

Oct 03, 2024

Explainability of Foundation Models

DOI

dx.doi.org/10.17504/protocols.io.4r3l2qw3pl1y/v1

Paola Di Maio^{1,2}

¹CSKRNS; ²Igdore



Paola Di Maio

CSKRNS

OPEN  ACCESS



DOI: **dx.doi.org/10.17504/protocols.io.4r3l2qw3pl1y/v1**

Protocol Citation: Paola Di Maio 2024. Explainability of Foundation Models. **protocols.io**

<https://dx.doi.org/10.17504/protocols.io.4r3l2qw3pl1y/v1>

License: This is an open access protocol distributed under the terms of the **[Creative Commons Attribution License](#)**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Protocol status: Working

We use this protocol and it's working

Created: October 03, 2024

Last Modified: October 03, 2024

Protocol Integer ID: 109143

Abstract

A metamodeling approach to support the development of a specifications for explainability in foundation models, adopting a combination of visualization, natural language representation.



- 1 Establish Ground Truth
- 2 Define Alignment
- 3 Define concept, key terms
- 4 Select Knowledge Representation
- 5 Visualise KR
- 6 Black box testing