



Seminario di

Cultura Digitale
Informatica Umanistica

Informatica umanistica: nuove e vecchie frontiere

Il caso del Tesserarum Sisciae Sylloge

di Marino Vazzana



Indice

<u>1 - Introduzione</u>	p. 3
<u>2 - Il progetto TSS: problematiche e soluzione</u>	
<u>2.1 - Estrazione e catalogazione dei glifi</u>	p. 4
<u>2.1.1 - Lo schema di codifica</u>	p. 4
<u>2.1.2 - Estensibilità dello schema</u>	p. 6
<u>2.1.3 - Ridefinizione delle fasi</u>	p. 7
<u>2.2 - Il calcolatore</u>	p. 7
<u>3 - Conclusioni: l'informatica umanistica</u>	p. 8

1 - Introduzione

Il progetto *Tesserarum Sisciae Sylloge*¹ (da qui in avanti *TSS*), diretto dalla Dottoressa Marion Lamé, verte sull'analisi storica digitale di un corpus di circa 1200 etichette romane in piombo, di datazione compresa tra il I e III secolo d.C., rinvenute in varie battute presso la città di Sisak, allora appartenente alla Pannonia romana e oggi in Croazia.

Le etichette, tutte riconducibili all'attività di una o più tingitorie (*folonicae*), conservano una grande mole di dati sui servizi offerti e relativi prezzi e sono attualmente custodite presso il Museo Archeologico di Zagabria.

Scopo del progetto è la realizzazione di una loro edizione digitale, che segue e amplia l'edizione cartacea prodotta dal Dottor Ivan Radman, che è anche editore scientifico della nuova edizione, finalizzata tanto al trattamento digitale dell'informazione che ad una fruizione chiara ed interattiva delle etichette da parte del pubblico.

Grazie ad una delle attività proposte dalla Professoressa Enrica Salvatori nell'ambito del suo corso di Storia Digitale, ho avuto modo di vedere de visu alcune di queste etichette, e di toccarle con mano presso il Visual Laboratory dell'ISTI-CNR di Pisa.

Quella del contatto con la storia, e con quella romana in specie, è stata un'esperienza complessa e mi ha subito suscitato vivo interesse per questo progetto e per il concetto di edizione digitale di fonti non bibliografiche.

La mia partecipazione, in qualità di tirocinante, al progetto *TSS* è cominciata ad una fase di realizzazione intermedia: già definiti gli obiettivi, gli strumenti e la maggior parte delle collaborazioni, il mio lavoro è cominciato a ridosso dell'inizio della fase operativa, subito dopo l'acquisizione delle immagini *RTI* di un set di 40 etichette pilota, selezionate dal Museo Archeologico di Zagabria tra quelle di lettura più difficile, e per il loro deterioramento e per una più complessa stratificazione del palinsesto.

Il mio lavoro è consistito di due fasi distinte: l'estrazione e la catalogazione, a partire dai disegni messi a disposizione dal Dottor Radman², dei glifi presenti su ciascuna delle etichette-campione, e la realizzazione di una calcolatrice di conversione basata sui dati monetari esatratte dal database precedentemente strutturato da altri collaboratori al progetto, da implementare nel sito web dell'edizione digitale.

¹ Marion Lamé, *Fonti in cerca di storici digitali*, Seminario di Cultura Digitale del 26/05/2014, abstract e video presso l'indirizzo http://labcd.humnet.unipi.it/seminario/cultura_digitale67-32/2014/05/26/fonti-in-cerca-di-storici-digitali/.

² Ivan Radman-Livaja, *Les plombs inscrits de Siscia*, tesi di dottorato, Ecole Pratique Des Hautes Etudes di Parigi, corso di Archeologia romana, 2010, rel. Michel Reddé.

Ognuna delle due fasi ha presentato problemi e soluzioni specifici ma comuni anche ad altri progetti e ad altri settori dell'informatica umanistica. In questo contesto, preziose sono stati i contributi apportati dai vari ospiti, accademici e non, del Seminario di Cultura Digitale.

2 - Il progetto *TSS*: problematiche e soluzione

2.1 - Estrazione e catalogazione dei glifi

La fase di estrazione e catalogazione dei singoli glifi presenti sulle etichette di Sisak, a partire, però, da una fonte secondaria quale i disegni messi a disposizione dal Dottor Radman, aveva diversi obiettivi:

- a) la creazione di un database dei segni presenti sulle etichette e quindi la creazione di uno schema di codifica dei glifi che riuscisse a tenere conto, in maniera esaustiva, di tutte quelle informazioni quali etichetta d'appartenenza, verso, layer del palinsesto, rigo, posizione, orientamento e tipo, tale da poter essere applicato a tutti i casi possibili, ma che fosse abbastanza semplice da poter essere interpretato rapidamente tanto da una macchina che da chiunque fosse adeguatamente informato sullo schema di codifica;
- b) l'estensibilità di questo schema anche ad etichette ulteriori rispetto a quelle appartenenti al set pilota e la verifica della sua applicabilità anche da parte di terzi con una formazione non specialistica;
- c) verifica della possibilità di definire fasi più precise di appartenenza delle etichette attraverso la datazione dei tipi che vi compaiono;

2.1.1 - Lo schema di codifica

La creazione di uno schema di codifica dei glifi ha presentato problematiche tanto quantitative che metodologiche.

Inizialmente, infatti, mi era stata prospettata una settimana di lavoro circa per l'analisi di tutti i glifi delle 40 etichette pilota, la cui media era stata stimata attorno alle 20 unità per etichetta. Invece, le etichette del campione contengono spesso tra i 30 e i 40 glifi, con una media finale di 35 glifi per etichetta.

Ciò ha sostanzialmente raddoppiato la mole di lavoro, e di conseguenza il tempo necessario al suo completamento. Infatti, essendo la codifica delle informazioni e il loro trattamento automatico le finalità di questo progetto, il lavoro ha anche avuto bisogno di revisioni successive, per eliminare gli inevitabili errori dovuti alla grande mole di glifi estratti e valutati “manualmente” (1400 su 40 etichette) e per adattare lo schema di codifica alle successive indicazioni del modellista del database, le cui specifiche erano anch'esse in concomitante evoluzione.

Ad esempio si era inizialmente deciso di indicare il rigo di appartenenza del glifo con un numerale romano, scelta in seguito abbandonata su richiesta del modellista in favore dell'indicazione R+numero, di computazione più semplice.

Da un punto di vista metodologico, invece, la definizione dello schema di codifica ha presentato diverse difficoltà, alcune delle quali abbastanza sottili. Ho avuto piena autonomia nel valutare quali informazioni sui glifi andassero codificate e quali meno.

Ad esempio, ho inizialmente provato, su un campione di 10 etichette, ad assegnare dei tag speciali ai glifi quand'essi erano in funzione di numero o di simbolo monetario, per valutare se non vi fosse una qualche differenza, statisticamente, nella loro grafia rispetto agli omologhi alfabetici. Questo lavoro si è rivelato infruttuoso e queste informazioni sono state eliminate dallo schema di codifica finale.

Altri problemi sono derivati dall'assegnazione di ogni glifo ad uno degli archetipi della lettera rappresentata: quelli individuati dal Dottor Radman si sono rivelati a volte insufficienti, e in specie nel caso della *p*: mancava nel suo lavoro un archetipo di *p* corsiva con pancia triangolare, in verità abbastanza frequente nel set pilota, così come una catalogazione dei tipi di simboli monetari.

Un'altra difficoltà è stata quella relativa all'interpretazione di un certo numero di glifi appartenenti agli strati anteriori del palinsesto, che sono stati solo occasionalmente trascritti nell'edizione cartacea, così come difficoltosa è stata l'interpretazione di alcuni glifi che costituivano nessi e di alcuni glifi dall'interpretazione incerta, che vengono indicati mediante un punto nelle trascrizioni, ma con una corrispondenza spesso inesatta tra il numero di punti e il numero di glifi discernibili sulle etichette.

Lo schema di codifica è riuscito a tenere conto di tutte queste problematiche, e la sua forma finale è:

[tag ID (xx.yy)]-[side (AV for front, REV for back)]-[layer (F for first, S for second)]-[line (R+integer)]-[prototype (e.g. A11, F3)]

Sono stati previsti dei tag speciali, da sostituire all'indicazione del prototipo, finalizzati appunto a risolvere le difficoltà interpretative e segnalarle agevolmente all'epigrafista, il Dottor Radman.

In specie, il tag NOT è stato applicato a tutti quei segni, spesso (ma non sempre) di facile interpretazione, che non comparivano nelle trascrizioni.

Il tag UND è stato applicato a tutti quei glifi la cui interpretazione rimane dubbia. E' l'unico tag che dovrebbe rimanere nella versione definitiva del database, in quanto, a meno di nuove scoperte, i glifi non interpretati sono solitamente troppo abrasivi per essere letti.

A cavallo tra questi si colloca il tag REV. Segnala la necessità di una revisione, da parte dell'epigrafista, di segni già da lui interpretati, ed è stato applicato ai casi in cui non si è riusciti a comprendere le sue scelte o è stato trovato un errore. La più alta concentrazione di questo tag, comunque temporaneo, si è solitamente avuta in caso di nesi complessi.

La collaborazione con l'epigrafista è stata comunque quanto mai cordiale ed orizzontale. Il Dottor Radman ha dimostrato sempre la massima attenzione alle opinioni altrui, ed è sempre stato solerte nel dissipare dubbi concettuali ed interpretativi.

2.1.2 - Estensibilità dello schema

In vista della collaborazione telematica con un gruppo di sei studenti canadesi, il cui compito era quello di analizzare 26 etichette ciascuno (tanti sono i gruppi di etichette individuati dal Dott. Radman in base a alle abbreviazioni o ai nomi citati³), e della loro formazione non specialistica (nella maggior parte dei casi i corsi di storia facevano parte di un *minor*), è stato necessario preparare un tutorial video⁴ e uno testuale, focalizzati il primo sul riconoscimento e l'analisi dei glifi, nonché sulle modalità della loro estrazione (realizzato da me), sullo schema di codifica il secondo (scritto in collaborazione con la Dottoressa Lamé).

Questa fase del progetto, denominata dalla Dottoressa Lamé *Digital Autoptic Process (DAP)*, ha influenzato non poco le metodologie da me usate nella creazione dello schema di codifica stesso e nell'estrazione dell'immagine: i risultati dovevano avere *consistency*, nonostante l'utilizzo di un processo semplificato, con strumenti *open source* o integrati nei principali sistemi operativi.

Gli strumenti necessari per l'estrazione e la catalogazione dei segni sono stati ridotti al minimo: nel processo, oltre alla documentazione e ai tutorial preparati all'uopo, la dotazione necessaria è stata ridotta ad un qualsiasi pdf viewer ed un qualsiasi editor d'immagini, con la sola restrizione di rispettare pedissequamente lo schema di codifica, il formato di output delle immagini così ricavate e le loro dimensioni.

Sono inoltre stati messi a disposizione degli studenti vari fogli di calcolo dove potevano presentare in itinere i propri risultati, e segnalare a me e agli altri collaboratori, ed in particolare al Dottor Radman, i propri dubbi interpretativi.

³ I. Radman, *ibid.*, p. 26.

⁴ Tutorial visionabile presso questo [link](#).

Le etichette così analizzate sono state ben 156. Lo schema di codifica si è rivelato sempre adeguato, e la maggior parte degli studenti ha lavorato presto e bene. Questo risultato è un aspetto molto importante del progetto: i fondi a disposizione per il suo completamento sono esigui, e seppur non siano state utilizzate tecnologie innovative, lo scopo del progetto *TSS* è una fruizione innovativa di questi reperti ed una loro nuova edizione scientifica. L'alternativa al trovare strumenti economici e di semplice utilizzo è la rinuncia al sapere stesso, e il progetto *TSS* è riuscito a trovare un metodo di lavoro funzionale, economico, modulare e con il massimo grado di interoperabilità.

2.1.3 - Ridefinizione delle fasi

Le etichette di Siscia hanno una datazione compresa tra il I e il III secolo d.C., desumibile da vari indizi. Il più evidente è il sistema monetario usato nelle transazioni, un sistema imperiale antecedente alla riforma di Diocleziano.

Datosi il prezzo mediamente basso delle transazioni, la moneta di riferimento è il denario, di medio valore.

Un altro elemento importante è la grafia di particolari lettere, tra cui marcatamente la *A*, la *C*, la *M*, la *R*, che ricalca fasi diverse, presentando forme anteriori e tarde.

La modellazione del database, che riconduce ogni glifo ad un archetipo, ci permetterà quindi di ipotizzare la datazione delle singole etichette, che può essere raffinata ulteriormente incrociando i dati così ottenuti con dati economici e con i database epigrafici.

Per valutare la fattibilità e la validità di questo approccio bisognerà aspettare il completamento del database e la sua estensione a tutte le etichette di Siscia.

Pur nell'impossibilità di arrivare a datazioni precise, qualora si riuscisse a contestualizzare le etichette in questo *continuum*, ciò ci darebbe la possibilità di valutarne l'evoluzione, soprattutto economica, che è poi l'evoluzione della vita quotidiana e lavorativa stessa.

2.2 - Il calcolatore

Come progetto collaterale, dato il carattere prevalentemente economico delle etichette, e la fitta presenza di prezzi, è stato ritenuto utile procedere alla programmazione di un calcolatore/convertitore, indirizzato al *general public*, che consentisse tanto la conversione tra numerali romani e numeri arabi che la conversione tra le monete, all'interno del sistema imperiale di riferimento.

La programmazione di questo tool, avvenuta in jQuery, è stata rapida e scevra da particolari problematiche, ed è in attesa di essere collocato sul sito web dell'edizione digitale, la cui pubblicazione avverrà comunque a Dicembre 2014.

Si discute sulla possibilità di ampliare o meno le funzionalità del calcolatore, ad esempio sulla possibilità di permettere di paragonare i prezzi con quelli di altre merci o epoche o zone geografiche.

Un tentativo in questo senso era stato fatto da me con lo studio dell'*Edictum De Pretiis Rerum Venalium* di Diocleziano (301 d.C.), rivelatosi però inadeguato: esso era infatti un decreto emergenziale, promulgato per far fronte ad una grave crisi inflazionistica e speculativa, frutto anche di un nuovo sistema monetario.

Pur facendo le dovute equivalenze tra i vari sistemi monetari, usando come termine di paragone il contenuto in metallo delle varie monete, bisogna ricordare che anche nel primo sistema imperiale vi furono degli aggiustamenti dei valori nominali delle monete: l'impossibilità di datare più precisamente le etichette di Siscia pone dei problemi metodologici ancora aperti anche in quest'ambito.

La finalizzazione di questo tool è ancora auspicabile, e le sue funzionalità dipenderanno dal livello di conoscenza delle etichette che sarà raggiunto. Anche nella sua forma compatta, comunque, rimane un utile ausilio per l'utenza non specialistica.

3 - Conclusioni: l'informatica umanistica

Il caso del Tesserarum Sisciae Sylloge è stato per me il primo impatto a tutto tondo con l'Informatica Umanistica.

Avevo già studiato materie e di ambito umanistico e di ambito informatico, ma, nonostante la natura collaborativa di questo settore sia spesso enfatizzata, devo ammettere che soltanto la partecipazione attiva ad un progetto mi ha dato la possibilità di rendermi conto delle implicazioni di questa disciplina-ponte.

In effetti, a mio avviso e per la mia esperienza, il ruolo dell'informatica umanistica è quello di mediare tra le diverse discipline specialistiche, trovare loro un linguaggio comune che permetta la definizione e la realizzazione di un progetto, di un nuovo metodo di fruizione e quindi di una nuova comprensione.

Mentre questa appare la caratteristica macroscopica delle *digital humanities*, molto sfaccettata e composita appaiono le modalità di questa mediazione, dipendenti in larga misura dall'oggetto di studio e dal target della fruizione. Il target, in particolare, è un elemento affatto secondario, e la sua definizione è prioritaria nell'ambito di un progetto, perchè seleziona le modalità di produzione e la complessità stessa dei risultati proposti.

Infatti l'informatica umanistica risponde tanto al bisogno di condivisione, presentazione e narrazione della conoscenza che al bisogno di strumenti innovativi di indagine.

Risponde, ad esempio, ad un bisogno diffuso da parte del pubblico di Storia, come esemplificato dal seminario “Web e discipline umanistiche: due esempi di integrazione”, tenuto da Luca Novarino⁵, dove ha presentato il progetto Memoro: una piattaforma di condivisione di testimonianze audio/video della storia recente, raccontate da protagonisti comuni, con più versioni localizzate e la possibilità di inserire commenti e materiale proprio, il che aiuta a rafforzare il senso di *community* e il suo coinvolgimento. In questo caso, dato anche un target il più ampio possibile e la necessità (quasi onnipresente) di misurarsi con un budget, non è stata l'introduzione di tecnologie innovative (né nella raccolta delle testimonianze né nella loro fruizione) a fare la fortuna del progetto, quanto una mediazione intelligente tra una disciplina, la Storia, solitamente molto poco *social*, con una piattaforma di condivisione adeguata: uno dei capisaldi della Storia Digitale.

Altri tipi di materiali pongono problemi metodologici e tecnologici di complessità maggiore. A questo proposito, esemplificativo è stato il seminario “Informatica umanistica e(/è) multidisciplinarietà: il caso del progetto Visionary Cross”, tenuto dal Professor Roberto Rosselli Del Turco⁶, attivo tanto nell'Università del Piemonte Orientale quanto nell'ateneo pisano.

Nel suo seminario il professore descriveva la sua esperienza nell'ambito del progetto *Visionary Cross*, il cui scopo è la produzione di un'edizione digitale complessiva di materiale epigrafico e bibliografico, quali la Croce di Ruthwell e quella di Bewcastle dell'VIII secolo e della Croce di Bruxelles e il Vercelli Book, entrambi del X secolo.

I materiali appaiono collegati dalla provenienza (a discapito dei nomi sono tutti reperti dell'Inghilterra sud-orientale) e dalla presenza di stralci del Poema della Vera Croce, in forma epigrafica sulla Croce di Ruthwell e in forma bibliografica nel Vercelli Book.

Le difficoltà metodologiche incontrate dal Professor Del Turco sono concettualmente simili a quelle affrontate da me e dagli altri partecipanti al progetto *TSS*, e sono stato confortato nella mia esperienza dalla sua, nella continua ricerca di soluzioni ai problemi ancora aperti.

A livello metodologico, sottolinea giustamente il Professor Rosselli Del Turco, un progetto di informatica umanistica è per definizione un progetto multidisciplinare, e coinvolge competenze e punti di vista diversi. In particolare, nel progetto *Visionary Cross* convergono filologia, storia dell'arte, *museum studies*, grafica e realtà virtuale.

⁵ Luca Novarino, *Web e discipline umanistiche: due esempi di integrazione*, Seminario di Cultura Digitale del 4 Dicembre 2014, abstract e video presso l'indirizzo http://labcd.humnet.unipi.it/seminario/cultura_digitale67-32/2013/12/10/luca-novarino-web-e-discipline-umanistiche-due-esempi-di-integrazione/.

⁶ Roberto Rosselli Del Turco, *Informatica umanistica e(/è) multidisciplinarietà: il caso del progetto Visionary Cross*, Seminario di Cultura Digitale del 26 Marzo 2014, abstract e video presso l'indirizzo http://labcd.humnet.unipi.it/seminario/cultura_digitale67-32/2014/05/26/informatica-umanistica-ee-multidisciplinarieta-il-caso-del-progetto-visionary-cross/.

Alla difficoltà di trovare un equilibrio fra le varie discipline, difficoltà comunque ineludibile a causa della natura non neutrale di qualsiasi interpretazione e narrazione, si aggiunge la difficoltà nel definire un target limitato, o, per converso, di confezionare un prodotto multi-target.

Infatti, una fruizione ottimale implica, a mio avviso, che il fruitore riceva la giusta quantità di informazione, sufficiente a consentirgli esplorazioni e percorsi semiautonoma ma non troppo abbondante tanto da risultare dispersiva.

La creazione di tali percorsi, anch'essa non neutrale, rappresenta già di per sé uno scoglio metodologico non indifferente, ma è ancora più difficile districarsi quando si voglia avere un target il più ampio possibile.

Vi sono varie possibilità, come la creazione di diverse edizioni finalizzate del proprio progetto, come è stato deciso di fare per il *Visionary Cross*, ove è stata prevista un'edizione per l'insegnamento ed una per gli specialisti, fino alla creazione di un'edizione unica, con tools che facilitino la comprensione, come nel caso del progetto *TSS*.

Queste scelte non dipendono solo dall'orientamento degli studiosi che vi lavorano, è anche il risultato di un equilibrio tra la committenza e l'oggetto di studio.

Nel caso del progetto *TSS*, ad esempio, il committente è il Museo Archeologico di Zagabria (un ente museale è interessato alla massima ampiezza del target), oggetto dell'edizione è un set di etichette omogenee per fattura e contenuti, le cui incisioni presentano relativamente pochi problemi interpretativi, che si spostano dal testo alla storia stessa e alla collocazione temporale delle etichette; la scelta di un'edizione digitale unica sembra quindi particolarmente indicata.

Una delle problematiche fondamentali dell'informatica umanistica, che forse andrebbe sviluppata più estesamente, però, è quella della formalizzazione di concetti di discipline specialistiche, tali da poter essere trattate informaticamente, soprattutto in progetti multidisciplinari in cui convergono specialisti di diverso ambito e soprattutto di diverso linguaggio. La creazione di un linguaggio comune è forse una delle frontiere più affascinante dell'informatica umanistica.

Vi accennava il Dottor Gabriele Gattiglia nel suo seminario “@MappaProject”⁷, quando parlava della difficoltà di formalizzare ed omogeneizzare a livello informatico concetti archeologici e geologici.

Anche nel progetto *TSS* vi sono state di queste difficoltà, in particolare sottile e di difficile computazione appariva la distinzione del Dottor Radman fra i prezzi dei servizi

⁷ Gabriele Gattiglia, @MappaProject, Seminario di Cultura Digitale del 6 Novembre 2013, abstract e video presso l'indirizzo http://labcd.humnet.unipi.it/seminario/cultura_digitale67-32/2013/11/06/gabriele-gattiglia-mappaproject/.

definti incerti e quelli definiti ipotetici. Entrambi si riferiscono ad etichette di lettura più difficile, con indicazione del prezzo del tutto o parzialmente illegibile. Con prezzi ipotetici il Dottor Radman si riferisce a prezzi ricostruiti, ma lo stesso tipo di integrazione è stata compiuta sui prezzi incerti, e una ridefinizione più precisa e la decisione di un metodo per il trattamento differenziale di queste informazioni hanno richiesto diversi giorni di discussione.

Il ventaglio di applicazioni dell'informatica umanistica è quanto mai ampio, e spazia in tutti gli ambiti, dalla linguistica computazionale alle tecnologie del linguaggio, alla produzione multimediale, alla storia digitale, alla filologia, alle biblioteche all'editoria digitali, tutti settori in cui l'informatica ha migliorato la fruizione e la comprensione di materiali e discipline già esistenti.

Quella che a mio avviso è la nuova frontiera della disciplina risiede appunto nel continuo sviluppo e affinamento di questo linguaggio comune, nella continua formalizzazione di concetti umanistici e specialistici in generale, tali che portino una riflessione a monte, da parte degli studiosi, e a valle, nelle infinite possibilità di combinazione fra discipline e settori che tale linguaggio consentirebbe, un'autentica rivoluzione nei modi e nelle forme del sapere.