

縄文土器の説明文生成 AI の評価データ作成システム

中田 豊久¹、近山 英輔¹、宮尾 亨²、藤田 晴啓¹

大規模言語モデル(LLM)の特徴の 1 つは、言葉を数のように取り扱うことである。例えば、「赤」+「白」=「ピンク」というような四則演算を行うことができる。この特徴により教師データには明示的に含まれていない文章であっても AI が自動的に作成してくれることになる。したがって縄文土器の説明文生成 AI の場合には、学習データには含まれていない新しい未知の土器に対しても一定程度の説明をしてくれることが期待される。そのため本稿で開発するシステムは、生成 AI の評価を行うための縄文土器データを作成することを目的としている。具体的には、任意に選択された 2 つの土器から片方は下部、もう一方は上部を取り出して結合して生成 AI の入力データを作成する。

図 1 に開発したシステムを示す。このシステムはブラウザをユーザインターフェースとして使用し、プログラムは JavaScript で作成されている。まず、3D モデリングソフトである Blender を使用して remesh, decimate 機能によってポリゴン数を約 5,000 から 8,000 まで削減した 3D データを準備する。システムはこれらの 3D モデルを読み込み、自動的に上部と下部に二分割する。二分割の手法は、各ポリゴンの法線ベクトルの縦軸成分の分散を使用して行われる。二分割されたそれぞれの部分の分散の和が最も大きくなるポイントで二分割が行われる。さらに 2 つの土器を選択すると、それらのスケールと位置を調整して画面の下に統合した土器が表示される。ダウンロードボタンを押すと、現在表示されている統合された土器の点群データがローカル PC にダウンロードされる。この開発されたシステムは、<https://www.pencode.com/suika/joumon.html> で公開されている。



図 1: 画面上部で 2 つの土器を選択すると、下に結合された土器が表示される

¹ 新潟国際情報大学経営情報学部

² 新潟県立歴史博物館