

Disambiguation – first proposal (as in the 20120829 version)

<p>

```
<span its-entity-type-source-ref="http://nerd.eurecom.fr/ontology"  
its-entity-type-ident-ref="http://nerd.eurecom.fr/ontology#Place"  
its-disambig-source-ref="http://dbpedia.org/"  
its-disambig-ident-ref="http://dbpedia.org/resource/Dublin"  
its-disambig-type="entity">Dublin</span>
```

is the

```
<span  
its-disambig-source-ref="  
http://www.w3.org/2006/03/wn/wn20/rdf/wordnet-synset.rdf"  
its-disambig-ident-ref="  
http://www.w3.org/2006/03/wn/wn20/instances/worsense-capital-noun-3"  
its-disambig-type="its:lexicalConcept">capital</span>
```

of Ireland.

</p>

Disambiguation – second proposal (last version)

```
<p><span
  its-disambig-class-ref="http://nerd.eurecom.fr/ontology#Place"
  its-disambig-ident-ref="http://dbpedia.org/resource/Dublin"
  its-disambig-granularity="entity"
>Dublin</span> is the <span
  its-disambig-source="Wordnet3.0"
  its-disambig-ident="301467919"
  its-disambig-granularity="lexicalConcept"
>capital</span> of Ireland.</p>
```

- The disambigSource can be derived from the identifier
- Allow non-URI identifiers
- Generalized entityType to targetType

Disambiguation – current issues

- The value of `disambigLevel` is (practically) entailed by `disambigIdent*`. Should it even exist?
 - Suggestion: Define a mapping for known classes to levels, inferring the value automatically, but keep it in.
- Calling the disambiguation target a "target" (as in "target type") is confusing, since "target" usually refers to "target language"
 - Suggestion: "disambiguation target type"?

Text Analysis Annotation (global)

```
<its:rules
  xmlns:its="http://www.w3.org/2005/11/its" version="2.0">
  <its:textAnalysisAnnotationRule selector="/text/its:rules/
[@id='dublin-disambig']"
    its:textAnalysisAgentRef="http://enrycher.ijis.si/"
    its:textAnalysisConfidenceScore="0.95" />
  <its:disambiguationRule id="dublin-disambig" selector="/text/
body/p/[@id='dublin']" its:disambigLevel="its:entity"
    its:disambigIdentRef="http://dbpedia.org/resource/Dublin" />
</its:rules>
```

- Global rules make it possible to precisely select and state what was processed with some tool

Text Analysis Annotation (local)

```
<body its-text-analysis-agent-ref="http://enrycher.ijs.si">  
  <p><span  
its-disambig-ident-ref="http://dbpedia.org/resource/Dublin"  
its-disambig-level="its:entity"  
its-text-analysis-confidence-score="0.95"  
>Dublin</span> is the capital of Ireland.</p>  
</body>
```

- The local variant makes it hard to select the scope of the agent.
- Current workaround: the default behavior selects “all its-* annotations contained within the annotated element and its descendants”
 - could be even less strict – “some its-* annotations may have been...”

Text Analysis Annotation (tool info consolidation)

- Listing the tools that touched the document
- Some loss of expressivity is ok – document-level annotation is what is usually necessary
- Currently, the most precise tool annotation is only possible via global rules (in order to select attributes) and quite verbose