Personal Information Tagging for Logs

https://tools.ietf.org/html/draft-rao-pitfol-02

Sandeep Rao, Shivan Sahib, Ryan Guest, Santhosh CN

IETF 108, PEARG

Jul 27 2020
Recap

Log and Privacy

● Privacy
  ○ Improve ways to identify personal information in logs

● Challenges
  ○ Subjective - many types of log data and formats
  ○ Dictionary, Heuristics query, Data-Set based training model
  ○ Vendor specific schema / privacy policy
  ○ No standard / guidelines - (what) data to be protected and (what) action to enforce
Document Goal

- Enable operators to explicitly detect and filter personal information in logs
- Reference model to enable tagging of personal information at source
- Illustrate ways to attach sensitivity/privacy tagging to log data
- Privacy Control Actions - Consumer Role-based / Case-based
Updates from -01

- Address review comments
- Defining a log privacy schema
- Proposal for a Personal Information Identifier Registry
- Log processing and Access Control
# Personal Information Identifier Registry

**Example:** Define “namespace” for personal information

<table>
<thead>
<tr>
<th>Name</th>
<th>Abstract Data Type</th>
<th>Description</th>
<th>Sensitivity [1-High 5-Normal]</th>
</tr>
</thead>
<tbody>
<tr>
<td>nationalIdentity</td>
<td>String</td>
<td>National IDs issued by sovereign governments. Eg., SSN</td>
<td>1</td>
</tr>
<tr>
<td>drivingLicense</td>
<td>String</td>
<td>Driving License number</td>
<td>1</td>
</tr>
<tr>
<td>taxIdentity</td>
<td>String</td>
<td>Tax identification numbers</td>
<td>1</td>
</tr>
<tr>
<td>creditCardNumber</td>
<td>String</td>
<td>Credit cards</td>
<td>1</td>
</tr>
<tr>
<td>bankAccount</td>
<td>String</td>
<td>Bank account number</td>
<td>1</td>
</tr>
<tr>
<td>dateOfBirth</td>
<td>Date</td>
<td>Date of Birth</td>
<td>2</td>
</tr>
<tr>
<td>personName</td>
<td>String</td>
<td>Person name</td>
<td>1</td>
</tr>
<tr>
<td>emailAddress</td>
<td>String</td>
<td>Email</td>
<td>2</td>
</tr>
<tr>
<td>phoneNumber</td>
<td>Number</td>
<td>Phone</td>
<td>1</td>
</tr>
<tr>
<td>zipCode</td>
<td>Integer</td>
<td>Zip codes</td>
<td>5</td>
</tr>
<tr>
<td>ipAddress</td>
<td>ipv4Address</td>
<td>IPv4 or IPv6 Address</td>
<td>4</td>
</tr>
<tr>
<td>dateTimeSeconds</td>
<td>dateTimeSeconds</td>
<td>Address seconds</td>
<td>5</td>
</tr>
<tr>
<td>age</td>
<td>Integer</td>
<td>Age</td>
<td>2</td>
</tr>
<tr>
<td>ethnicGroup</td>
<td>String</td>
<td>Ethnic group</td>
<td>1</td>
</tr>
<tr>
<td>genderIdentity</td>
<td>String</td>
<td>Gender</td>
<td>1</td>
</tr>
<tr>
<td>macAddress</td>
<td>macAddress</td>
<td>MAC Address</td>
<td>4</td>
</tr>
</tbody>
</table>
Privacy Tagging

Field level tagging

```plaintext
<120> Apr 18 16:32:58 10.0.1.11 QAUDJRN: [AF@0 event="AF-Authority failure" violation="A-Not authorized to object" actual_type="AF-A" vjrn_seq="1001363" timestamp="20120418163258988000" vjob_name="QPADEV000B" {personName="XYZZY" pii_sensitivity_level=1} {emailAddress="xyz@foo.com" pii_sensitivity_level=2} object_type="*FILE" pgm_name="" pgm libr="" workstation=""]
```

Log level tagging

```plaintext
<120> Apr 18 16:32:58 10.0.1.11 QAUDJRN: [AF@0 event="AF-Authority failure" violation="A-Not authorized to object" actual_type="AF-A" vjrn_seq="1001363" timestamp="20120418163258988000" vjob_name="QPADEV000B" personName="XYZZY" emailAddress="xyz@foo.com" object_type="*FILE" pgm_name="" pgm libr="" workstation=""}, {pii_sensitivity_level=1}]
```
Privacy Tagging

Redaction Action

```
<120> Apr 18 16:32:58 10.0.1.11 QAUDJRN: [AF@0 event="User Logged In"
timestamp="20200418163258938000"
   personname="xyz@foo.com" ip_addr="10.21.23.1" phonenum="+912232422213"],
   pii_metadata= {
      "piiDescriptor": {
         "field": "username",
      },
   "action": {
         "SensitivityLevelTag": "4",
         "Action": "FullAnonymization",
   }
}
```
Future Work

- Privacy preservation across log transformations
- Change of Privacy marking policy / classifications
- Out-of-band mechanism to notify privacy schema
Open Questions

- Continue to pursue this work in PEARG?
- Does the scope need to change?
- Request for RG adoption?