FHIR R4 AI Attribution: Complete Implementation Guide

Overview

When implementing AI-generated content in FHIR R4, you need to capture two distinct pieces of information:

1. **That an AI was used** - a universal flag indicating AI involvement
2. **Which specific AI was used** - detailed attribution for the particular AI system

This guide provides a scalable solution that works across all FHIR resource types.

Final Recommendation

1. Universal AI Flag: AIAST Security Label

Use the official HL7 **AIAST (Artificial Intelligence Asserted)** security label to mark any AI-generated FHIR resource:

```json
{
  "resourceType": "Observation", // or any FHIR resource type
  "meta": {
    "security": [
      {
        "system": "http://terminology.hl7.org/CodeSystem/v3-ObservationValue",
        "code": "AIAST",
        "display": "Artificial Intelligence asserted"
      }
    ]
  }
}
```

Benefits:

- Official HL7 standard specifically designed for AI attribution
- Works universally across all FHIR resource types
- Cross-standard compatibility (FHIR, HL7 v2, CDA, DICOM, XDS)
- Enables AI systems to identify prior AI content (prevents model collapse)
- Simple binary flag that answers "was AI involved?"

2. Specific AI Details: Provenance Resource

Create a Provenance resource to capture detailed information about which AI system was used:

```json
{
  "resourceType": "Provenance",
  "target": [
    {
      "reference": "Observation/ai-generated-obs-123" // or any resource
    }
  ],
  "agent": [
    {
      "type": {
        "coding": [{
          "system": "http://terminology.hl7.org/CodeSystem/provenance-participant-type",
          "code": "author"
        }]
      },
      "who": {
        "reference": "Device/gpt4-clinical-v2-config-abc"
      }
    }
  ],
  "activity": {
    "coding": [{
      "system": "http://terminology.hl7.org/CodeSystem/v3-DataOperation",
      "code": "CREATE"
    }]
  },
  "occurredDateTime": "2024-08-18T10:30:00Z"
}
```

Why Provenance over resource-specific elements:

- **Universal scalability**: Works with any FHIR resource type
- **Future-proof**: No need to modify approach for new AI-generated resource types
- **Semantically appropriate**: Provenance is designed for "who did what, when"
- **Rich attribution**: Can capture comprehensive AI system details
- **Audit-ready**: Built for regulatory and compliance requirements

Device Resource for AI Systems

The Provenance resource references a Device resource containing detailed AI system information:

json


```
{
  "resourceType": "Device",
  "id": "gpt4-clinical-v2-config-abc",
  "deviceName": [
    {
      "name": "GPT-4 Clinical Assistant",
      "type": "manufacturer-name"
    }
  ],
  "manufacturer": "OpenAI",
  "version": [
    {
      "type": {
        "coding": [{
          "system": "http://hl7.org/fhir/device-version-type",
          "code": "model-number"
        }]
      },
      "value": "gpt-4-clinical-v2.1.3"
    }
  ],
  "type": {
    "coding": [{
      "system": "http://snomed.info/sct",
      "code": "706687001",
      "display": "Software"
    }]
  },
  "property": [
    {
      "type": {
        "coding": [{
          "system": "http://your-org.com/ai-properties",
          "code": "training-cutoff"
        }]
      },
      "valueString": "2024-01-31"
    },
    {
      "type": {
        "coding": [{
          "system": "http://your-org.com/ai-properties",
          "code": "configuration-id"
        }]
      },
      "valueString": "config-abc-clinical-high-precision"
    }
```

```
  ]
}
```

Complete Implementation Examples

Example 1: AI-Generated Observation

```json
{
  "resourceType": "Observation",
  "id": "ai-generated-obs-123",
  "meta": {
    "security": [
      {
        "system": "http://terminology.hl7.org/CodeSystem/v3-ObservationValue",
        "code": "AIAST"
      }
    ]
  },
  "status": "final",
  "code": {
    "coding": [{
      "system": "http://loinc.org",
      "code": "72133-2",
      "display": "Risk assessment"
    }]
  },
  "subject": {
    "reference": "Patient/patient-123"
  },
  "valueString": "High risk for cardiovascular events based on clinical indicators"
}
```

Example 2: AI-Generated Appointment

json

```
{
  "resourceType": "Appointment",
  "id": "ai-scheduled-456",
  "meta": {
    "security": [
      {
        "system": "http://terminology.hl7.org/CodeSystem/v3-ObservationValue",
        "code": "AIAST"
      }
    ]
  },
  "status": "booked",
  "serviceType": [{
    "coding": [{
      "system": "http://terminology.hl7.org/CodeSystem/service-type",
      "code": "124",
      "display": "General Practice"
    }]
  }],
  "participant": [
    {
      "actor": {
        "reference": "Patient/patient-123"
      },
      "status": "accepted"
    }
  ]
}
```

Example 3: Universal Provenance Pattern

```json
{
  "resourceType": "Provenance",
  "target": [
    {"reference": "Appointment/ai-scheduled-456"}
  ],
  "agent": [
    {
      "type": {
        "coding": [{
          "system": "http://terminology.hl7.org/CodeSystem/provenance-participant-type",
          "code": "author"
        }]
      },
      "who": {
        "reference": "Device/scheduling-ai-assistant-v2"
      }
    },
    {
      "type": {
        "coding": [{
          "system": "http://terminology.hl7.org/CodeSystem/provenance-participant-type",
          "code": "verifier"
        }]
      },
      "who": {
        "reference": "Practitioner/supervising-physician-789"
      }
    }
  ],
  "activity": {
    "coding": [{
      "system": "http://terminology.hl7.org/CodeSystem/v3-DataOperation",
      "code": "CREATE"
    }]
  },
  "occurredDateTime": "2024-08-18T10:30:00Z",
  "reason": [{
    "coding": [{
      "system": "http://your-org.com/scheduling-reasons",
      "code": "optimal-availability",
      "display": "Optimized based on provider availability and patient preferences"
    }]
  }]
}
```

Scalability Across Resource Types

This approach scales seamlessly to any FHIR resource type that might be AI-generated:

Resource Type	AI Use Case	Implementation
Observation	AI diagnosis, risk assessment	AIAST label + Provenance
DiagnosticReport	AI radiology interpretation	AIAST label + Provenance
Appointment	AI scheduling optimization	AIAST label + Provenance
CarePlan	AI treatment recommendations	AIAST label + Provenance
MedicationRequest	AI prescription suggestions	AIAST label + Provenance
Encounter	AI clinical documentation	AIAST label + Provenance
DocumentReference	AI-generated clinical notes	AIAST label + Provenance

Implementation Benefits

Standards Compliance

- Uses official HL7 terminology (AIAST)
- Leverages standard FHIR R4 resources (Provenance, Device)
- No custom extensions required

Scalability

- Universal pattern works across all resource types
- Consistent implementation regardless of AI application domain
- Future-proof for new AI use cases

AI-Aware Processing

- AIAST flag enables AI systems to identify prior AI content
- Prevents AI feedback loops and model collapse
- Supports downstream decision-making processes

Audit and Compliance

- Comprehensive provenance tracking
- Regulatory compliance support
- Clear attribution chain for AI-generated content

Query Patterns

Find all AI-generated resources

```
GET /fhir/Observation?_security=http://terminology.hl7.org/CodeSystem/v3-ObservationValue|AIAST
```

Find resources by specific AI system

```
GET /fhir/Provenance?agent:Device.identifier=gpt4-clinical-v2
```

Complex queries combining both

```sql
-- Find all AI-generated Observations from GPT-4 systems
SELECT o.* FROM Observation o
JOIN Provenance p ON p.target = o.id
JOIN Device d ON p.agent.who = d.id
WHERE o.meta.security CONTAINS 'AIAST'
AND d.deviceName LIKE '%GPT-4%'
```

Summary

The combination of **AIAST security labels** and **Provenance resources** provides a comprehensive, scalable, and standards-compliant approach to AI attribution in FHIR R4:

1. **AIAST security label**: Universal flag that AI was involved (works on all resource types)
2. **Provenance resource**: Detailed AI system attribution (scales to any resource type)
3. **Device resource**: Rich AI system metadata and configuration details

This approach ensures consistent AI attribution across your entire FHIR ecosystem while maintaining semantic accuracy and regulatory compliance.