CSCI 7000-IR Fall 2008 Quiz 3

Name: _________________________

On my honor, as a University of Colorado at Boulder student, I have neither given nor received unauthorized assistance on this work. _____________________________

1. (5 points) True or False: It’s gotten pretty cold out.

2. (10 points) One of the main capabilities of information extraction systems is to be able to locate and analyze the entities of interest in a text. Describe a way that such ability might be used to improve the performance of an ad hoc information retrieval system.

3. (5 points) Describe two problems in information extraction and sentiment analysis that Microsoft had to solve to deploy the system shown on the screen.

4. (5 Points) True or False: Basic techniques from text-based information retrieval (i.e. Lucene’s functionality) can be adapted to build content-based recommender systems.

5. (10 Points) Briefly describe the basic User x Item model used in collaborative filtering approaches to building recommender systems.

6. Amazon frequently makes recommendations to Jim for travel guides, intro programming texts, and child-rearing books where he’s already purchased adequate books in those categories and doesn’t need new ones.

   a) (5 points) Briefly describe why it’s likely that this happening.

   b) (15 points) Briefly describe (1) what kind of information would be needed to avoid this problem and (2) how you might deploy it.