HealthDevHub Meetups

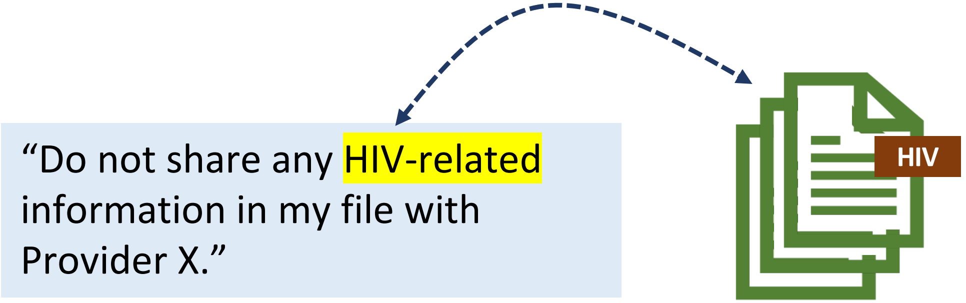
Data Segmentation for Privacy and Consent

April 30, 2024

Mohammad Jafari, Independent Subject-Matter Expert

Security Labels

- metadata
  - associated with a unit of data
    - bundle, resource, or portion of a resource
  - marks additionally protected information



“Do not share any HIV-related information in my file with Provider X.”

- can be referenced in policies

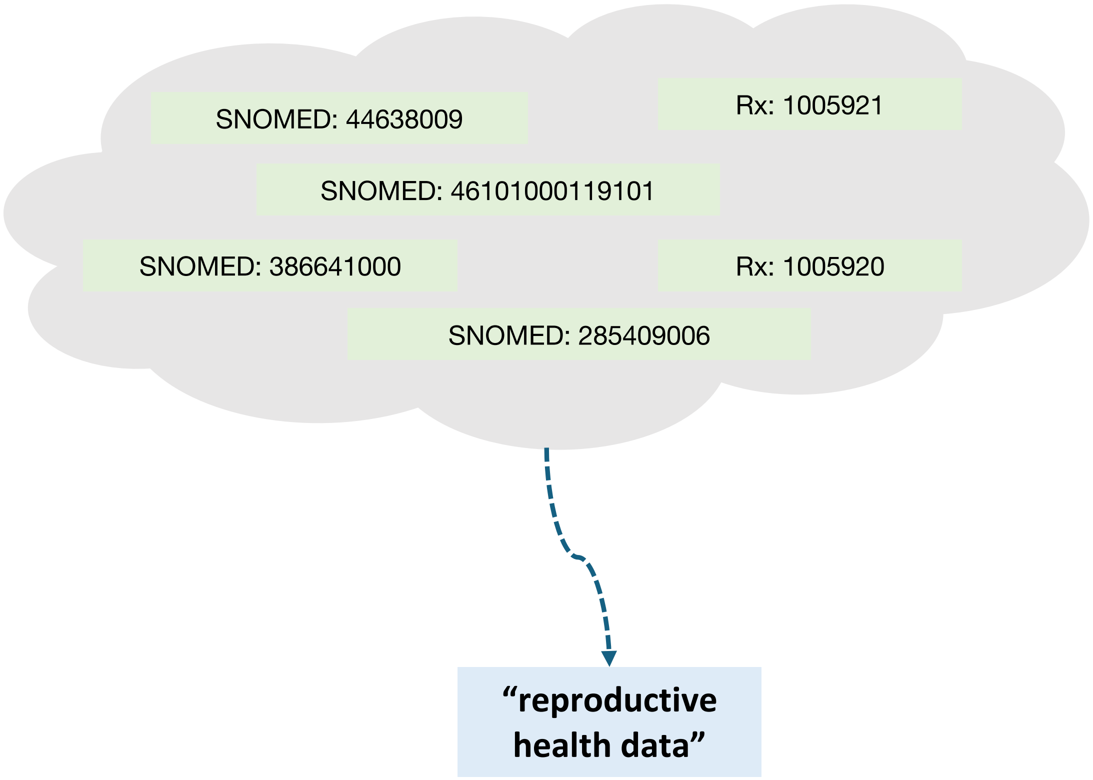
- Identifying data elements that are subject to additional privacy/security controls.
- Examples from US jurisdictions
  - *Substance Use Data*, *Psychotherapy Notes*, *Behavioral Health Data*, *Reproductive Health Data*

Key Components

- Tag data
  - Security Labeling Service (SLS)
- Record labels
  - `meta.security` and inline labels
  - Standard codes for labels
    - HL7 terminology, FHIR DS4P IG
  - Label metadata
- Process Labels
  - incorporate in authorization decision e.g., consent enforcement
  - incorporate in workflow, e.g., prevent sensitive information from access
  - incorporate in UI/UX, e.g., mark sensitive data

Security Labeling Service

- Rudimentary labeling plainly based on pre-determined value sets
- Probabilistic labeling
- More sophisticated technologies
  - Related resources
  - Encounter context
  - Facility type
  - Unstructured text: NLP and LLM



SNOMED: 44638009

Rx: 1005921

SNOMED: 46101000119101

SNOMED: 386641000

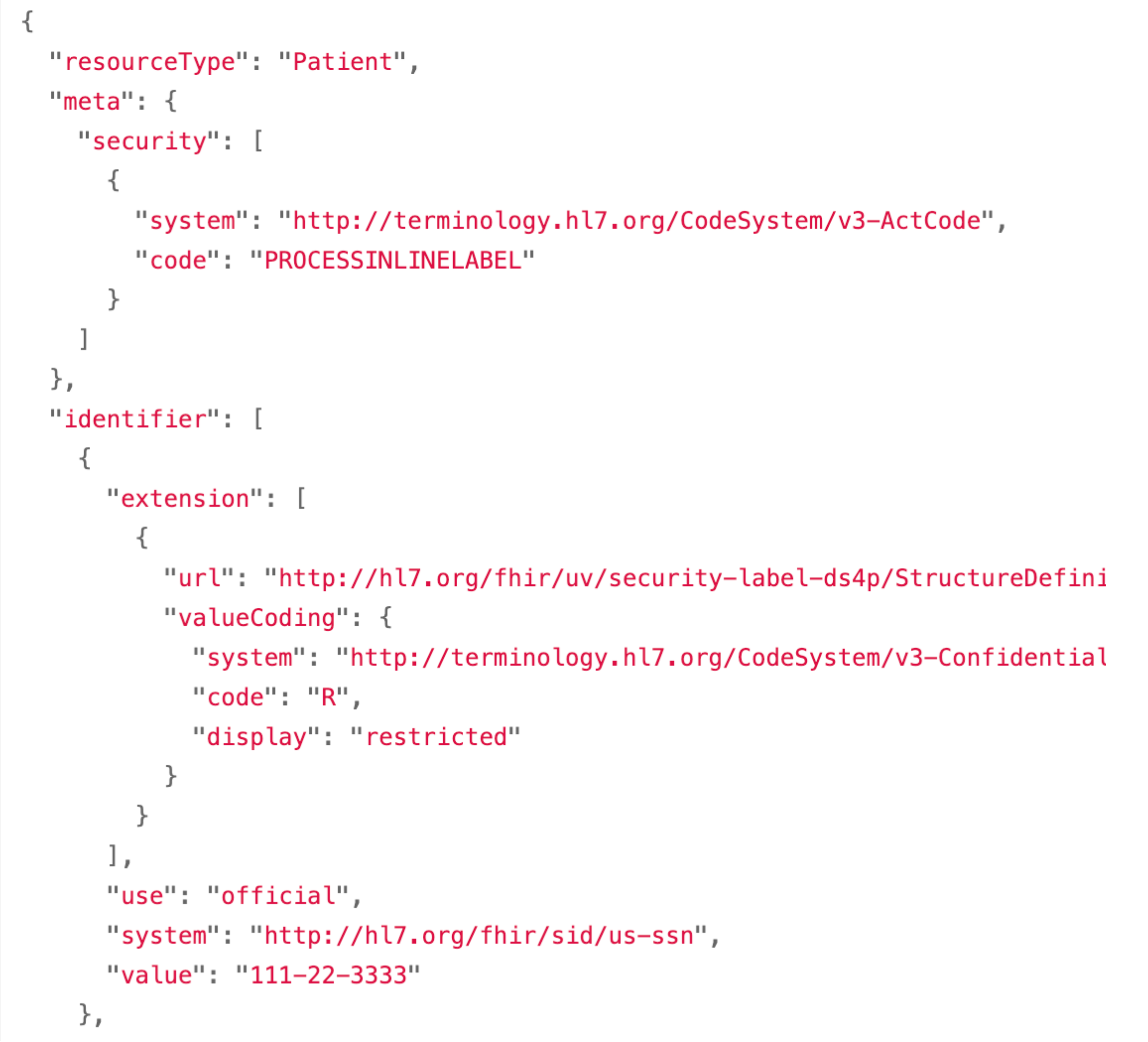
Rx: 1005920

SNOMED: 285409006

“reproductive health data”

Inline Labeling

- Granular assignment of a label to a portion of a resource
    - e.g., residential address is confidential (but not the mailing address)
- Resource-level marker to process inline label
- Extension to record the label on a portion of the resource



```
{
  "resourceType": "Patient",
  "meta": {
    "security": [
      {
        "system": "http://terminology.hl7.org/CodeSystem/v3-ActCode",
        "code": "PROCESSINLINELABEL"
      }
    ]
  },
  "identifier": [
    {
      "extension": [
        {
          "url": "http://hl7.org/fhir/uv/security-label-ds4p/StructureDefini
          "valueCoding": {
            "system": "http://terminology.hl7.org/CodeSystem/v3-Confidential
            "code": "R",
            "display": "restricted"
          }
        }
      ],
      "use": "official",
      "system": "http://hl7.org/fhir/sid/us-ssn",
      "value": "111-22-3333"
    },
```

Metadata

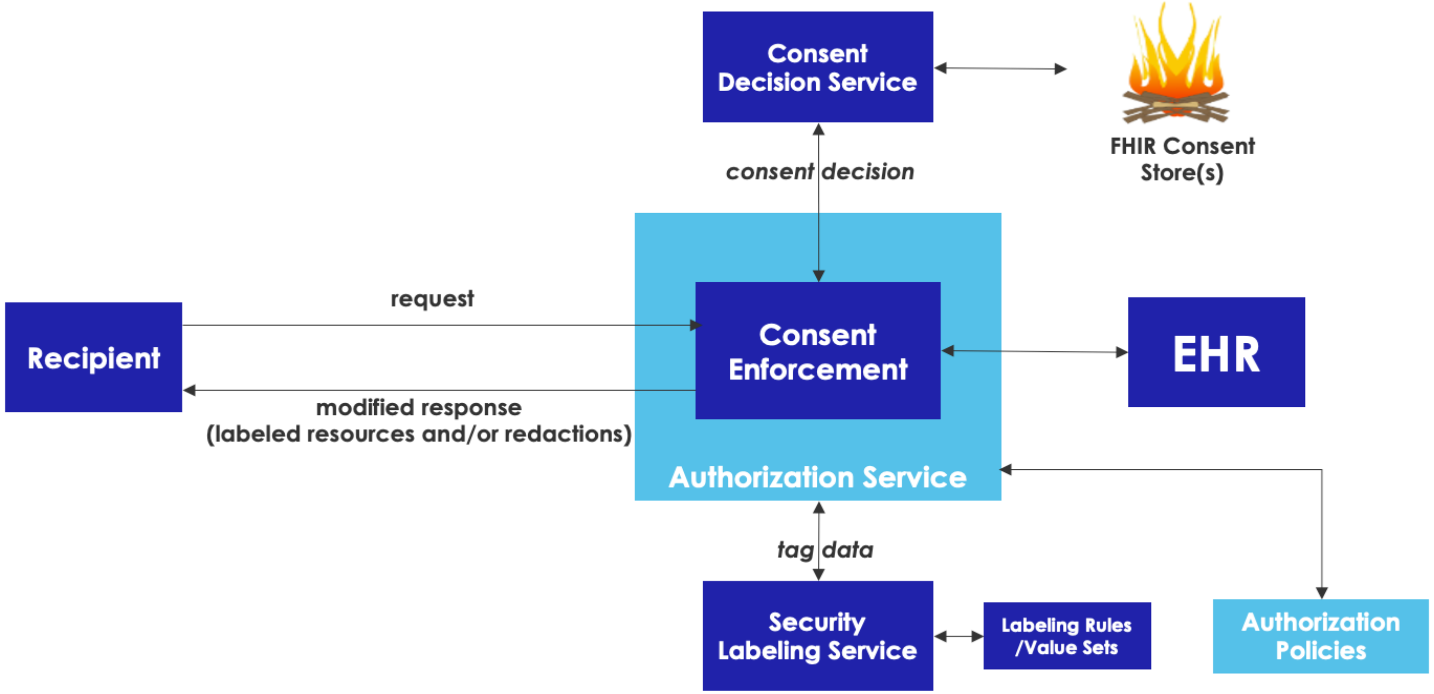
- Who performed the labeling
  - Identity of the entity that applied the label
  - `sec-label-classifier`
- The basis for labeling
  - The law or regulation behind the labeling
  - `sec-label-basis`
- Time stamp
  - Determining whether the data has changed since the decision to label
  - No extensions in the FHIR DS4P IG right now.

Technical Architecture Considerations

- Where does the labeling service reside?
  - EHR, HIE, Third-party service
- Standard API for labeling
- When does the labeling take place?
  - At the time of transaction
    - Always label the latest version of the documents, no need to persist labels or re-label
    - Response-time challenges
  - Offline
    - Batch or bulk labeling of data at rest and persist the labels
    - Advanced processing (e.g., unstructured text) is possible because of the offline nature.

Labeling and Consent

- Consent rules can be based on labels
  - “do not share reproductive health data with provider X”
- Consent decision should incorporate labels
- Data must be labeled at the time of the decision (at the latest)



Challenges and Gaps

- HL7 specifications are available but need to be actively updated and maintained
- HL7 terminology for sensitive categories need to be overhauled
  - More granular codes
  - Deprecate old codes
  - Update definitions
- More implementation guidance on:
  - Standard HL7 codes to use for different classes of sensitive data identified in US regulations
  - Value sets (of clinical codes) tied to each sensitivity category
  - Standard API for SLS