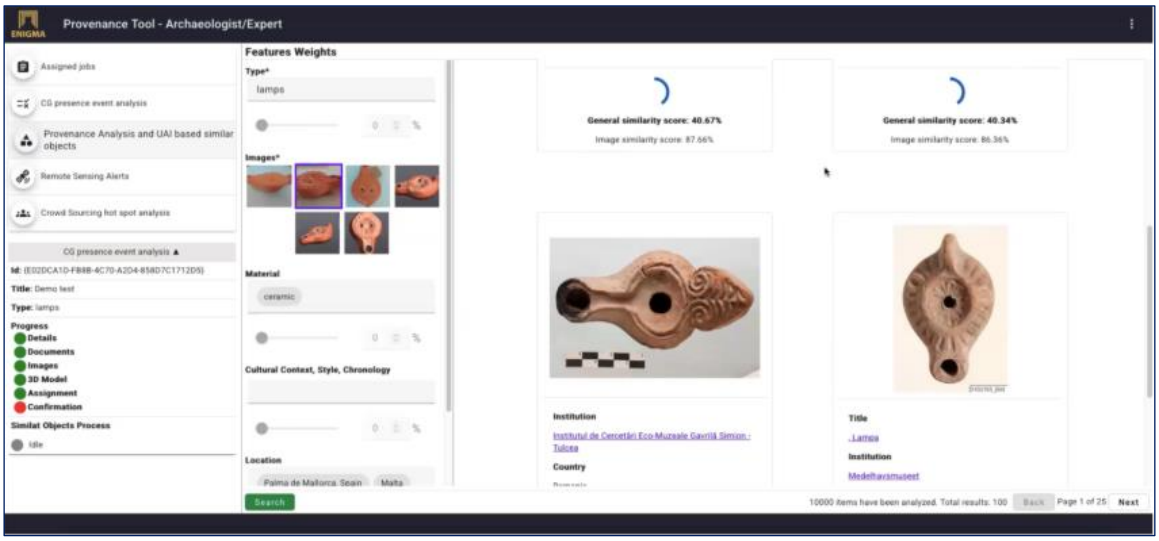
Best Practice: SIMILARITY SCORE

DESCRIPTION

ENIGMA includes a similarity score as part of the Unique Authentication Identifier (UAI) and Provenance Analysis tools. Using AI, the similarity compares input of a suspicious and held object with artefacts from external databases, and provides a list of potential matches. The similarity can help LEAs with verifying the legal status, or may help Cultural Heritage Experts in determining authenticity and provide supplementary provenance data. Thus, the similarity score helps to identify the authenticity, provenance, origin, and legal status of the held object.

IDEAL SCENARIO

For example, a LEA officer uses ENIGMA to input data of a seized artefact. They provide a full set of metadata, including the object's category, material, and a detailed description. In addition, they upload a series of images. The AI processes a similarity score based on the available data and generates a report showing a few highly similar objects from the main database. This allows the officer to quickly identify a potential match and obtain crucial provenance information.



TIPS

Do

- Use the vocabularies from ENIGMA for better search results
- Include a clear object type, as this increases a successful similarity
- Upload fine quality images
- Be patient, the similarity score may take a while

Don't

- Don't solely rely on similarity, and compare matches to understand the context of the object
- Don't rely on only one search algorithm, use the three search options for better similarity scores
- Don't assume a low score means the object is authentic or unique

***Figure**: UAI and similarity results for the Expert provenance search*

Funded by the European Union

MAIN IMAGE

Example of clustered artifacts generated as part of the similarity score analysis. Each three cluster groups in the image shows five example amulets grouped using clustering algorithm with deep learning features. Images contain also, for example, material, origin and period information (created with the description texts and GPT-5 model). This visualization demonstrates how clustering can reveal patterns and similarities supporting provenance and authenticity assessment.

FURTHER RESOURCES:

Vähämäki, BSc Santeri, Luca Zelioli, and Henri Nyberg. "leveraging text extraction and language models for coin image classification." (2025).

Troiano, Maurizio, et al. "A comparative analysis of machine learning algorithms for identifying cultural and technological groups in archaeological datasets through clustering analysis of homogeneous data." *Electronics* 13.14 (2024): 2752.

Pereira, Javier, et al. "Clustering and semantics preservation in cultural heritage information spaces." *RIAO*. 2010.