

---

# ”How similar, how different? The problems of quantification of similarity in archaeological research- the example of ancient pottery ”

Agnieszka Kaliszewska\*<sup>1</sup>, Rafał Bieńkowski<sup>1</sup>, and Monika Syga<sup>2</sup>

<sup>1</sup>Systems Research Institute [Warsaw] – Poland

<sup>2</sup>Faculty of Mathematics and Information Science [Warszawa] – Poland

## Abstract

This contribution aims at presenting issues arising with recognition and description of similarity in the archaeological research. This topic might seem trivial at the beginning but it can be the source of numerous difficulties in the research process. This is in major part related to the fact that each person, even a theoretically objective researcher, perceives reality in a slightly different way. These issues have been partly addressed in archaeology and geological sciences with tools such as the Munsell colour system or the Moh scale. The necessity to use objective and unequivocal parameters and terminology in the description of the subjects of the study is one of the big challenges in archaeological research.

For the purpose of this study, we use the example of archaeological pottery. Pottery is the most common and numerous find during archaeological excavations, and hence constitutes a good marker for chronological periods, cultural groups, sometimes even ethnicity of the people who created or used the pots. The study of ancient pottery requires a careful study of the similarities and differences between the pots in one assemblage as well their relation to the pottery found elsewhere. Oftentimes the definition of whether two objects are different or similar and to what degree poses a challenge. The differences can often be very subtle and the recognition of the right parameters that distinguish the pottery fragments requires highly specialised knowledge.

The research on archaeology is heavily based on typologies, created mostly on the base of the shape and sometimes decoration of the vessels. Typologies constitute an efficient tool to deal with the variability of ancient ceramics. They can easily be referenced and we can assign a ”new” vessel to one of the known groups. However, researchers often encounter the situation when they want to quantify how much two vessels are similar. This, in turn, requires a definition of similarity as well as a way to measure it. The easiest definition of similarity is that it is a set of characteristics shared by the objects. Among these characteristics are size, weight, capacity, or the material of which the object is made. A commonly used characteristic is the shape of the object. Although it is very natural to describe an object by its shape, it creates a series of problems when trying to compare two objects.

In archaeological research, there are several standards of conveying as many information as possible about pottery and pottery fragments. These are used for the purpose of preparing data for analysis and to enable comparison between fragments. Among these are: archaeological drawings or a standardised form for the description of specific aspects of pottery. With

---

\*Speaker

the introduction of digital photography and microscope photography, this kind of documentation is also often added. The need to create highly precise and complete documentation is a part of the whole archaeological research, not only ceramology.

The process, described above is true for an "analogue" form of research. Many problems arise when the process is transferred to a digital environment. In order to perform any kind of analysis digital data must be complete, consistent and compatible. If we are to look for similarity/differences in the investigated objects, similarity must be well defined, and, contrary to the analogue process there is no place for judgement calls made on the basis of experience or knowledge. The second issue which arises when transferring the analysis to the digital environment is the question of compatibility of data. If the documentation is in the form of descriptions, there is no guarantee that characteristics perceived as similar are described in a similar way. Such inconsistencies may include terms such as ovoid, oval, cylindrical, rounded-elongated. To further make things more complicated these terms are also used together with quantifiers such as slightly, a little, sharply, very etc. All these expressions, although very natural in written language, do not translate in the computer-aided analysis. Many of these can be tackled through tagging of data, dictionaries or ontologies, but the question of "How similar are the objects?" will be still dependent on the adopted definition and measure of similarity. Moreover, such definition and measure will differ from study to study, as they depend on the research questions that need to be answered through the analysis.

In the last part, the authors would like to present an example of an approach to determine similarity and automatically create typologies in archaeological ceramics through the application of mathematical tools and examine how it relates to the discussion presented in the contribution.

**Keywords:** similarity, typologies, archaeology, data analysis