



MultilingualWeb-LT: Quality Data Categories

Target audience

- Localization Service Providers doing Quality Assessment tasks
- Content creators doing quality verification
- Authors marking up authoring errors
- Posteditors providing information on efficiency/MT quality

Motivations

- Streamline linguistic review process
 - Reduce multiple reviews of similar content
 - Use provenance of segments to reinforce assessment
- Build real-time workflow rules based on past translator performance.
- Build content profile from translation agent, MT confidence, domain, purpose, etc.
- Use Linked Data to capture cross-silo process profile of content.

Description

- Two (complex) data categories:
 - Errors themselves
 - Overall error profiles
- Provides flexible way to use different error models in a consistent, standardized way.
- Implement standard marking regardless of metric used.

qualityerror purpose

- Mark quality issues against a specified framework
- Provide an audit trail for quality resolution
- Flexible framework for various tasks

qualityerror description

- type?
- ruleset?
- severity?
- note?
- agent?

The `<span`

```
its-qa-type="syntactic error"
```

```
its-qa-ruleset="SAE J2450"
```

```
its-qa-severity="major"
```

```
its-qa-note="bad grammar"
```

```
its-qa-agent="ABCReview">verbs agrees</span> with the subject.
```

qualityprofile purpose

- Defines the overall profile(s) used.
- Allows for flexibility in picking an appropriate profile/metric
- Does not require local tagging: can be summary only (e.g., the traditional LISA QA Model software or SAE J2450 were used with no in-line tagging)
- Defaults to LISA QA Model

qualityprofile description

- name?
- uri?
- pass?
- score?
- agent?

```
<meta  
  name="qualityProfile"  
  content="name:LISA QA Model;  
  uri:http://www.example.com;  
  pass:no;score:85%;  
  agent:ABCReview"  
>
```

We know this would not work exactly as is, but it shows the concept.

Issues and Notes

- Verboseness
- Would these be used in general? (More sophisticated than most current processes)
- Implementation commitments and timeframe
- Represents perspective of upcoming QT Launchpad project (which aims to standardize/unify error metrics).