

NLP R&D and Commercial Deployment

Mark Wasson
LexisNexis

Tutorial Outline

1. Academic and commercial perspectives on NLP research: (a.) The role of academic research; (b.) The role of R&D in industry; (c.) Knowledge and technology transfer.
2. What industry wants from NLP technology: (a.) Information overload and the silver bullet; (b.) Costs, productivity, competitive advantage, profit; (c.) Why industry likes simple alternatives to NLP.
3. Evaluating NLP: (a.) Evaluation considerations for commercial deployment; (b.) TREC, DUC, MUC, ETC: Pros and cons; (c.) What to measure, what to report.
4. Test data: (a.) Standard corpora aren't enough; (b.) Using representative data of sufficient scale; (c.) Where to get the data.
5. Functionality: (a.) What does the NLP component do; (b.) Turning NLP functionality into product functionality; (c.) Understanding the end user .
6. Performance and scale considerations: (a.) How fast is fast; (b.) How large is large scale; (c.) Why throwing hardware at it isn't a solution.
7. Integration: (a.) The production environment(s); (b.) Customizing the application; (c.) Ongoing maintenance and support; (d.) What 24x7 service really means.
8. Selling NLP to industry: (a.) Know your customer; (b.) Know your technology; (c.) Know your competition; (d.) Showing how your technology benefits your customer; (e.) How good salespeople go bad.
9. What industry wants from NLP technology - specific R&D areas.
10. Why industry should value and support academic NLP research, even that with no direct commercial value (and why we too often don't).
11. Concluding remarks.

Abstract

Over the past ten years, researchers in computational linguistics and information retrieval have made a number of advances in document retrieval, categorization, entity recognition and other areas. This drew the attention of venture capitalists who provided the money needed to commercialize this work. There have been some successes, but the landscape is littered with failed startups and applications that didn't live up to expectations. The value of NLP research is not based on its commercializability. But for those who seek to commercialize their research, good research alone is not enough.

The purpose of this tutorial will be to examine the role of NLP research from the perspective commercial deployment. Specifically, it will focus on issues and concerns that must be addressed to meet the needs of potential customers for NLP technology, customers who are eager for text processing and retrieval solutions, but who often are disappointed with what they find.