



The W.A.L.(L) Project

**Francesca Buscemi
CNR - ISPC Catania**

The idea and the context

The W.A.L.(L) Project was born within the framework of the ‘Laboratori Archeologici Internazionali Congiunti’ (Joint International Archaeological Laboratories) of the CNR-ISPC and in the scientific context of the topography and survey activities of the Archaeological Mission in Festòs of the University of Catania.

The Project is aimed at the application of quantitative analysis methods and Machine Learning to ancient architecture and the creation of a dedicated research infrastructure, based on open-source technology, according to the principles of Open Science.

The results of the 2015-2018 excavation campaigns at Festòs, with the identification of an entire phase of reoccupation of the Palace in the Sub-Minoan-Geometric period (11th-8th century BC), have suggested using as sample structures some walls from this period, until recent years little investigated in Cretan sites where the magnificent Minoan phase (II millennium) had attracted, by preference, the attention of scholars.

A multidisciplinary research team, starting with 3D photogrammetric models, has created a training dataset of around 1,300 digital stone models belonging to twelve walls datable from the Late Minoan IIIC (1200-1050 BC) to the entire Geometric period (8th century BC) and located in four Cretan archaeological sites (Festòs, Haghia Triada, Sissi, Anavlochos).

The aim

Given the high potential of building material as a historical, social and cultural information tool, the challenge of the research is to make the three-dimensional digital data (in our case, the virtual models of the wall stones) interrogable by means of a special script that, on the basis of algorithms constructed by archaeologists and computer scientists, extracts significant numerical characteristics. All this in order to: statistically identify construction practices (methods of processing and setting stones) (Fig. 3, Geometric Period indicators identified on the basis of recurrence); assess continuity/change phenomena of the practices

themselves, attributable to: tradition, group identity, chronology; contribute to the definition of a relative intra-site chronology of the walls; identify restoration patterns.

A dedicated web database allows the end user to query the walls through a user friendly interface. The elaboration of the database's Logical Model was the occasion for important research on the conceptualisation and semantic classification of the data: the absence of databases specifically addressed to ancient architectural data, in fact, entailed, in particular, careful work on vocabularies, fundamental for the very communication between the archaeologists who, coming from different parts of the world, are active in Crete.

The research team

Principal investigator: Francesca Buscemi (CNR-ISPC), Classical Archaeologist and Topographer; Partner Coordinator: Jan Driessen (Université Catholique de Louvain), Aegean Archaeologist; Partner: Maude Devolder (Ghent University), Aegean Archaeologist; Marianna Figuera (University of Catania), Aegean Archaeologist; Florence Gaignerot-Driessen (University of Cincinnati), Aegean Archaeologist; Giovanni Gallo (University of Catania), Mathematician and Computer Scientist; Angelica Lo Duca and Andrea Marchetti (CNR-IIT), Computer Scientists.

Methodology

Artificial Intelligence and Prehistoric and Protohistoric Architecture

The extraction of qualitative features (Fig. 2, Segmentation, semantic classification and recurrence analysis of building material types. Right: minimal bounding box for analysing the orientation axes of the stones) from the 1,300 virtual stone samples identified as significant by the archaeologist (size, location, orientation, position within the wall, type, degree of workmanship) enabled the start of some tests to study ancient architecture using Machine Learning (Fig. 4, Comparison of the 'feature importance' according to the four algorithms tested). The latter is based on 'training' the computer in reasoning by inputting a large amount of data that are then analysed.

The experiment of the W.A.L.(L) Project has highlighted some particularly challenging characteristics of ancient architecture (and specifically of the prehistoric one) for these procedures, first and foremost concerning the poor predictivity of masonry (i.e. the uniqueness or extreme irregularity of each building element and of the walls as a whole) which requires on

the one hand particularly large training datasets, and on the other hand the elaboration of complex algorithms for the definition of very specific numerical features.

Nevertheless, some promising initial results have been achieved, for example in the automatic recognition of unprocessed construction material types (wedges and rubble stones) (Fig. 5, Mesh views of unworked facing stones (a,b,c) and wedges (d,e,t)): of the four algorithms tested, the best result was achieved by the random Forest classification, with an accuracy of 83.85%.

Collaborative processes and open science

All the processing and use of the data of the W.A.L.(L) Project are conceived in the direction that research is going in today at the European level: that of open knowledge, open access, open publishing, or in a single concept, Open Science. Data, that is, to be truly useful, must be 'findable', 'accessible', 'searchable', and 'reusable'.

The applications used in the Project are therefore Open Source: the Blender software (Fig. 6, Visualisation of a query in Blender), which was employed for the segmentation of the 3D models of the walls, from which the individual stones were virtually separated, as well as for the Machine Learning operations; and the Django framework, employed to orchestrate various modules of the Project: from the import of spreadsheets from Blender into the database, to the generation of data entry forms (interfaces) for its manual populating (Fig. 1, Project workflow).

Finally, the web interface enabling the query of digital resources, which is currently accessible with credentials, enables a collaborative data implementation process by the W.A.L.(L) Project partners and is prepared for public use of the database's calculation potential.



ΤΟ ΠΡΟΓΡΑΜΜΑ W.A.L.(L)

Francesca Buscemi
CNR - ISPC Κατάνια

Η ιδέα και το πλαίσιο

Το έργο W.A.L.(L) γεννήθηκε στο πλαίσιο των "Laboratori Archeologici Internazionali Congiunti" (Κοινά Διεθνή Αρχαιολογικά Εργαστήρια) του CNR-ISPC και στο επιστημονικό πλαίσιο των δραστηριοτήτων τοπογραφίας και έρευνας της Αρχαιολογικής Αποστολής στη Φαιστό του Πανεπιστημίου της Κατάνια.

Το έργο αποσκοπεί στην εφαρμογή μεθόδων ποσοτικής ανάλυσης και μηχανικής μάθησης στην αρχαία αρχιτεκτονική και στη δημιουργία ειδικής ερευνητικής υποδομής, βασισμένης σε τεχνολογία ανοικτού κώδικα, σύμφωνα με τις αρχές της ανοικτής επιστήμης. Τα αποτελέσματα των ανασκαφικών εκστρατειών της περιόδου 2015-2018 στη Φαιστό, με τον εντοπισμό μιας ολόκληρης φάσης επανακατοίκησης του ανακτόρου κατά την υπομινωική-γεωμετρική περίοδο (11ος-8ος αιώνας π.Χ.), πρότειναν να χρησιμοποιηθούν ως δειγματικές κατασκευές ορισμένοι τοίχοι αυτής της περιόδου, που μέχρι τα τελευταία χρόνια ελάχιστα είχαν ερευνηθεί σε κρητικές τοποθεσίες, όπου η μεγαλειώδης μινωική φάση (II χιλιετία) είχε προσελκύσει, κατά προτίμηση, την προσοχή των μελετητών.

Μια διεπιστημονική ερευνητική ομάδα, ξεκινώντας από τρισδιάστατα φωτογραμμετρικά μοντέλα, δημιούργησε ένα σύνολο εκπαιδευτικών δεδομένων από περίπου 1.300 ψηφιακά μοντέλα λίθων που ανήκουν σε δώδεκα τοίχους που χρονολογούνται από την Υστερομινωική ΙΙΙ (1200-1050 π.Χ.) έως ολόκληρη τη Γεωμετρική περίοδο (8ος αιώνας π.Χ.) και βρίσκονται σε τέσσερις κρητικούς αρχαιολογικούς χώρους (Φαιστός, Αγία Τριάδα, Σίσι, Αναβλόχος).

Ο στόχος

Δεδομένης της μεγάλης δυναμικής του οικοδομικού υλικού ως εργαλείου ιστορικής, κοινωνικής και πολιτιστικής πληροφόρησης, η πρόκληση της έρευνας είναι να καταστήσει τα τρισδιάστατα ψηφιακά δεδομένα (στην περίπτωση μας, τα εικονικά μοντέλα των πέτρινων τοίχων) επεξεργάσιμα μέσω ενός ειδικού σεναρίου το οποίο, βάσει αλγορίθμων που έχουν κατασκευαστεί από αρχαιολόγους και επιστήμονες πληροφορικής, εξάγει σημαντικά αριθμητικά χαρακτηριστικά. Όλα αυτά προκειμένου: να προσδιοριστούν στατιστικά οι κατασκευαστικές πρακτικές (μέθοδοι επεξεργασίας και τοποθέτησης των λίθων) (Εικ. 3, δείκτες γεωμετρικής περιόδου που προσδιορίζονται με βάση την επανάληψη), να εκτιμηθούν φαινόμενα συνέχειας/αλλαγής των ίδιων των πρακτικών, που αποδίδονται σε: παράδοση, ομαδική ταυτότητα, χρονολογία, να συμβάλουν στον καθορισμό μιας σχετικής ενδοτοπικής χρονολογίας των τοίχων, να εντοπιστούν πρότυπα αποκατάστασης.

Μια ειδική διαδικτυακή βάση δεδομένων επιτρέπει στον τελικό χρήστη να αναζητήσει τα τείχη μέσω μιας φιλικής προς το χρήστη διεπαφής. Η εκπόνηση του Λογικού Μοντέλου της

βάσης δεδομένων αποτέλεσε την αφορμή για σημαντική έρευνα σχετικά με την εννοιολόγηση και τη σημασιολογική ταξινόμηση των δεδομένων: η απουσία βάσεων δεδομένων που απευθύνονται ειδικά σε αρχαία αρχιτεκτονικά δεδομένα, στην πραγματικότητα, επέβαλε, ιδίως, προσεκτική εργασία σε λεξιλόγια, θεμελιώδους σημασίας για την ίδια την επικοινωνία μεταξύ των αρχαιολόγων που, προερχόμενοι από διαφορετικά μέρη του κόσμου, δραστηριοποιούνται στην Κρήτη.

Η ερευνητική ομάδα

Κύριος ερευνητής: Francesca Buscemi (CNR-ISPC), Κλασική Αρχαιολόγος και Τοπογράφος- Συντονιστής εταίρος: Jan Driessen (Université Catholique de Louvain), αρχαιολόγος του Αιγαίου: Maude Devolder (Πανεπιστήμιο της Γάνδης), αρχαιολόγος του Αιγαίου- Marianna Figuera (Πανεπιστήμιο της Κατάνιας), αρχαιολόγος του Αιγαίου- Florence Gaignerot-Driessen (Πανεπιστήμιο του Σινσινάτι), αρχαιολόγος του Αιγαίου- Giovanni Gallo (Πανεπιστήμιο της Κατάνιας), μαθηματικός και επιστήμονας πληροφορικής- Angelica Lo Duca και Andrea Marchetti (CNR-IIT), επιστήμονες πληροφορικής.

Μεθοδολογία

Τεχνητή νοημοσύνη και προϊστορική και πρωτοϊστορική αρχιτεκτονική

Η εξαγωγή ποιοτικών χαρακτηριστικών (Εικ. 2, Τμηματοποίηση, σημασιολογική ταξινόμηση και ανάλυση επανάληψης των τύπων δομικών υλικών. Δεξιά: ελάχιστο πλαίσιο οριοθέτησης για την ανάλυση των αξόνων προσανατολισμού των λίθων) από τα 1.300 εικονικά δείγματα λίθων που αναγνωρίστηκαν ως σημαντικά από τον αρχαιολόγο (μέγεθος, θέση, προσανατολισμός, θέση εντός του τοίχου, τύπος, βαθμός επεξεργασίας) επέτρεψε την έναρξη ορισμένων δοκιμών για τη μελέτη της αρχαίας αρχιτεκτονικής με τη χρήση Μηχανικής Μάθησης (Εικ. 4, Σύγκριση της "σημασίας των χαρακτηριστικών" σύμφωνα με τους τέσσερις αλγορίθμους που δοκιμάστηκαν).

Η τελευταία βασίζεται στην "εκπαίδευση" του υπολογιστή στη συλλογιστική με την εισαγωγή μεγάλου όγκου δεδομένων τα οποία στη συνέχεια αναλύονται. Το πείραμα του προγράμματος W.A.L.(L) ανέδειξε ορισμένα ιδιαίτερα απαιτητικά χαρακτηριστικά της αρχαίας αρχιτεκτονικής, και ιδίως της προϊστορικής περιόδου, για τις διαδικασίες αυτές, με πρώτο και καλύτερο την κακή προβλεψιμότητα της τοιχοποιίας, δηλαδή τη μοναδικότητα ή την ακραία ακανόνιστη μορφή κάθε δομικού στοιχείου και των τοίχων στο σύνολό τους, η οποία απαιτεί αφενός ιδιαίτερα μεγάλα σύνολα δεδομένων εκπαίδευσης και αφετέρου την

επεξεργασία πολύπλοκων αλγορίθμων για τον ορισμό πολύ συγκεκριμένων αριθμητικών χαρακτηριστικών.

Παρ' όλα αυτά, έχουν επιτευχθεί κάποια πολλά υποσχόμενα πρώτα αποτελέσματα, για παράδειγμα στην αυτόματη αναγνώριση των τύπων ανεπεξέργαστων δομικών υλικών (σφήνες και θραύσματα) (Εικ. 5, Ματιές από μη επεξεργασμένους λίθους επένδυσης (a,b,c) και σφήνες (d,e,t)): από τους τέσσερις αλγορίθμους που δοκιμάστηκαν, το καλύτερο αποτέλεσμα επιτεύχθηκε από την ταξινόμηση με τυχαίο δάσος, με ακρίβεια 83,85%.

Συνεργατικές διαδικασίες και ανοικτή επιστήμη

Όλες οι επεξεργασίες και η χρήση των δεδομένων του έργου W.A.L.(L) σχεδιάζονται προς την κατεύθυνση που ακολουθεί σήμερα η έρευνα σε ευρωπαϊκό επίπεδο: αυτή της ανοικτής γνώσης, της ανοικτής πρόσβασης, της ανοικτής δημοσίευσης ή, με μια ενιαία έννοια, της ανοικτής επιστήμης. Τα δεδομένα, δηλαδή, για να είναι πραγματικά χρήσιμα, πρέπει να είναι "ανακτήσιμα", "προσβάσιμα", "αναζητήσιμα" και "επαναχρησιμοποιήσιμα".

Οι εφαρμογές που χρησιμοποιήθηκαν, επομένως, στο Έργο είναι Ανοιχτού Κώδικα: το λογισμικό Blender (Εικ. 6, Οπτικοποίηση ενός ερωτήματος στο Blender), το οποίο χρησιμοποιήθηκε για την κατάτμηση των τρισδιάστατων μοντέλων των τοίχων, από τα οποία διαχωρίστηκαν εικονικά οι μεμονωμένες πέτρες, καθώς και για τις λειτουργίες Μηχανικής Μάθησης- και το πλαίσιο Django, το οποίο χρησιμοποιήθηκε για την ενορχήστρωση διαφόρων ενοτήτων του Έργου: από την εισαγωγή των υπολογιστικών φύλλων από το Blender στη βάση δεδομένων, έως τη δημιουργία φορμών εισαγωγής δεδομένων (διεπαφών) για τη χειροκίνητη συμπλήρωσή τους (Εικ. 1, Ροή εργασιών του Έργου).

Τέλος, η διαδικτυακή διεπαφή που επιτρέπει την αναζήτηση ψηφιακών πόρων, η οποία είναι επί του παρόντος προσβάσιμη με διαπιστευτήρια, επιτρέπει μια συνεργατική διαδικασία υλοποίησης δεδομένων από τους εταίρους του Προγράμματος W.A.L.(L) και προετοιμάζεται για τη δημόσια χρήση των δυνατοτήτων υπολογισμού της βάσης δεδομένων.