Brief Bio and (PR)^2: Problems & Pitches – Rants & Raves by Ying Ding

About me
Dr. Ying Ding is an Associate Professor at the School of Library and Information Science, Indiana University. Before she worked as a senior researcher at the University of Innsbruck, Austria and as a researcher at the Free University of Amsterdam, the Netherlands. She has been involved in various NIH and European-Union funded Semantic Web projects. She has published 150+ papers in journals, conferences and workshops. She serves as a Program Committee member for 120+ international conferences and workshops. She is the coeditor of book series called Semantic Web Synthesis by Morgan & Claypool publisher. She is co-author of the book "Intelligent Information Integration in B2B Electronic Commerce" published by Kluwer Academic Publishers. She is also co-author of book chapters in the book "Spinning the Semantic Web" published by MIT Press and "Towards the Semantic Web: Ontology-driven Knowledge Management" published by Wiley. She is the editorial board member of four ISI indexed top journals in Information Science and Semantic Web. Her current interest areas include social network analysis, Semantic Web, citation analysis, knowledge management and application of Web Technology.

Questions

1. What are your main interests in attending the workshop?
I would like to understand the major challenges in translational medicine and how semantic technology can help solve some issues, especially data integration and data discovery.

2. What ideas, methods and tools would you like to share at the workshop?
I would like to share some of the tools our (together with David Wild's team) team has developed:
- Bio-LDA: to calculate the topic distribution of bio entities
- subgraph mining: given any two nodes, we will find the diversity subgraph which connects them
- SLAP (semantic link association prediction): to predicate the connection between any given drug and target.
- PubMed Entity graph: enable literature discovery via connected entity graph in PubMed literature.

3. What do you think are the biggest opportunities or unmet needs in any of: translational medicine, drug discovery, semantic technologies, data visualization, or healthcare information? (feel free to pick those with which you have the most interest/experience)
Data analytic power is not yet deployed in major tools. Sparql is a search method, not the data analytic method. Semantic inference is far too slow and generates high recall and low precision which is not ideal for translational medicine.

How to combine experimental data and literature data in translational medicine? How to enable powerful and efficient data integration?
4. What are the biggest road blocks to realizing these opportunities?
Unique identifiers for major bio-entities are not standardized, more in-house solutions, which will significantly hinder data integration. It is the conflict between distributed vs. central, not easy to find a quick solution.

5. In which of the main areas of emphasis of the workshop (semantics, translational medicine, drug discovery, big data, semantic technologies, visualization and networks) do you work?
semantic technologies, and data mining.

6. What are the biggest challenges in your work currently?
How to combine literature discovery with data discovery? How to discover new knowledge from public data, how to rank results?

7. What are the main sources of funding for your work? How difficult do you consider it to get funding in your area, and why?
Government funding

8. What would you like to learn and achieve at the workshop?
I want to learn other best practices in translational medicine. I want to learn how data and literature discovery can play a role in personalized medicine.