MARTI Metadata Requirements Outline (Version 1.0)

Defining Metadata Standards for the MARTI Framework
Document Version Information:
Last edited: Carrie Bickner. Nov 13, 2024 12:00 PM EST

1. Introduction

This document outlines the metadata requirements for the MARTI (Metadata for AI Responsibility, Transparency, and Integrity) framework. These requirements establish a structured approach to documenting both simple and complex AI-generated objects, including generative AI outputs, computer programs, and large language models (LLMs). By setting precise standards for metadata across different domains and applications, MARTI aims to maintain transparency, integrity, and accountability in AI systems.

2. Criteria for Metadata

The MARTI metadata framework follows two main criteria:
1. Task-based requirements: Metadata elements that specify the intended purpose or usage context of an AI object, adapting to different applications such as marketing, creative arts, scientific research, and more.

2. Suite or environment requirements: Metadata elements that describe the environment or tools used to create the AI object, addressing specifics like development settings, training data changes, and hardware configurations.

3. Task-Based Metadata

Task-based metadata requirements in MARTI adapt to the distinct needs of each application area. These requirements include elements that help users understand the purpose, scope, and performance of each AI output:

- Marketing: Metadata that ensures authenticity, fidelity, and ethical compliance in AI-driven campaigns.
- Creative Arts: Elements for attribution, originality metrics, and the creative intent behind each output.
- Scientific and Product Innovation: Detailed records of versioning, modifications, and test results to track scientific rigor and product evolution.
  This approach ensures that metadata supports each domain's unique standards while aligning with MARTI's core values of transparency and accountability.

4. Suite or Environment Metadata

MARTI's suite or environment metadata requirements capture essential context for AI objects by detailing the creation environment. Key elements include:

- Development and Training Data: Records of dataset versions, data diversity, and changes across iterations.
- Hardware and Software Specifications: Details on hardware (e.g., GPUs, memory) and runtime environments, as well as software versions.
- API Changes and Compatibility: Metadata documenting any API or compatibility changes between versions to inform developers and users.
  These elements ensure users understand the technical context of an AI object, which is crucial for interpreting its capabilities and limitations.

5. Preservation and Provenance

Preservation is central to MARTI's metadata requirements. By applying museum and library practices, MARTI captures:

- Provenance: Comprehensive lineage documentation that establishes the origin and ownership of each AI output.
- Chain-of-Custody: A sequential record of custody and modifications throughout an object's lifecycle.
- Archival Standards: Metadata practices aligned with established standards to ensure long-term reliability and accessibility of AI-generated objects.

6. Instance and Variation Delineation

MARTI metadata requirements support clear delineation between instances and variations of AI objects. Key elements include:

- Instance Identification: Unique identifiers that distinguish each version, ensuring accurate traceability and comparison.
- Variation Metadata: Links to adaptations or special-use versions, specifying derivative objects and highlighting their relationships to the original.
  This delineation allows users to interpret each AI output accurately, understanding both its unique features and its connection to related versions.

7. Quality Standards for Metadata

High-quality metadata is essential to MARTI's goals. Key standards for metadata quality include:

- Accuracy and Completeness: All metadata elements must be precise and complete, allowing comprehensive documentation.
- Consistency and Standardization: Consistent formats, terminology, and structure across metadata records ensure easy interpretation.
- Ethical Integrity: Metadata should reflect ethical considerations and avoid misrepresentation, especially in contexts like marketing or research.

8. Future Considerations and Versioning

As AI technology evolves, MARTI's metadata requirements must adapt to new standards and developments. This document will undergo periodic reviews to incorporate emerging needs and technologies, ensuring that MARTI's framework remains relevant and robust. Versioning will reflect each major update, with records of changes to provide transparency and traceability.

```
MARTI Metadata Record: Metadata Requirements Outline (Version 1.0,
Edited)

1. Record ID: MARTI_REQUIREMENTS_1_0_EDITED

2. Title: MARTI Metadata Requirements Outline

3. Version: 1.0 (Edited)

   - Primary Agent: Carrie Bickner
   - Conceptual Architect: Carrie Bickner
   - Original Draft: ChatGPT (Orla)
   - Edited by: Carrie Bickner

5. Date of Version: November 13, 2024

6. Purpose: To define the metadata requirements within the MARTI
framework, establishing adaptable standards for both simple and
complex AI objects and systems, including large language models
(LLMs) and static AI outputs.

7. Core Changes:
   - Minor edits to refine language and clarify the adaptability of
metadata requirements across different domains.

8. Chain of Custody:
   - Version 1.0: Created by ChatGPT (Orla) on November 13, 2024.
   - Version 1.0 (Edited): Edited by Carrie Bickner on November 13,
2024.
```

9. Provenance: This record captures the initial metadata requirements outlined for MARTI, with minor edits to improve clarity and ensure alignment with MARTI's core values.

10. Summary:
The Metadata Requirements Outline document serves as a foundational guide for establishing metadata standards across MARTI, detailing criteria for task-based and suite-based requirements, preservation, and quality standards for AI-generated outputs. This version has been edited for precision and consistency.