A Definitions and Examples for the Annotation Scheme

We give the definitions and the examples for the **Resource Roles** and **Resource Functions** for the resource citations. To show the evidence, we further highlight the arguments (which are mostly the key verbs before the citation) for identifying the function and the arguments (which are mostly the target nominals ahead the citation) for identifying the role.

A.1 Resource Roles

A.1.1 Material

- **Data** The data consists of datasets, databases or corpus.
  E.g. The selection of C and the RKHS has been done as indicated in <CITE> for Adult and Web data sets and in <CITE> for Banana, Diabetes and Splice data sets.

A.1.2 Method

- **Tool** The tool consists of toolkits, softwares, systems or projects.
  E.g. We use a CRF++ based POS tagger for Hi, which is freely available from <CITE>. For En, we use the Twitter POS tagger.

- **Code** The code consists of codebases, libraries or implementations.
  E.g. The SVM computations are performed using the freely available Spider Matlab machine learning package available at <CITE>.

- **Algorithm** The algorithm consists of methods, models or solutions.
  E.g. The model uses a maximum entropy learner <CITE>, training one binary classifier per sense.

A.1.3 Supplement

- **Website** The website consists of homepages, services, on-line platforms or interfaces.
  E.g. Answers and Live Search QnA <CITE>, have been rapidly gaining popularity among Web users interested in sharing information online.

- **Document** The document consists of supplements, tutorials, specifications or guidelines.

E.g. Unfortunately there exists substantial disagreement regarding the interpretation of existing approaches see <CITE>.

- **Paper** The paper is a short/long conference paper taken from sites.
  E.g. In this section, we propose a use case of application of Linked Aoki-Kinoshita et al. Journal of Biomedical Semantics 2014, 6:3 Page 4 of 13 <CITE>.

- **Media** The media consists of games, musics or videos.
  E.g. Formulating MI using Renyi entropy, and Gaussian Example video clips can be viewed at <CITE>.

- **License** The license provides access to and details of the used licenses.
  E.g. The Creative Commons Public Domain Dedication waiver (<CITE>) applies to the data made available in this article, unless otherwise stated.

A.2 Resource Functions

- **Use** The cited resource is used in this paper's work.
  E.g. We use a local search engine, <CITE>, which accepts the SearchTerm and LocationTerm as two query fields and returns the search results from a business listings database.

- **Produce** The cited resource is first produced or released by this paper's work.
  E.g. Finally, to aid other AAC researchers, we have publicly released our crowdsourced AAC collection, word lists and best-performing language models <CITE>.

- **Introduce** The background, characteristic or applications of the cited resource is described in the context sentences. It mostly appears in the introductory parts of a paper, such as *Introduction*, *Background* and *Related Work*.
  E.g. The central and most widely-used resource in the field is CyanoBase (<CITE> Nakao et al., 2010).

- **Compare** The resource is compared with other resources. It often appears in the experimental parts of a paper.
  E.g. Our logistic regression linear parser

and re-implementation of Chen and Manning (2014) give comparable accuracies to the perceptron ZPar <CITE> and Stanford NN Parser, respectively.

- **Extend** The resource is the foundation of this paper's work or some improvements are made based on the resource. Note that the definition for **Extend** class and the definition for **Use** class are substantially distinguishing. The **Extend** means that the cited resource is improved, upgraded or changed to work for other problems. While the **Use** means in most cases there are no changes on the original cited resource.
  E.g. We integrate a phonetic-based encoding scheme, UrduPhone, a feature-based similarity function, and a clustering algorithm, ... and Double Metaphone <CITE>.

- **Other** The function do not belong to the above 5 categories will be classified into Other.
  E.g. We would also like to acknowledge the Cognitive Rhythms Collaborative (<CITE>), ...