphy*, un settore delle *Digital Humanities* che in questi ultimi anni ha varato numerosi progetti di banche dati, edizioni digitali e applicazioni di diverse tecnologie all’archiviazione, codifica e pubblicazione di oggetti e monumenti iscritti. Molto è stato fatto dai tempi in cui, per consultare il *Corpus Inscriptionum Latinarum*, ci si doveva fisicamente recare in una biblioteca, ma molto, a mio avviso, resta ancora da fare prima che, anche nell’era della cultura digitale, ci si possa dire del tutto indipendenti dal ricorso a metodi di analisi più tradizionali. Attraverso alcuni esempi concreti metterò, dunque, in evidenza l’insostituibile contributo del ricercatore coinvolto in progetti di ricerca integrata tra *digital* e *humanities*.

### *Georeferenziazione*

Si può dire che quasi non esista progetto di digitalizzazione che non preveda anche l’uso di un sistema di georeferenziazione che consenta di localizzare su una mappa virtuale luoghi di rinvenimento o di attestazione di un monumento, un reperto, una fonte scritta, una persona, una parola...<sup>5</sup> Questo prezioso strumento ci ha dato un’ulteriore conferma dell’importanza di considerare i dati, per valutarli correttamente, nel loro contesto geografico e topografico. La cartografia digitale, inoltre, ci permette di visualizzare in modo efficace la distribuzione di un fenomeno, lo spostamento di uomini e merci, la definizione degli spazi. Ma perché tutto questo acquisisca il valore di uno strumento di lavoro per la ricerca storica, le possibilità tecniche vanno raffinate non solo e non tanto nel senso di una

---

<sup>5</sup> Impossibile elencare tutti i progetti di georeferenziazione esistenti, sia come siti autonomi, che come *gazetteer* all’interno di altri progetti. Mi sia consentito citare solo due tra gli strumenti più utilizzati e onnicomprensivi in questo campo: ‘Pleiades’ (<https://pleiades.stoa.org/>) e ‘Trismegistos Places’ (<https://www.trismegistos.org/geo/>).

maggior precisione, quanto piuttosto di un loro uso combinato o integrato con altri dati e informazioni che ne restituiscano pienamente il senso.

Fonti epigrafiche che hanno coordinate geografiche che coincidono al secondo possono non avere lo stesso significato o, meglio, possono essere lette e interpretate in modo che quella stessa coincidenza topografica assuma una valenza storicamente significativa.

Penso, ad esempio, ai *tituli picti* di Pompei, che contengono i manifesti elettorali dipinti sui muri esterni delle case e degli edifici pubblici della città, in cui si leggono i nomi dei candidati alle varie cariche locali e anche dei loro sostenitori.<sup>6</sup> È possibile – e avviene in più di un caso – che su uno stesso muro si sovrappongano, in realtà, più strati di intonaco, frutto del succedersi di più campagne elettorali, che comportavano la copertura dei ‘manifesti’ precedenti con una nuova mano di calce, su cui dipingere i nomi dei nuovi candidati. La posizione geografica dei due manifesti che si sovrappongono è chiaramente la stessa, ma i pochi millimetri di spessore che li separano sono un’informazione-chiave che consente di distinguere cronologicamente i personaggi in essi menzionati (il candidato dello strato superiore appartiene necessariamente ad un anno successivo) e di ricostruire così, comparando la posizione reciproca dei vari nomi e mettendoli a confronto con fonti antiche di altra natura, l’albo municipale di Pompei, con il succedersi dei titolari della varie cariche anno per anno (Fig.1).<sup>7</sup>

---

<sup>6</sup> I *tituli picti* sono pubblicati in un’apposita sezione del IV volume del *Corpus Inscriptionum Latinarum*, dedicato alle iscrizioni parietali di Pompei, il cui ultimo supplemento è uscito nel 2011. Le immagini ad essi relative sono raccolte in A. Varone e G. Stefani (a cura di), *Titulorum Pictorum Pompeianorum qui in CIL Vol. IV collecti sunt Imagines*, Roma, L’Erma di Bretschneider, 2009. In particolare, ai *tituli picti* di carattere elettorale è dedicato il volume di C. Chiavia, *Programmata. Manifesti elettorali nella colonia romana di Pompei*, Torino, Silvio Zamorani Editore, 2002.

<sup>7</sup> Si vedano, ad esempio, i risultati ottenuti e illustrati da M. Stefanile, *La schedatura dei tituli picti pompeiani per EDR: novità sulla cronologia di Suedio Clemente e di alcuni candidati*, in S. Antolini, S.M. Marengo, G. Paci (a cura di), *Colonie e municipi nell’era digitale. Documentazione epigrafica per la conoscenza delle città antiche*. Atti del Convegno di studi (Macerata, 10-12 dicembre 2015), Tivoli, TORED, 2017, p. 601-613.



*Fig. 1: Manifesti elettorali di Pompei con i nomi di due candidati che si sovrappongono<sup>8</sup>*

### *Machine reading*

Un'altra straordinaria possibilità offerta dall'applicazione delle tecnologie digitali in campo umanistico è il cosiddetto *machine reading*, cioè la lettura automatica dei testi da parte di un computer opportunamente addestrato. L'attuale possibilità di utilizzare a questo scopo una grande quantità di dati derivanti dai molti progetti di digitalizzazione e codifica dei testi esistenti in rete rende questa possibilità una tentazione alla quale è effettivamente difficile resistere. Grandi passi avanti sono stati fatti in questo senso nella lettura di papiri e manoscritti, anche se conviene sempre tenere a mente che

---

<sup>8</sup> Foto di Massimo Osanna, Direttore del Parco Archeologico di Pompei, pubblicata sulla pagina *Facebook* di Pompei, Parco Archeologico, al momento della scoperta, nel giugno 2018: <https://www.facebook.com/pompeisoprintendenza/photos/a.1523717371268809/1831449097162300/?type=3&theater>

fonti storiche di questo genere richiedono un approccio che vada al di là della pura codifica testuale, e tenga conto anche degli aspetti materiali del documento (supporto, tecnica di scrittura, eventuali errori e peculiarità...). Recentemente, uno studio sviluppato da un team di studiosi dell'Università di Oxford ha illustrato le nuove, sorprendenti possibilità che si aprono grazie all'applicazione delle tecniche dell'Intelligenza Artificiale e del *Deep Learning* a un campione significativo di testi – letterari ed epigrafici – in lingua greca.<sup>9</sup> L'algoritmo così ottenuto, infatti, è in grado di proporre un numero di possibili integrazioni di testi epigrafici frammentari molto più elevato di quello che menti umane, anche molto allenate, sono in grado di elaborare. Spetta in ogni caso allo studioso il compito di vagliare tali possibilità e di scegliere tra queste non solo quelle statisticamente più rilevanti, ma soprattutto quelle storicamente più attendibili. Resta, tuttavia, il problema posto dal riconoscimento automatico di segni non alfabetici, il cui uso non è raro in epigrafia, o di segni che, pur avendo lo stesso aspetto grafico, abbiano significati diversi a seconda del contesto. Dando semplicemente lo stesso significato agli stessi segni si finirebbe, infatti, per travisare il senso dell'intero testo e, quindi, anche l'uso di dispositivi sofisticati in questo campo continuerebbe, verosimilmente, a richiedere delle cautele intese ad evitare fraintendimenti non secondari.

È quel che accade, ad esempio, in un'iscrizione latina di Reggio Calabria,<sup>10</sup> in cui un personaggio dell'élite locale viene onorato per una serie di benemerienze che vengono puntualmente elencate nell'epigrafe che doveva accompagnare la statua eretta in suo onore. Queste comprendevano il dono di una serie di oggetti preziosi tra cui spiccano una tazza d'argento decorata e un gruppo di statuette di Lari, anch'esse d'argento. In entrambi i casi i doni sono accompagnati, sulla pietra, dai segni P I I S (Fig. 2).

---

<sup>9</sup> Y. Assael, T. Sommerschild, J. Prag, *Restoring Ancient Text Using Deep Learning: A Case Study on Greek Epigraphy*. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Association for Computational Linguistics, Honkong, 2019: <https://www.aclweb.org/anthology/D19-1668/>

<sup>10</sup> Pubblicata in *CIL*, X 6 e consultabile nella banca dati EDR (Epigraphic Database Roma: [www.edr-edr.it](http://www.edr-edr.it)) al record EDR171375.

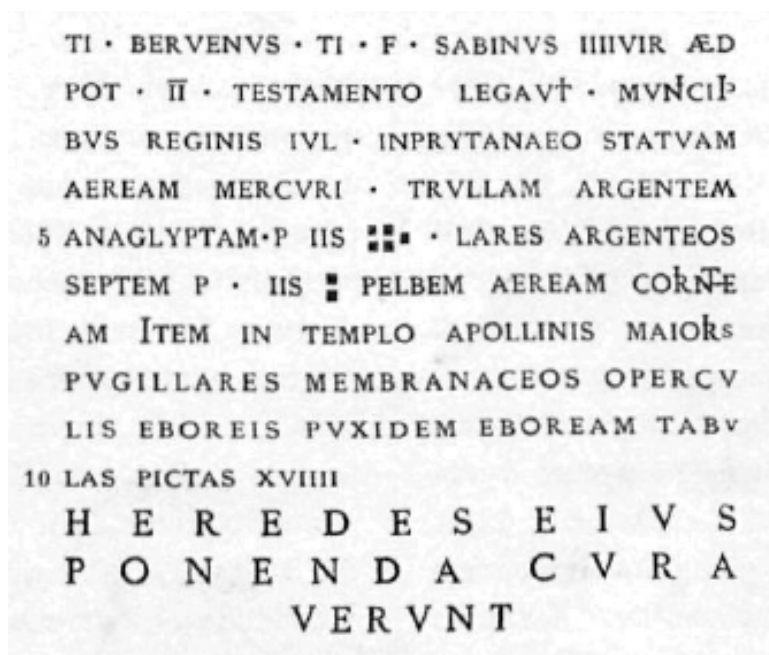


Fig. 2. Iscrizione onoraria da Reggio Calabria, nell'edizione in *CIL*, X 6

Questi segni sono stati letti e trascritti come il dativo plurale *piis* in una delle edizioni digitali di questo testo.<sup>11</sup> In realtà, un occhio esperto sa che in un elenco destinato a sottolineare la generosità di un benefattore, i doni preziosi sono spesso accompagnati dall'indicazione del loro peso, espresso in libbre e sottomultipli di libbra. E infatti anche nel nostro caso gli stessi segni acquistano ben altro significato se interpretati e trascritti come l'abbreviazione del termine *p(ondo)* (cioè libbre) seguito dal numerale II e dall'indicazione dei sottomultipli: S che sta per *semis* (mezza libbra) e 5 onces, cioè dodicesimi di libbra, espressi da segni non alfabetici: 5 punti, che non possono essere ignorati nella lettura perché non sono né accidentali né semplici divisori, ma, appunto, simboli di unità di misura corrispondenti a un'oncia. Siamo sicuri che anche il più sofisticato metodo di lettura automatica dei testi e/o di scioglimento automatico delle abbreviazioni sia

<sup>11</sup> Il record EDCS-11400093 della banca dati Clauss/Slaby ([www.manfredclaus.de](http://www.manfredclaus.de)). Cfr. anche il record EDCS-13201164, in cui si assiste alla stessa omissione dei segni non alfabetici nella trascrizione dell'epigrafe sepolcrale di Roma *CIL*, VI 22879 (corrispondente a EDR174059).

in grado di percepire la differenza di significato che uno stesso segno – alfabetico e non – assume nell’ambito di un diverso contesto? O è legittimo chiedersi se sia ancora insostituibile l’occhio dell’epigrafista, che sa andare al di là dell’aspetto esteriore di un carattere per ricostruirne la sua portata storica?

*Visual recognition system*

Sempre per rimanere nell’ambito dell’aspetto esteriore delle iscrizioni, un’autentica rivoluzione nell’ambito delle possibilità di ricerca e, più in generale, dell’approccio alle iscrizioni, è stata l’applicazione a questo tipo di manufatti del *visual recognition system* messo a punto dagli informatici del CNR-ISTI di Pisa,<sup>12</sup> prima nell’ambito del progetto europeo ‘EAGLE’ ([www.eagle-network.eu](http://www.eagle-network.eu)), e poi con la collaborazione della ‘International Digital Epigraphy Association’ (<https://www.eagle-network.eu/about/who-we-are/>). Questo sistema consente l’individuazione di un’epigrafe, sia in loco che dal computer di casa, in base al riconoscimento delle immagini ad essa associate archiviate in un server e trasformate in altrettanti algoritmi, permettendo, così, non solo di ricercare un testo a prescindere dalle parole da cui è composto, ma anche di reperire più facilmente confronti tipologici per oggetti iscritti caratterizzati dalla stessa forma o decorazione. Il sistema, tuttavia, per funzionare in modo ottimale, richiede un elevato numero di foto, possibilmente riprese in condizioni di luce e posizioni diverse, e in ogni caso non raggiunge – o almeno non raggiunge ancora – un livello di sofisticazione tale da distinguere quei casi – più numerosi di quanto si possa pensare – in cui uno stesso testo epigrafico è stato redatto in più copie, che si distinguono tra loro per dettagli minimi, che solo un occhio esperto, consapevole dell’esistenza di questa possibilità, riesce a cogliere. È così, ad esempio, che è stata individuata l’esistenza di due esemplari quasi identici di un’iscrizione sepolcrale di Roma (pubblicata in *CIL*, VI 24776) che ricorda l’acquisto, da parte di un privato, di più loculi in una tomba collettiva, per cui dobbiamo immaginare che ogni loculo fosse, appunto, contrassegnato da

---

<sup>12</sup> G. Amato, F. Falchi, F. Rabitti, L. Vadicamo, *Inscriptions Visual Recognition. A Comparison of State-of-the-art Object Recognition Approaches*, in *Information Technologies for Epigraphy and Cultural Heritage. Proceedings of the First EAGLE International Conference*, a cura di S. Orlandi, R. Santucci, V. Casarosa, P.M. Liuzzo, Roma, Sapienza Università Editrice, 2014, p. 117-131.

uno stesso testo iscritto che rinvia al titolare delle sepolture. Confusi per lungo tempo dagli editori precedenti, i due testi, oggi conservati rispettivamente ai Musei Capitolini e nel Duomo di Anagni,<sup>13</sup> solo recentemente, grazie all'osservazione attenta di alcuni particolari dell'incisione e dell'impaginazione, sono stati riconosciuti come due oggetti distinti, e quindi più correttamente interpretati anche nel loro significato giuridico di attestazioni di proprietà diverse.<sup>14</sup>

Tutto questo, naturalmente, non mira affatto a 'demonizzare' il ricorso alla tecnologia digitale come strumento ormai imprescindibile nel campo di tutte le discipline umanistiche, ma intende richiamare l'attenzione sull'importanza, ancora irrinunciabile, della componente 'umana' delle *Digital Humanities*, cioè quell'insieme di competenze disciplinari accompagnate da capacità di analisi, di riflessione e di critica senza le quali non si è in grado di produrre validi risultati scientifici. Una componente che ci espone, in quanto esseri fallibili per natura, a un margine di errore, ma che consente anche da un lato di cogliere sfumature e particolari che l'intelligenza artificiale ancora non attinge, e dall'altro di mettere a disposizione della comunità scientifica informazioni di varia natura mediate da un patrimonio di esperienze e di saperi,<sup>15</sup> sfuggendo a quello che è stato recentemente definito "l'inganno contemporaneo della conoscenza disintermediata".<sup>16</sup>

---

<sup>13</sup> Corrispondenti rispettivamente al record EDR100016 e a EDR032608, dove si trovano dati, bibliografia e immagini dei due pezzi.

<sup>14</sup> Vedi in proposito I. Gabrielli, "EDR – Effetti collaterali", *Scienze dell'Antichità*, 21 (1), 2015, p. 248-249.

<sup>15</sup> Quello che Luciano Floridi ha definito "capitale semantico":  
<https://lifeinabyte.com/2019/07/18/il-capitale-semantico-nellera-del-digitale-il-punto-di-vista-di-luciano-floridi/>

<sup>16</sup> In un articolo di Vittorio Pelligra uscito su *Il Sole 24 ore* del 28 luglio 2019 e consultabile al link:

[https://www.ilssole24ore.com/art/l-inganno-contemporaneo-conoscenza-disintermediata-ACuu1hb?fbclid=IwAR2ld16ic5ezUYqKhC1g5Jf2LFIKk\\_xATH6zJnwtVvcs8c8373GSgy07vUI&refresh\\_ce=1](https://www.ilssole24ore.com/art/l-inganno-contemporaneo-conoscenza-disintermediata-ACuu1hb?fbclid=IwAR2ld16ic5ezUYqKhC1g5Jf2LFIKk_xATH6zJnwtVvcs8c8373GSgy07vUI&refresh_ce=1)

REFERENCES:

- Amato, Giuseppe – Falchi, Fabrizio – Rabitti, Fausto – Vadicamo, Lucia, *Inscriptions Visual Recognition. A Comparison of State-of-the-art Object Recognition Approaches, in Information Technologies for Epigraphy and Cultural Heritage*. Proceedings of the First EAGLE International Conference, a cura di S. Orlandi, R. Santucci, V. Casarosa, P.M. Liuzzo, Roma, Sapienza Università Editrice, 2014, p. 117-131.
- Assael, Yannis – Sommerschild, Thea – Prag, Jonathan, *Restoring Ancient Text Using Deep Learning: A Case Study on Greek Epigraphy*. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Association for Computational Linguistics, Honkong, 2019:  
<https://www.aclweb.org/anthology/D19-1668/>
- Chiavia, Catherine, *Programmata. Manifesti elettorali nella colonia romana di Pompei*, Torino, Silvio Zamorani Editore, 2002.
- Gabrielli, Ilaria, “EDR – Effetti collaterali”, *Scienze dell’Antichità*, 21 (1), 2015, p. 248-249.
- Harper, Kyle, *Il destino di Roma. Clima, epidemie e la fine di un impero*, Torino, Einaudi, 2017.
- Scheidel, Walter (ed.), *The Science of Roman History. Biology, Climate and the Future of the Past*, Princeton, Princeton University Press, 2018.
- Stefanile, Michele, *La schedatura dei titoli picti pompeiani per EDR: novità sulla cronologia di Suedio Clemente e di alcuni candidati, in Colonie e municipi nell’era digitale. Documentazione epigrafica per la conoscenza delle città antiche*. Atti del Convegno di studi (Macerata, 10-12 dicembre 2015), a cura di S. Antolini, S.M. Marengo, G. Paci, Tivoli, TORED, 2017, p. 601-613.
- Varone, Antonio – Stefani, Grete (a cura di), *Titulorum Pictorum Pompeianorum qui in CIL Vol. IV collecti sunt Imagines*, Roma, L’Erma di Bretschneider, 2009.
- Woodall, Philip – Borek, Alexander – Gao, Jing – Oberhofer, Martin – Koronios, Andy, *An Investigation of How Data Quality is affected by Dataset Size in the Context of Big Data Analytics, in 19th International Conference on Information Quality (ICIQ)*, Xi’an, 2014, p. 24-33: <https://ualr.edu/informationquality/iciq-proceedings/iciq-2014/>

SILVIA ORLANDI

Sapienza Università di Roma

[silvia.orlandi@uniroma1.it](mailto:silvia.orlandi@uniroma1.it)