International Journal on Advances in Systems and Measurements, vol 11 no 1 & 2, year 2018, http://www.iariajournals.org/systems_and_measurements/

Scientific Challenges in Archaeology

Is modern computer science ready for archaeology?

Lutz Schubert
IOMI, University of Ulm
Ulm, Germany
lutz.schubert@uni-ulm.de

Gill Hunt
Hunt Lancaster
Lancaster, UK
ghunt@huntlancaster.co.uk

Keith Jeffery
Keith G Jeffery Consultants
Faringdon, UK
keith.jeffery@keithgjefferyconsultants.co.uk

***Abstract*—Digital Humanities is receiving growing interest in academia. However, in most cases it is understood as teaching computer science in humanities, much rather than actually merging the subjects. In fact, most computer scientists regard humanities as a "trivial" challenge without "hard, scientific" problems. However, many areas in humanities pose hard challenges to which current computer science cannot provide sufficient solution approaches. In this paper, we examine the specific scientific problems in archaeology that pose new hard and interesting challenges to computer science.**

Keywords – advanced applications; archaeology; high performance computing; physics; simulation; network analysis; social networks; agent systems; theoretical computer science.

I. INTRODUCTION

In this paper we examine the challenges presented at the INFOCOMP conference in 2017 [1] and expand them with new insights and a concrete case study:

Dealing with one of the oldest subjects, archaeology is one of the youngest scientific fields, which developed really only in the late 19th century [2]. To many people it comes as a surprise how methodological and scientific such a discipline is that seemingly still works with shovels, pen and paper. In fact, archaeology is an early adopter to many new technologies, such as radiocarbon dating, remote sensing, LIDAR, photogrammetry etc. [3][4].

Yet what computational tasks may arise from a field that deals with interpretation much more than with hard facts? Why should statistical methods, databases and current analytical tools be insufficient? Let alone pose new problems to computer science?

To illustrate the problems faced by archaeology and the types of analysis typically performed, it is best to examine a typical, yet slightly more baffling case in archaeology (Section II). We will use this concrete case study to elaborate the analytical tasks of archaeology and examine the scientific foundations it builds up upon. We will analyse each of these methods in light of current computer science and its deficiencies, as well as potential solution approaches. We cannot propose a solution here: there is none as yet! All these challenges remain to be addressed by computer science and for this, more interdisciplinary research is needed. We will discuss this briefly in Section III.

Figure 1. Archaeologist mapping the layout of a trench. Source: Wikipedia.

II. CASE STUDY: THE MYSTICAL COIN

Figure 2. The excavation site at Okinawa. Source: Uruma City Educational Board.

In 2016, a group of 10 roundish, metal objects was found in Okinawa, Japan, 1m deep in the ground among a set of symmetrically shaped, roughly geometrically aligned set of stones [5]. To be precise, two types of shapes and geometries could be observed, but we will ignore this for simplicity sake. In Figure 2, you will probably recognise the ground plan of a building right away, and you may suppose that the metal plates are coins – but bear in mind that Japanese coinage in the past was quite distinctive (see Figure 8).

But we are already jumping ahead of ourselves: past? We are only talking about something in the ground! Building? Aligned stones! Why dig here in the first instance?

A. Geophysics

Many readers will have visited one archaeological site or the other and will probably have a mental image such as of the Acropolis (Figure 3), yet this is obviously basing on decades of reconstruction.

Figure 3. The Acropolis of Athens. Source: Savin, Wikimedia Commons.

Instead, archaeologists are happy if some indicator of the site is visible from the ground (see Figure 4). The most typical approach even nowadays still is to probe the ground at a "promising" location, viz. basing on historical records, sporadic finds or indeed indicators on the ground.

Figure 4. Archaeological site on Samothraki, GR. Try to find the buildings. Photo by the author.

Over recent years, other methods have evolved to detect buildings and structures that were otherwise covered up or undetectable. One such method, LiDAR recently just made the headlines, when the technique helped to discover a Maya "Megalopolis" in the jungles of Guatemala [6]. LiDAR bases on the capability of laser to penetrate the green foliage, thus displaying the elevation otherwise hidden by the trees.

Another, slightly less known technique exploits the fact that the density, magnetic or electric characteristics are different between materials. Thus, by applying a current, or injecting radar or sound waves into the ground, we can measure different "replies" depending on the material properties and the depth (distance, current, run-time) investigated. GPR (Ground Penetrating Radar) is one such technique, which is also exploited for oil exploration [7].

Such techniques have the major advantage that they are faster than a test dig and do not destroy the evidence in the ground. In other words, the archaeologists can safely explore the area to decide on where and how to excavate. On the negative side, the technique is expensive and still time consuming enough to not be executed on a more global scale, but is more localised.

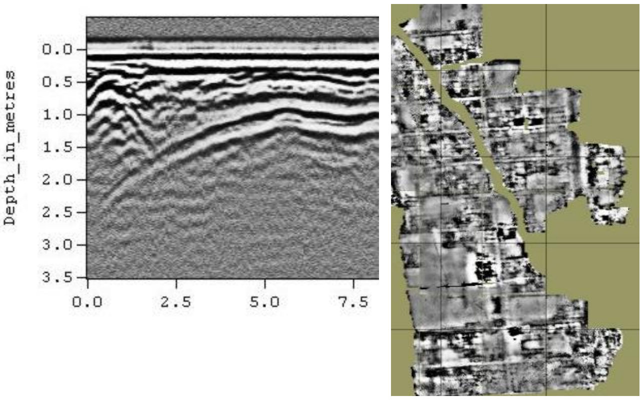

Figure 5. Ground Penetrating Radar (left) and map resulting from a resistance survey (right). Source: Wikipedia.

Scientific Challenge(s). The technique can only observe differences in material properties and allows (so far) little information about the material itself (see Figure 5, left). In general, the runtime evaluation is already computationally intensive [7]. What is more, however, there is still a wide gap to better assess the types of material in the ground and to assess the results in the first instance (see Figure 5, right).

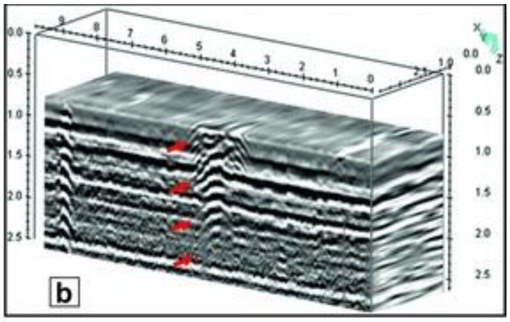

Figure 6. GPR survey of a Roman Stadium floor in Turkey. Source: [8].

Geology can provide additional information about the material constitution in the respective area. This provides a first insight into the expected conductive / reflective behaviour of the ground material (as is used in oil exploration in fact), and thus any deviation is highly indicative.

At the moment, computation of the results generates a pseudo 3d image that indicates the depth of material changes over the area covered (see Figure 6), but requires a considerable amount of interpretation.

In order to improve interpretation of the data, the different layouts and principle material characteristics need to be put into relationship with regional / context information. The granularity of the technique implies that only structures of larger granularity (typically buildings) can be identified. Nonetheless, as with any human (or in fact animal) artefacts, non-regarding all past and present standardisation efforts, individual structures will differ from each other – be it only for aesthetic preferences. In addition to this, local circumstances, as well as the differences in collapse and deterioration will lead to strong individual deviations.

This deviation needs to be taken into consideration when trying to interpret the structures under ground: changes in material properties may(!) represent structural elements which need to be compared against possible structures in the region – bearing the regional and individual differences in mind. Physical collapse as well as human influence in destruction (both see below) thereby equally play a role – we will come back to that later again.

Matching the 3d information needs a different way of representing and matching 3d structures than currently exists: the shape is not only rotation and scale invariant, but may also change considerably between instances – not only in terms of different collapse and destruction, but also in terms of individual builders' preferences.

Another way with related challenges consists in "space archaeology" [9], i.e., analysis of satellite images for potential features on the ground. Considering the amount of images and the details that could mark a site, this exceeds current image analysis techniques by far: both in terms of identifying the features, as well as the speed of execution.

B. *Stratigraphy?*

Though analysis of the GPR and satellite data may have given us an indicator for the site to dig, we still need somehow assign a date to the location:

Anyone who has ever dug a hole will have noticed that the constitution of the ground changes in sometimes clearly visible layers (or "strata"). Archaeologists (and in fact all geoscientists) can use this information to measure time – or more correctly, sequences of actions. To illustrate this, think about your own trashcan (see Figure 7): as you fill it over the days, it will accumulate garbage that reflects your past actions: at the lowest layer, the bottom of the trashcan, you can find whatever you threw away the very first day after the trashcan was emptied – let us say it's a set of different organic matter from a nice meal you prepared for friends. The next day you cleaned your house and threw away dirt from the vacuum cleaner, and assembled dirt from around the house – you may not even have eaten a lot, but only had some leftovers from the previous day, which generated little extra trash. The day after you threw away the old broken cup and some ugly ceramics that was gathering dust in the cupboard… and so on.

Figure 7. Stratigraphic layers in a trashcan. Source: Canterbury Archaeological Trust Ltd.

If you could cut your trashcan in half and look at it from the side, you could now see distinctive layers of different types of trash – each reflecting your daily activities and the garbage you produced. Of course, you will object that these layers are hardly that different over a normal week, which is one of the reasons that only major changes are actually observable (as changes) in the archaeological context. Since these layers decompose and compress, a meter can represent hundreds of years, or as Jared Diamond put it

"we can travel in time, it is just in a different direction than we expected."

Obviously, the relationship is not as simple as x meters down equals y years: strata build up differently in different regions (think forest floor vs. desert) and human (and animal) interference will change the layout, such as when digging a hole and filling it up with material from another location – we will look at this again later, in the context of "big data" below.

For now, let us assume that this gives us sufficient information to date the set of metal objects to the 15th century. Where was the computer science challenge in this?

Scientific Challenge(s). How material builds up and behaves over time is still mostly unknown. The information available so far bases mostly on experimental data and experience. In fact, chemical properties, composting, even compression of organic material is still too complex for modern computer simulations. Though we can model air and water flow, electro-chemical properties and deformation of metal etc. we still struggle with composite material that combines different properties and, which is worse, reacts with its environment and thereby changes its properties.

As computer scientists, we are used to *precise* computation – in fact, much algorithmic work is invested into increasing the precision without increasing the computation time. However, whilst this is of scientific interest, it is hardly of any practical value. Many computer scientists before have already noted that we could potentially achieve more if we would change the way we are asking questions – if we invest into "what" we want to achieve, rather than "how". The EC Cloud Computing Expert Group published one report to the end that software engineering needs to refocus the way we are thinking about problems [10]. Similarly, the whole area of "imprecise computing", as a rising IT field, is investigating whether we cannot achieve the same results with less precision [11].

What is more, the timescale of a typical simulation ranges in the order of milliseconds (such as for protein folding) to minutes (such as in airflow simulation). The longest time-span is probably in the order days with weather simulations, but bear in mind that this already belongs in the realm of imprecise and statistical computing.

Multi-level computing is still on the rise and problems such as the Virtual Physiological Human [12] cannot be solved before we cannot reduce the computational complexity. Organic composting, compression and composite material, though all highly important for building management, engineering etc. are so far only scientifically investigated by archaeology and build up on human expertise and experience.

C. *Multidisciplinary Statistical Big Data*

Until now, we identified the site itself and have a rough feeling for the time to which the metal plates belonged. As we will discuss in more details below, timing just by stratigraphy is insufficient as it differs too strongly between regions and circumstances.

One of the most frequent and most challenging tasks in archaeology consists in data analysis (and as we shall see, reasoning): in our example we have identified so far a location, the structure itself and a set of round metal objects. Analysing them goes far beyond the traditional material probing (though this can help dating). Primarily, any interpretation in archaeology is trying to link data with known facts. Let us concentrate on the building itself first:

We have known parameters: location, rough dimensions and ground plan, as well as depth (and therefore indicative age). With this, we can start looking for registries, historical records, architectural analysis etc. Other potential information involve the geological source of the stones, the shape and methods in which the stones were shaped (as observable by rough and microscopic surface analysis), the overall layout (as in cultural building traditions), any signatures, stamps, inscriptions and of course all other objects found on site. Archaeology is one of the most interdisciplinary scientific fields and combines information from a multitude of sources – we will come across this issue repeatedly in this text. For now, however, let us assume that historical records of the respective building exist – in this case, a Japanese castle from the 14th century.

Scientific Challenge(s). Cross-relating such information is a typical task and requires searching through different sources, ranging from geology, over climatology to anthropology, architecture, biology etc. Archaeology thereby has way less hard data at its disposal than other scientific fields – already the constant changes in radiocarbon-dating [13] or the constant reinterpretation of genetic data [14] is witness to the instability. In most cases, these re-examinations were in fact triggered by archaeologists, when the data generated by the so-called "hard scientific methods" was in conflict with the archaeological evidence or contradictory to reasoning.

The main challenges is however not so much the scope and interdisciplinarity, as one might expect, but the fact that most of the data is not concrete and that there can be only "interpretative" links, as the actual "facts" may change at any time. Just like in other scientific fields, once enough evidence is gathered that is in conflict with current thinking, the current facts need to be rethought. In the context of humanities, however, the actual subject (humans) are way more complex and facts considerably less stable.

In other words, the main challenge is to reason over different data and assess likelihood of certain events or situations. As we shall see, a simple stochastic reasoner is insufficient, as we must also consider stability of a belief and thus its impact on the overall conclusions, should it change.

Data mining hence needs to be extended with 3 aspects: interpretation (as any social data is subject to interpretation rather than a simple hard fact), reasoning (see below) and likelihood. In terms of likelihood, we need to distinguish between direct facts (i.e., the evidence gathered from whichever means – be that the stones or the metal plates on site, or the genome sequence from a human bone) and inferred facts and their relationship. For example, the *best* evidence suggests that the rock assemblage was indeed a Japanese castle (given the location and the rough age), but the documents may be false, the site may have been reused within a short timespan etc. In almost all archaeological cases no historical data exists in the first instance, thus making the interpretation even more subject to likelihood based on inference and assumptions.

The archaeological pictures (and in fact our understanding of ourselves as humans) gets constantly updated, as we gather more knowledge and gain more insight. In every step, all knowledge acquired needs to be considered and revised – potentially affecting *all* data.

Let us look back at the metal plates found in what we know presume to be a Japanese castle of the 14th century. By logical inference, we may first assume that the metal plates must be as old as the castle, but we have to be careful about this interpretation: the castle may have been built in the 14th century, but it was in use until the 15th century. Even then, it only slowly disintegrated and thus the coverage of dirt may indeed be from the 18th century and the metal plates of according age. In other words, the finds can only be as old, as the first layer on top of it – at the same time, it cannot be older than the first layer underneath it. Bear in mind that this applies to the age of last use of the object, not to the age of the object itself.

In this specific case, the evidence suggests indeed that the objects were or discarded some time in the 14th or 15th century. But what does that make them?

D. *Advanced Image Recognition*

As already noted in "Geophysics" above, one of the typical challenges consists in interpreting evidence –

visually. Though more elaborate analysis methods exist, they typically base on material constitution, chemical consistency etc. but only visual analysis helps for interpreting shapes and shapes can be interpreted through similarity: in rare cases are two objects identical – in particular if manufactured by hand. Yet we can infer producer (artist) and cultural traits from artistic decoration and shape [15] – for example, the “art nouveau” period left a distinctive mark in visual cues and it is pretty easy to distinguish individual artistic schools. Such classification is subject to a lot of research in image recognition, yet in our case here, we not only have to recognize styles, we also have to deal with deterioration and destruction.

In the case of our set of metal plates, the first association by inference was that these may have served as armor plates, as local coinage would have had a distinctive square hole (see Figure 8).

Figure 8. Examples of Japanese coinage from the 14th-17th century. Source: Wikipedia.

It was sheer luck that one of the archaeologists recognized a similarity with Roman coins based on experience gained in Egypt. What makes the challenge much harder is the fact that not only knowledge about the other context (here: Roman) is needed, but also that artefacts deteriorate. In case of the coin, Figure 9 depicts how the coin was found in Japan (left) and how it could have looked like if it would have been preserved better (right).

Figure 9. The Roman coin found on the Japanese site (left, Source: Uruma City Educational Board) and a well preserved version of (probably) the same coin (right, Source: www.wildwinds.com).

Scientific Challenge(s). As we have already noted multiple times, identification of finds, their classification and timing requires a big data comparison over multiple different sources. The case here clearly illustrates how ignorance of all information can lead to misinterpretation of the artefacts and thus to misinterpretation of whole region and its past. We will see later how the (correct) interpretation is actually not very helpful in this case.

Years are spent in analysing the finds from a site and even with the best archaeologists, mistakes in interpretation can happen. A computer aided interpretation process could greatly support this endeavour but would require that image recognition is expanded to consider deterioration processes. Much like we will discuss below in “Physics” and in “Geophysics” above, these processes can be modelled to a certain degree.

Differences in the manufacturing process (such as stamping coins) can already to a certain degree be analysed, yet the process is still very crude and does not really allow to compare coins of different origin. Analysis of microabrasions, experimental archaeology and simple experience so far help much more than the analysis capabilities of modern software. Again, combination of approaches are needed to expand the algorithmic quality.

E. 3d Recognition and Matching

More typically, though not the case in our example, archaeological finds are in fragments, such as broken pieces of ceramics, or the famous Lionman, which consists of more than 300 fragments of ivory and still is not complete – the smallest of the fragments is thereby only a few millimetres.

Next to the general layout of the finds, the actual material and shape of the fragments themselves provide indicators for their relationship. Consider the various forms of pottery that can be found in archaeology: shape, material and texture, respectively decoration are good indicators as to whether two sherds may have belonged to the same object. This also applies to (human) bones, larger sculptures etc.

Generally, parts are missing, scattered, or even archived in a completely different city / country due to different excavation processes, movement after excavation etc. Furthermore, due to the vast amount of similar fragments, identification of corresponding parts is close to impossible.

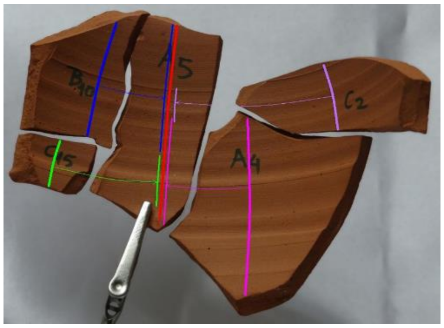

Figure 10. Matching sherds on basis of profile information [16].

Scientific Challenge(s). In itself, this is a considerable big data process where multiple factors need to be compared and cross-correlated to identify potential matches. These in turn will have to be matched in shape and against types of objects. Ideally, the fragments touch and thus have a common breakage area. Though this may sound like “just” a fitting task, one needs to consider (a) the amount and size of fits and (b) that natural processes change the breakage area by smoothing and reducing it etc., so that no perfect fit can be achieved. Such processes have to be taken into account when asserting whether two fragments match [17].

In most cases, no 3d models are available, let alone sufficient details to attempt a match in the first instance. As the number of available models increases, so does the complexity to match all the available finds – but even just within a single excavation, the effort is considerable.

As noted, though, in our example no such matching needs to take place.

F. Physics & Material Deterioration

We already mentioned in the context of Geophysics, above, that physical (there: biochemical) processes need to be understood when examining a site or the artefacts found therein. Of specific interest here are thereby the deterioration processes from exposure to the elements – this equally includes rusting of the metals objects, stratigraphic built-up, as well as structural collapse.

In the case of our castle, it is not only covered up by increasing layers of compost, loose dirt and rubble. The building probably also collapsed due to structural weakness and, what is more, human intervention: people always have and always will reuse buildings for new structures, as a material source, or just use the space for other purposes, thereby flattening the remainders. This poses extremely hard challenges on reconstructing the processes and simple rigid body physics simulations are insufficient to take all these factors into account.

It is however not only structural analysis that can benefit from physics simulation: humans and animals having been killed violently and / or moved after death will end up in certain positions and orientation. For example, skeletons in the Tollense valley have been moved by water slides and thus ended up in a collective heap [18]. Knowing the shape of the land, the flow of water and intensity of rainfall allows reconstructing where the bodies originated from, and (to a degree) their original positions. As the process is irreversible, this is not entirely possible – but the order of skeletons already indicates how they must have been flooded down the hill. Notably, the state of decomposition makes a major difference with this respect.

Related to this, marks in the skeleton give an indicator for strength and direction of a blow or projectile. Arrow-heads embedded in bones tell something about the position of the opponents relative to each other, but also about how the weapon was used and the force that the respective weapon can transmit. Human factors have to be taken into consideration, such as whether the force could be created by muscle strength (spear) or whether additional means would have been needed (bow). Given e.g., the Tollense layout, a reconstruction of the event can be attempted.

Scientific Challenge(s). As noted, rigid body physics simulations are good to get some feeling for how buildings can collapse, and thus good for educative purposes, but they are far from providing insight into the processes that actually took place on site. Just as with geophysics, what is needed is multi-level simulations in order to take all the different factors into account.

By nature, however, physics is a chaotic process with little influences leading to major changes in the outcome. Therefore, it is highly unlikely that we can recreate all the circumstances that led to a specific constellation as found in an archaeological site (or in fact any site). Archaeologists essentially invert physics and try to find logical explanations for which forces may have acted to result in this outcome.

Physics inversion is mathematically impossible, but stochastic relationships between forces and outcomes are possible [19][20][21]. The shape and layout of the elements in the heap allow reasoning over the possible original structures. By comparing these possibilities with existing, similar ones, we can make even reasonable assumptions about the factors that led to the final distribution.

Such methods essentially can only be used for verifying (respectively falsifying) certain assumptions, such as that the likelihood of a specific distribution of stones can be the result of natural phenomena, or that the distribution of food remainders has not been tampered with [19][22].

Notably, we need an assessment for the forces that may have acted on the site – one of the most difficult to measure and analyse being human intervention.

With all the information gathered so far we know now that the metal plate is indeed Roman, and given the deterioration, as well as the expected initial shape, we can assume that the coin is from the Roman Emperor Constantine. We also know that the coin must have been lost some time in the 14th or 15th century. Now let us apply our big data analysis again: Emperor Constantine lived from 272 to 337 AD – certainly way before our Japanese castle was built. His coins may have been in use past his reign, but since every emperor decreed new coins, it is unlikely that this lasted for very long. Obviously, the coins did not get in full disuse, yet at the 14th century, Roman coinage was not a currency anymore. Though the Roman empire still existed at this point, it had lost most of its influence and was quickly succumbing to the Arabian conquest (see Figure 11).

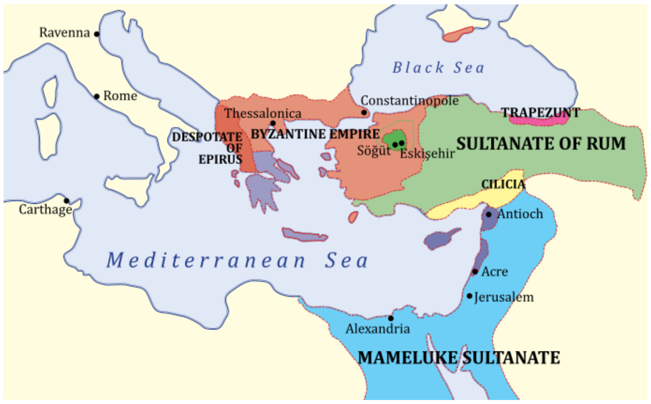

Figure 11. The Roman empire around the 14th century. Source: Wikipedia.

Which leads us to a simple question: how did a more than 1.000 year old coin end up in a place more than 10.000 km away from its origin (see Figure 12)? Japan was never part of the Roman empire and had its own coinage, but some human contact must obviously have taken place: we have hard scientific evidence in the undeniable coin's existence and location.

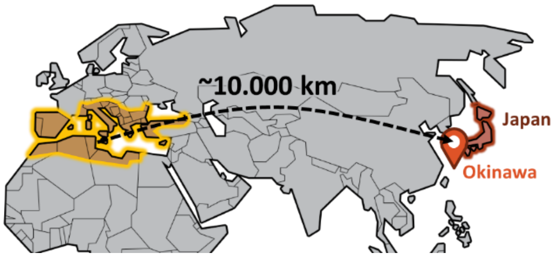

Figure 12. Map of Eurasia showing the main Roman empire and Japan.

What can we say about the human agents that must have been involved?

G. *Simulating Human Behaviour*

Archaeology is about humans: how they lived, what they have done, when and why. However, in an illiterate society, ways of thinking leave no traces and even in literate societies, written evidence should not be confused with facts [3]. The challenge for archaeology therefore consists in relating finds to potential behaviour, intentions and way of thinking. Some of this behaviour is obvious and straightforward: a ceramic pot indicates that (a) someone was there to leave the pot behind and (b) someone made the pot. However, was the pot used as a domestic item, was it an item of worship, was it just decorative, was it discarded right away? All this cannot be gathered from the pot alone.

As seen (data mining, above), a considerable amount of information has to be cross-linked. What is more, though, is that human behaviour, intentions and beliefs, capabilities and knowledge etc. stand at the middle of the explanation chain and form the basis for any conjecture. As indicated above, this can obviously take different levels:

Presence. Straight-forward, remains are "just" indicators for human presence and actions, such as that someone must have brought the find to the location, must have made it etc. Notably, not always is a find clearly of human origin, as e.g., is the case with some Palaeolithic "tools" [23]. This is the level of direct archaeological evidence.

Capabilities. At an intermediate level of complexity, human capabilities must be taken into consideration. This defines whether it was e.g., possible to reach a location, build a structure etc. How humans reached the American continent is one such unsolved question. At this level we talk about the assumptions that can be substantiated by archaeological evidence (existence of boats), but not fully proven.

Belief and Intention. At the most complex level we need to argue over belief and actions that are behind the evidence. It is a frequent cliché that archaeologists classify any evidence without clear functionality as "ritualistic". Indeed, it is difficult to assess the intention of an object that has no comparison in modern context. At this level, all "evidence" is pure conjecture and may change on basis of new theories.

Whereas knowledge at level 1 and partially at level 2 falls clearly into big data management, i.e., cross-checking facts, most of level 2 and in particular level 3 are conjecture and base on logical possibilities alone. Aspects such as movement of peoples require that the behaviour is simulated and the likelihood assessed on basis of this simulation.

Scientific Challenge(s). Even the best swarm simulation software cannot accurately model human behaviour beyond simple crowd movement. The typical approach consists in agent based simulations, which model multiple entities and their interactions on a simplified level [24]. There is a considerable amount of criticism of these models, as they must be incomplete and error prone – it is currently not possible to appropriately simulate how even just a small settlement would behave [25].

Human behaviour is complex and cannot be easily abstracted, so a major question relates to which human aspects have to be modelled in the first instance and how. Much can be learned from social network interactions, but care must be taken when applying modern contexts to ancient circumstances, as behaviour and mindset are in constant flux [3]. As indicated in the context of simulating physics, we can however argue about the likelihood of a specific behaviour (see below).

Statistical analyses can reduce the computational effort, even though they have a high error margin. They can help to eliminate *unlikely* situations, such as for the Roman "tourist" in Japan, which would necessitate the according means of travel, communication etc. In [26] the authors suggest an analysis based on throwing angles and strengths to assess the layout of shell middens. This is a highly simplified human behaviour model but already allows for some degree of feasibility assessment.

H. *Network Analysis*

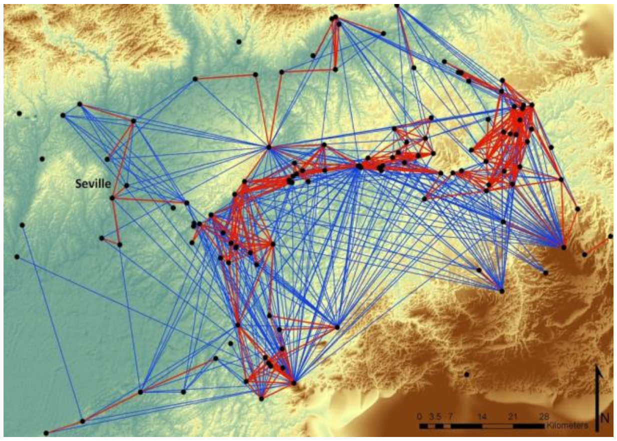

Figure 13. A social network diagram for Facebook. Source: Archaeological Networks.

A special and indirect form of human behaviour analysis consists in social network analysis (SNA). SNA gets a growing interest in archaeological analysis, as it is taken as an objective evaluation of the hard evidence. It belongs into the domain of big data analysis, but given its prominence and growing interest, it deserves its own subsection.

A network analysis assesses the statistical likelihood of two entities being related or connected. It is frequently used in social analysis to identify social relationships between people all over the world. Typical parameters are thereby simple knowledge of each other such as through Facebook (see Figure 13). Such a network can be analysed in different ways: if only social relationship is observed, clusters of high connectivity depict the relational intensity, i.e., "friends with common friends" – outliers or loose connections between clusters indicate isolated groups with little knowledge of each other. Such a relationship is interesting e.g., for target market analysis, to see which people may most influence each other's taste.

Frequently, other parameters are used for analysis, such as indeed shopping behaviour: people with similar taste are more likely to buy the same items. In other words: the more your shopping behaviour links to other people, the more likely their taste is comparable to yours. Whilst this works on average, it certainly fails in the specifics, as everyone looking at "personalised recommendations" knows.

In archaeology, other "similarity factors" are typically used, such as the similarity between art forms, genetic relatedness, or in our case, appearance of Roman coins. To be meaningful, the information typically needs to be projected onto a map, ideally also across time. Not surprisingly, we would expect that the occurrence of Roman coins diminishes over distance and time – and in fact, no other occurrence of Roman coins in Japan is known.

Scientific Challenges. SNA is far from a hard analysis tool and there are many problems with such an analysis. Unfortunately, results from SNA are frequently mistaken for facts, leading to potential misinterpretations.

Let us look at some of the base issues with network analysis – specifically in the context of archaeology:

1. Encoding "soft" parameters: similarity and thus potential relationship based on e.g., similarities of shape need to be encoded in a parametric form. Try a network analysis of all different forms of art nouveau and its appearance over the globe and you will notice that defining similarity is a hard task. By using too many parameters, relationship will diminish to nothing and by using too few (e.g., "coin") you will end up with too many relationships. Depending on who defines the parameters and how, completely different graphs can be produced – and hence completely different interpretations.

2. "Absence of evidence is not evidence of absence": just because no Roman coins have been recorded in Japan, this does not mean that they may not have been misinterpreted (or simply not found yet). Obviously, the network will immediately change, when more evidence is uncovered, but the reason for evidence being available is manifold and ranges from archaeological accuracy, destruction, investment to preservation conditions.

3. Different analysis methods lead to different results, but their interpretation is once again up to the human user: the obvious connection between the Roman empire and a Japanese castle does not imply that a Roman time traveller lived in a Japanese castle – we will need to turn back to that in "Reasoning" below.

Network analysis in the context of archaeology needs to be extended with the capability to assess the "stability" of the relationships. In other words: which impact would other analysis methods, new evidence, or change in parameters have on the network shape and in particular on its clustering.

Another, growing trend, is to assign weights to the relationships in the network graph – mostly in order to reflect the transportation or communication cost implied by distance (i.e., when project the graph onto a map). This provides an additional interpretation factor, which relates to the likelihood aspects mentioned above: the more difficult (costly) the connection, the less the impact and hence the weaker the relationship.

In context of the Roman coin, the cost for travelling 10.000 kms at Roman time would be tremendous. However, we thereby neglect a few factors, which we shall turn to next:

I. Reasoning

With the data we have from the Japanese site so far we may infer multiple interpretations, such as that a Roman travelled to Japan to live there and leave the coin as a heritage to his children; that a 2nd century Japanese travelled to Rome and returned with a coin; that a 14th century Japanese travelled through time and space and so on. Though some of these explanations will immediately seem ridiculous, we must note that none of them contradicts the data or interpretation we have so far.

The key point here being that some of the interpretations stand to *reason* that they cannot be correct – such as for the time traveller. In fact, if we consider the data carefully enough, we will even notice that the Roman travelling 10.000 kilometers is highly unlikely given the situation and travelling options at the time. This form of restricting the options (and thus weighing the relationship) is a form of big data analysis, but implies logical constraints and derivations – for example, by applying the average travelling speed at given times as a factor.

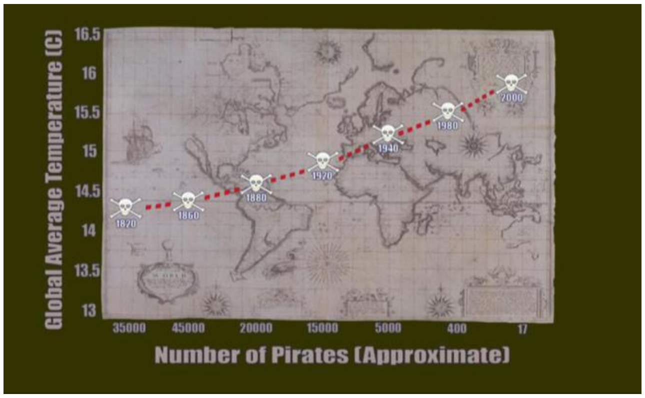

Figure 14. Relationship between numbers of pirates and global average temperature. Source: Forbes.

Modern analysis methods are bad at such behaviour and generally just generate statistical relationships. A classic example of such a "maladjusted" interpretation is the relationship between pirates and global warming: because a correlation between numbers of pirates and global temperature can be observed (see Figure 14) [28]. Since there is no logical check involved, this assumption is fair –

we may even be inclined to take the inversion of the false cause-and-effect chain and thus presume that we need more pirates to counteract global warming. It is only because it is so obviously wrong, that this assumption is not taken up – what, however, if the logical mistake is not so obvious?

Scientific Challenge(s). Reasoning in AI methods is still a new field, but an increasing amount of voices raise concerns about potential misinterpretations of data. Note though that reasoning itself is a type of big data challenge as even though the logical operators are constraint, the number of potential logical assertions is infinite.

Such attempts relate strongly to the domain of semantic reasoning and are still under development, but should not be left aside when interpreting data.

What does all this leave us for interpreting the data at the Japanese site? Obviously, we cannot get a definite answer, but several explanations reach a certain plausibility: (1) there are more Roman coins in Japan that just have not been discovered, (2) the most likely candidate for a transfer route is the silk road, which equally existed in Roman and medieval times. We can only make assumptions about why the coin ended up here and obviously, anything from a collector, over heirloom to an exotic gift may be valid.

Obviously, we can come up with many more alternative routes, but we hopefully managed to reproduce the archaeologist's way of thinking in an algorithmic fashion, as well as showing which challenges this still poses for computer science.

For the sake of completeness, we can identify many more such challenges, such as:

J. *Simulating Climate*

Climate is constantly changing – not only due to human interference, but also due to the earth's rotation and movement, leading to glacial and hot periods. The implication of such weather changes is obvious and can already be observed today: different plants grow in different climatic zones, animals (and certainly humans) move to different areas, clothing changes etc. In times before Air Conditioning, this hit doubly strong and will have caused (and prevented) massive movement and settlement patterns, following game or reacting to environmental pressure.

Climate completely changes the face of the earth, from rising (and sinking) sea levels to landscapes covered in ice sheets or turned into steppes. These changes leave their marks and are sometimes directly measurable, such as in tree growth (dendrochronology) or remains of marine life in the dessert, respectively vice versa [3] [4].

In the archaeological context climate is only of interest insofar as it influences humans [27]. As such, it is only a contributing factor to Simulating Human Behaviour (see Section II.G) and can serve equally as an explanation, as well as an obstacle. For example, the movement of Homo Sapiens to the American continent is frequently explained by the possibility of a connection between North America and Siberia (the Bering land bridge) [29]. This land bridge could have existed due to a massive amount of water being locked in ice, thus causing the sea-level to sink considerably. Similarly, the movement of hominins into central Europe from Africa may have been made possible by fluctuations (inter-pluvial arid periods) in the temperature of the Sahara [30].

Climate conditions apparently play a role in any discussion about behaviour influenced by weather, such as clothing, foodstuff etc. Therefore, modelling the weather and in particular the climatic changes over history is a relevant aspect of the argumentation chain related to Simulating Human Behaviour (see Section II.G).

Approaches. It is well-known that weather simulation belongs to the most difficult tasks in advanced applications [31]. While meteorological simulations try to accurately predict local, minute changes in the weather, climate models can be more coarse-grained, identifying patterns of general weather trends over longer periods of time. However, already the overall climatic changes in the glacial and interglacial periods are difficult to predict and not all factors are known. Such models base more on observed factors, such as glacial movements and encapsulated CO^2, than on calculations [4].

Nonetheless, different models are under development [32] and particularly try to provide more local and fine-grained climatic conditions, so as to assess the size and distribution of ice sheets, but also just to predict shorelines, climatic zones etc. Such models can be validated partially against archaeobotanical finds, i.e., seeds that have been preserved under anaerobic conditions.

K. *3d images*

3d scanning is a growing field of interest in general, but also more and more archaeologists make use of photogrammetry to document the excavation [33]. There is a high risk that this is considered sufficient documentation, though it cannot replace profile drawings or good maps, but we shall not follow this discussion in this paper.

Generating 3d models from pictures taken in the open field is still time consuming and error prone, where missing pictures can only be identified after generation of the point cloud, which can take days in itself. Since the excavation will have progressed by then, this can lead to considerable problems. Better methods are needed to assess quality and potential gaps right at the time of taking the pictures, and the process in general needs to become more flexible – both require new algorithmic approaches that are highly related to performance optimisation in general.

One should also not ignore the fact that 3d scanning generates massive amount of data (i.e., the 3d points) that so far cannot be easily processed. Identifying an object in 3d space, i.e., which points belong to each other to form an artefact of its own, is still basically impossible. Similar challenges exist in 2d image analysis, where major progress has been made. So far most approaches simply generate a mesh of the whole scan, thus not allowing to (re)move individual objects, let alone perform an analysis on this level.

Since the advent of LIDAR scanning [34], processing of 3d images becomes an important factor for detecting hidden and obscured structures, very similar to identifying hidden structures in geophysical data.

Approaches. So far, most approaches rely on methods from 2d image processing, such as similarity of colour, identification of key features and of their relationship etc., but application in 3d is still very limited – not alone because the size of data is considerably larger (at least from n^2 to n^3).

Google and Microsoft already try to incorporate scans and 3d data from multiple (social) sources, but the sheer amount and computational complexity is still an unsolved challenge. Ideally, however, multiple sources are integrated in scanning, but notably, these will all have to be calibrated individually and the data then has to be cross-correlated first.

Some attempts also try to make use of additional data, such as arising from the accelerometer to pre-assess the quality and usability of the images, but there is no general good solution as yet and the amount of data will only increase.

III. CONCLUSIONS

The list of issues presented in this paper is far from exhaustive but already demonstrates the shortcoming of current computer science methodologies with respect to the needs of archaeology. Specifically, by addressing these challenges and incorporating knowledge from archaeology, the following improvements could be achieved:

- improved geological modelling: archaeology has knowledge about more short-term processes, such as soil deposition and collapse that can be exploited for engineering, city planning etc.;
- better human and agent models: anthropology and archaeology have information about human movement that is not reflected in simulation, thus leading to unrealistic movement and agency models;
- prospecting can benefit from prediction models and material knowledge gained from excavations;
- data mining and big data do not address complexities raised by such interdisciplinary fields as archaeology, which develops such methods for 100 years now;
- statistical analysis is an important field in archaeology and needs to be applied differently for network analysis, clustering etc. The feedback is rarely incorporated (see e.g., [35]);
- structure from motion is constantly being improved through landscape archaeology and field surveys [33] – new more robust methods and better object recognition are still being researched;
- most simulations model time forward from a given situation– in archaeology, time needs to be modelled backwards, i.e., leading from effect to cause, which in turn improves simulation performance and analysis capabilities [26];
- dealing with incomplete data by adding assumption models: archaeology is using methods for this on a daily basis, yet big data still struggles with it;
- both fields need better methods to capture the probability and likelihood of complex data to be correct and to identify logical and improbable errors;
- reasoning needs to improve beyond stochastic data mapping and in particular needs to include the probability that two actions are related. Artificial Intelligence concepts from the 90ies already approach such issues on a limited scale.

Not only can computer science improve archaeology further, but also knowledge from archaeology can help advance computer science capabilities in particular for application in any human-centric simulation or modelling.

We hope that this paper has shown that there remain many challenging tasks for IT in archaeology and that computer science still has many things to learn from the approaches in archaeology. The authors directly contribute to such a collaboration via a dedicated working group in the CAA International (Computer Applications and Quantitative Methods in Archaeology).

REFERENCES

[1] L. Schubert, K. Jeffery, and G. Hunt, “Challenges for Advanced Applications in Archaeology,” In Proceedings of The Seventh International Conference on Advanced Communications and Computation, June 25-29, 2017, Venice, Italy, C.-P. Rückemann, I. Flood, I. Schwerdtfeger, C. Simmendinger, G. Beckett, (Eds.), International Academy, Research, and Industry Association, 2017, pp. 29-34, ISSN: 2308-3484, ISBN: 978-1-61208-567-8, URL: http://www.thinkmind.org/index.php?view=article&articleid=infocomp_2017_3_30_60034 [retrieved: March 2018].

[2] K. Greene and T. Moore, “Archaeology: An Introduction,” Routledge, 2010. ISBN: 978-0415496391

[3] M. Sutton, “Archaeology: The Science of the Human Past,” Routledge, July 2015. ISBN: 978-0205572373

[4] M.J. Aitken, “Science-Based Dating in Archaeology,” Routledge, February 2014. ISBN: 978-0582493094

[5] E. Jozuka, “Ancient Roman coins found in ruined Japanese castle,” CNN, 2017 [Online]. Available from: http://edition.cnn.com/2016/09/27/luxury/ancient-roman-coins-japan/ [retrieved: May 2018]

[6] T. Clynes, “Ground Penetrating Radar Theory and Applications,” National Geographic News, 2018 [Online]. Available at: https://news.nationalgeographic.com/2018/02/maya-laser-lidar-guatemala-pacunam/ [retrieved: May 2018]

[7] H.M. Jol (Ed.), “Ground Penetrating Radar Theory and Applications,” Elsevier, 2008. ISBN: 978-0-444-53348-7

[8] V. Karabacak, Ö. Yönlü, E. Dökü, N. Kıyak, E. Altunel, S. Özüdoğru, C.C. Yalciner, and H.S. Akyuz, “Analyses of Seismic Deformation at the Kibyra Roman Stadium, Southwest Turkey,” Geoarchaeology. 28, 2013, pp. 531-543. https://doi.org/10.1002/gea.21456

[9] R. Lasaponara and N. Masini, “Satellite Remote Sensing: A New Tool for Archaeology,” Springer, 2012. ISBN: 978-9400796157

[10] L. Schubert and K. Jeffery, "New Software Engineering Requirements in Clouds and Large-Scale Systems," in IEEE Cloud Computing, vol. 2, no. 1, pp. 48-58, Jan.-Feb. 2015. https://doi.org/10.1109/MCC.2015.13

[11] J. M. Cadenas and M. C. Garrido, “Intelligent Data Analysis, Soft Computing and Imperfect Data,” in Soft Computing Based Optimization and Decision Models, D. A. Pelta and C. Cruz Corona, Eds. Cham: Springer International Publishing, 2018, pp. 25–43. ISBN: 978-3319642857

[12] M. Viceconti, “The Virtual Physiological Human: Computer Simulation for Integrative Biomedicine: Papers of a Theme

Issue," Philosophical transactions of the Royal Society of London / A: Philosophical transactions of the Royal Society. Royal Society, 2010. https://doi.org/10.1098/rsta.2010.0096

[13] Y. Asscher, S. Weiner, and E. Boaretto. "A new method for extracting the insoluble occluded carbon in archaeological and modern phytoliths: Detection of 14C depleted carbon fraction and implications for radiocarbon dating," Journal of Archaeological Science 78, 2017, pp. 57-65. https://doi.org/10.1016/j.jas.2016.11.005

[14] R. Eisenhofer, A. Anderson, K. Dobney, A. Cooper, and L.S. Weyrich, "Ancient Microbial DNA in Dental Calculus: A New method for Studying Rapid Human Migration Events," The Journal of Island and Coastal Archaeology, 2017, pp. 1-14. https://doi.org/10.1080/15564894.2017.1382620

[15] A. M. W. Hunt, "The Oxford Handbook of Archaeological Ceramic Analysis," Oxford University Press, 2017. ISBN: 978-0199681532

[16] M.I. Stamatopoulos and C.-N. Anagnostopoulos, "3D digital reassembling of archaeological ceramic pottery fragments based on their thickness profile," ResearchGate 2016. [Online] Available at: https://www.researchgate.net/publication/301873740_3D_digital_reassembling_of_archaeological_ceramic_pottery_fragments_based_on_their_thickness_profile [retrieved: May 2018]

[17] F. Stanco, S. Battiato, and G. Gallo, "Digital Imaging for Cultural Heritage Preservation: Analysis, Restoration, and Reconstruction of Ancient Artworks," CRC Press, 2011. ISBN: 978-1439821732

[18] A. Curry, "Slaughter at the bridge: Uncovering a colossal Bronze Age battle," Science Magazine 2016. [Online] Available at: http://www.sciencemag.org/news/2016/03/slaughter-bridge-uncovering-colossal-bronze-age-battle [retrieved: May 2018]

[19] L. Schubert and K. Jeffery, "Interpreting Shell Middens through Physics Simulation," In Proceedings of the Computer Applications & Quantitative Methods in Archaeology (CAA) International Conference, 2017. In print.

[20] L. Schubert, S. Metz and K. Jeffey, "Inverting Physics – a future for the past," In Proceedings of the ICNAAM Conference, 2017. In print.

[21] L. Schubert, R. Warden, and T. Guderjan, "Site Reconstruction from Partial Information," In Proceedings of the Computer Applications & Quantitative Methods in Archaeology (CAA) International Conference, 2018. In print.

[22] L. Schubert, A. Predoi, and K. Jeffery, "Analysing the archaeological context: Reconstructing stratigraphic layers," In AIP Conference Proceedings, 2017. https://doi.org/10.1063/1.4992196

[23] K. Wong, "Wild Monkeys' Stone 'Tools' Force a Rethink of Human Uniqueness," Scientific American, 2016 [Online]. Available at: https://www.scientificamerican.com/article/wild-monkeys-stone-tools-force-a-rethink-of-human-uniqueness/ [retrieved: May 2018]

[24] G. Wurzer, K. Kowarik, and H. Reschreiter (eds.), "Agent-based Modeling and Simulation in Archaeology," Springer 2015. ISBN: 978-3319342825

[25] D. Thalmann and S.R. Musse, "Crowd Simulation," Springer 2012. ISBN: 978-1447144496

[26] L. Schubert and K. Jeffery, "Modelling Physics in Shell Middens," Proceedings of the CAA 2017, in print

[27] M. Tallavaara, M. Luoto, N. Korhonen, H. Järvinen, and H. Seppä, "Human population dynamics in Europe over the Last Glacial Maximum," PNAS, vol. 112, no. 27, pp. 8232-8237, July 2015. https://doi.org/10.1073/pnas.1503784112

[28] E. Andersen, "True Fact: The Lack of Pirates Is Causing Global Warming," Forbes Online, 2012. Available at: https://www.forbes.com/sites/erikaandersen/2012/03/23/true-fact-the-lack-of-pirates-is-causing-global-warming/ [Retrieved: May 2018]

[29] S. Armstrong Elias, "First Americans Lived on Bering Land Bridge for Thousands of Years," Scientific American 2016 [Online]. Available at: https://www.scientificamerican.com/article/first-americans-lived-on-bering-land-bridge-for-thousands-of-years/ [retrieved: May 2018]

[30] W.J. Burroughs, "Climate Change in Prehistory: the end of the reign of chaos," Cambridge University Press, 2007. ISBN: 978-0521070102

[31] U. Gähde, S. Hartmann, and J. Henning (Eds.), "Models, Simulations, and the Reduction of Complexity," de Gruyter, 2013. ISBN: 978-3110313604

[32] J.d'A. Guedes, S.A. Crabtree, R.K. Bocinsky, and T.A. Kohler, "21st-Century Approaches to Ancient Problems: Climate and Society," PNAS, vol. 113, no.51, pp. 14483-14491. https://doi.org/10.1073/pnas.1616188113

[33] M. Llobera, "Archaeological Visualization: Towards an Archaeological Information Science (AISc)," J Archaeol Method Theory, vol. 18, no. 3, pp. 193-223, September 2011. https://doi.org/10.1007/s10816-010-9098-4

[34] A.J. Clark, "Seeing Beneath the Soil, Prospecting Methods in Archaeology," Batsford London, 1990. ISBN: 978-0415214407

[35] A. Collar, F. Coward, T. Brughmans, and B.J. Mills, "Networks in Archaeology: Phenomena, Abstraction, Representation," J Archaeol Method Theory, vol. 22, no. 1, pp. 1–32. January 2015. https://doi.org/10.1007/s10816-014-9235-6