

## **RDFa and Personally Identifiable Information (PII)**

### **Table of Contents**

RDFa and Personally Identifiable Information (PII).....	1
RDFa in the Header.....	2
Conflicts with the Dublin Core Syntax.....	2
Conflicts with GRDDL Specification.....	2
“Double Truth”, Rights and Copyrights.....	2
RDFa in the Document Body.....	3
Privacy for the Individual and Social Networking.....	3
Redaction Workaround.....	3
References.....	5

## ***RDFa in the Header***

### **Conflicts with the Dublin Core Syntax**

The Candidate Recommendation [RDFa] for the usage of RDFa in the <meta> and <link> elements is incompatible with the Dublin Core Metadata Initiative recommendations for the same purpose [DCMI-XHTML].

### **Conflicts with GRDDL Specification**

The Candidate Recommendation [RDFa] header syntax conflicts with the Gleaning Resource Descriptions from Dialects of Languages [GRDDL] W3C Recommendation of 11 September 2007.

### **“Double Truth”, Rights and Copyrights**

The use of similar attributes for the meta element, for example @{name} and @{property} can lead to processing distillations which suffer from a heresy, called Double Truth. In its prime, around the 12<sup>th</sup> Century, Double Truth got many a heretic burned at the stake; pagans (with the wrong Truth) having been likewise “educated” for several previous millennia.

The problem is in the nature of meta data – normally hidden; it gives no evidence of additions, modifications or redactions (deletions) when an e-document is transported or processed. This creates an issue for Copyright Claims, Licenses and other rights claims to Intellectual Property because the locus of the “source document” is less certain.

This problem is *manageable for every class of entity* when the meta data is in the (displayed visible) <body> of the document and rendered in “What You See Is What You Get” (WYSIWYG) fashion, although there remain privacy issues for persons. As befits its long history, hidden meta data is problematic for every class of entity, Public or Private.

## ***RDFa in the Document Body***

### **Privacy for the Individual and Social Networking**

It is the right of every individual in our connected society to not to leave an electronic trail of activities, proclivities, preferences, associations etc., to disconnect from the society at large. Public Figures may have surrendered that right, but had it as children, and may reclaim the right for reasons of old age or diminished mental capacity – signs of decency and respect rarely if ever challenged. All human societies protect the weak in one fashion or another.

The express purpose of RDFa is to facilitate the composition of an electronic trail. This is not to say that RDFa is evil and should be banned, rather, RDFa is a tool for use only in the Public Domain. It is commonly believed that once information is in the Public Domain, there is no way to privatize the information. This is true, for example, when a Law Enforcement entity has an unrestricted ability to follow an electronic trail. However, with the decency and respect mentioned above, and a bit of precautionary help from authors, it is possible to mark Personally Identifiable Information (PII) for automated redaction, privatization, when an individual chooses, for whatever reason, to reclaim their privacy.

### **Redaction Workaround**

Markup encapsulates and segregates meta data. Using the existing RDFa attribute syntax and the Personally Identifiable Information Namespace [PII], a BLOG author, for example, can tag in-line or block segments of XHTML for automated redaction at a later date.

Redactions can be accomplished with XSLT processing and the application of the Redact Unless Static Text [RUST] Rule.

The largest encapsulating node (div, span, p) only bears the RDFa attributes which reference the PII namespace. If possible, use `@{instanceof}`, or if that attribute already exists, `@{href}`. These two options should cover almost every case. Although it increases the processing complexity, the use of the `xsl:fallback` element or an additional RDFa `@{fallback}` attribute seems possible.

### Example 1 (from [RDFa]):

```
<html
  xmlns="http://www.w3.org/1999/xhtml"
  xmlns:cal="http://www.w3.org/2002/12/cal/ical#"
  xmlns:pii="http://purl.org/pii/terms/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <head><title>Jo's Friends and Family Blog</title></head>
  <body>
    <p>
      I'm holding
      <span property="cal:summary">
        one last summer Barbecue
      </span>,
      on
      <span property="cal:dtstart" content="20070916T1600-0500"
        datatype="xsd:datetime" instanceof="pii:witness">
        September 16th at 4pm
      </span>.
    </p>
  </body>
</html>
```

### Example 2 (from [RDFa]):

```
<html
  xmlns="http://www.w3.org/1999/xhtml"
  xmlns:cal="http://www.w3.org/2002/12/cal/ical#"
  xmlns:pii="http://purl.org/pii/terms/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  >
  <head><title>Jo's Friends and Family Blog</title></head>
  <body>
    <p style="font-family:monospace;color:Blue;" instanceof="cal:Vevent"
href="pii:witness">
      I'm holding
      <span property="cal:summary">
        one last summer Barbecue
      </span>,
      on
      <span property="cal:dtstart" content="20070916T1600-0500"
        datatype="xsd:datetime">
        September 16th at 4pm
      </span>.
    </p>
  </body>
</html>
```

## **References**

[RDFa] W3C, 2008-06-20, RDFa in XHTML: Syntax and Processing, A collection of attributes and processing rules for extending XHTML to support RDF, <<http://www.w3.org/TR/rdfa-syntax>>

[DCMI-XHTML] DCMI, 2003-11-30, Expressing Dublin Core in HTML/XHTML meta and link elements, <<http://www.dublincore.org/documents/dcq-html/>>

[GRDDL] W3C, 2007-09-11, Gleaning Resource Descriptions from Dialects of Languages, <<http://www.w3.org/TR/grddl/>>

[PII] RUSTPrivacy.org, 2007-, The Personally Identifiable Information Namespace, <<http://purl.org/pii/terms/>>

[RUST] Gannon J. Dick, 2006-2008, <<http://www.RUSTPrivacy.org/>>