

The Read The Web homework assignment for the the week of March 2-9 has two parts:

1. Due by \*wednesday noon\* : each group will post a brief project writeup (1-2 pages) to the class Kiva site
2. Before class thursday: Read every report, and come prepared to discuss opportunities for synergy among groups.

**Reminder:** the projects we heard about in today's class included:

- General purpose bootstrap learner for relation extractions
- Active learning: which questions to ask a trainer given a limited budget
- Co-reference resolution.
- Mine extracted knowledge base for regularities and provide feedback to bootstrap learning methods.
- Semantic role labeling, especially nominalization.
- Extract associations between typed entities; e.g., the very general relation: Co-occursMostFrequentlyWith(<entity\_type1, entity\_type2>)

**What to provide in your report:**

- A concise, precise description of the functionality you intend to provide. The closer you come to a tentative API the better. Be sure to describe what your system will learn, from what, using what redundant sources of information
- Any assumptions and requirements of your code (e.g., only applies to plain ASCII text; requires NER from others)
- a short bibliography of relevant research papers your will read (really) and build on.
- an illustration/scenario of how your system would apply to and learn in the following problem: learning to extract information about a biology faculty member, her publications, conferences, advisees, and their research topic, using any combination of their website, online bio, and online publications. (this will allow us to develop a joint scenario involving all projects, so we can identify synergies among our groups)