Marti Output-Focused Metadata Rubric (Version 1.1)

Subtitle: Quality Standards for Generative AI Outputs with Anticipation of Scaling Policies
Last edited: Carrie Bickner. Nov 14, 2024 10:45AM EST

1. Introduction

This document provides a rubric for evaluating and applying metadata quality standards within the MARTI (Metadata for AI Responsibility, Transparency, and Integrity) framework. The rubric emphasizes key criteria—accuracy, fidelity, and ethical integrity—that guide metadata creation for generative AI outputs. By applying the rubric, contributors ensure that AI-generated content is transparently documented and ethically aligned with MARTI's principles. This version also anticipates Responsible Scaling Policies, where metadata standards scale alongside AI model capabilities and potential risks with increasing AI autonomy.

2. Key Metadata Criteria

The MARTI Output-Focused Metadata Rubric includes the following primary criteria:

- Accuracy: Metadata must precisely reflect the content and context of the AI output, avoiding misrepresentation or oversimplification.
- Fidelity: The metadata should provide a faithful account of the AI's processes, including notable dependencies, adaptations, or biases.
- Ethical Integrity: Contributors are responsible for ensuring metadata reflects the ethical considerations surrounding the AI output, including considerations of bias, fairness, and transparency.

3. Loose-to-Strict Scale with Scaling Policies

The loose-to-strict scale in MARTI provides flexibility in applying metadata standards based on the intended purpose, risk level, and context of the AI output. In alignment with Responsible Scaling Policies, this scale can be adapted to match graduated safety standards, such as AI Safety Levels, where higher model capabilities and risks require stricter metadata documentation.

The scale ranges from 'Loose' (for minimal impact outputs) to 'Strict' (for high-impact outputs or those with greater ethical implications):

- **Loose:** Minimum metadata requirements are met, suitable for basic outputs with limited impact (e.g., simple illustrations).
- **Moderate:** An intermediate level of metadata detail, offering essential contextual information without exhaustive documentation. Suitable for standard outputs or content with moderate public reach.
- **Strict:** Full and comprehensive metadata documentation covering all aspects of provenance, accuracy, and ethical responsibility. Ideal for complex, high-impact outputs, such as advanced LLMs or outputs requiring regulatory compliance and transparency.

As models increase in capability and potential impact, MARTI's scale should be applied in conjunction with relevant Responsible Scaling Policies, ensuring that metadata rigor matches the risk and scope of the AI output.

4. Applying the Rubric by Output Type and Scaling Level

To determine the appropriate metadata rigor, contributors should consider both the nature and impact of the output and any applicable scaling standards. The following table provides examples of how the rubric aligns with Responsible Scaling Policies for different types of generative AI outputs:

Illustrations/Artworks: Typically require Loose to Moderate adherence, with higher levels for broader or more sensitive applications.

Marketing Content: Requires Moderate to Strict adherence, covering ethical implications, audience reach, and transparency.

Scientific/Research Outputs: Demand Strict metadata for accuracy, fidelity, and ethical integrity, with possible scaling to the highest levels as model complexity increases.

LLM/High-Impact Outputs: Necessitate the Strictest adherence, fully documenting provenance, ethical responsibility, dependencies, and in line with Responsible Scaling Policies.

5. Ethical Considerations and Compliance

Ethical integrity is a foundational principle within MARTI, especially as model capabilities grow. Contributors must ensure metadata aligns with MARTI's ethical guidelines and any Responsible Scaling Policies, such as AI Safety Levels, particularly for high-impact outputs. Metadata should accurately reflect known biases, limitations, or ethical concerns of the AI, while complying with relevant standards like C2PA to maintain transparency and accountability.

6. Quality Standards for Metadata

MARTI establishes rigorous quality standards for metadata, with a focus on:

- Consistency: Ensuring terminology, formatting, and detail levels adhere to MARTI guidelines and Responsible Scaling Policies.
- Completeness: All required fields are populated as per the loose-to-strict scale and scaling policies, ensuring full contextual capture.
- Adaptability: Metadata practices evolve with AI advancements, with periodic reviews to keep MARTI standards aligned with Responsible Scaling Policies.

7. Future Considerations and Versioning

As generative AI and scaling standards evolve, MARTI's metadata rubric will be reviewed periodically to incorporate Responsible Scaling Policies and best practices. Each update will be accompanied by a versioning record, ensuring transparency and traceability in MARTI's metadata standards as capabilities and ethical requirements grow.

```
MARTI Metadata Record: Output-Focused Metadata Rubric (Version
1.1, Edited)


1. Record ID: MARTI_OUTPUT_METADATA_RUBRIC_1_0_EDITED


2. Title: MARTI Output-Focused Metadata Rubric


3. Version: 1.0 (Edited)


4. Author/Contributor:
- 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 establish a rubric for applying metadata quality
standards across generative AI outputs in MARTI, focusing on
accuracy, fidelity, and ethical integrity.


7. Core Changes:
   - Minor edits for clarity and alignment with MARTI's standards
on ethical integrity and metadata consistency.
```

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

9. Provenance: This record establishes the initial metadata rubric for generative AI outputs within MARTI, capturing key standards for quality, consistency, and compliance.

10. Summary:
The Output-Focused Metadata Rubric provides quality standards for metadata applied to generative AI outputs, incorporating a loose-to-strict scale for adjusting rigor based on output type and impact. This document emphasizes MARTI's commitment to ethical transparency and high standards for metadata integrity.