Text Mining Special Interest Group

Ron Behling, Bristol-Myers Squibb

Novartis Institute for Biomedical Research,
Cambridge, MA 6-8th October 2004

The PRISM Forum

At Bristol-Myers Squibb, Text Mining is considered a solution for…?

- General searching of
  - Documents associated with creating and managing clinical studies
  - Documents for Filings
  - “Knowledge desktops”
  - Integrating information from multiple areas
- Considering more focused solutions
  - Safety related information
  - Reuse of protocols in research and exploratory development (preclinical safety, MAP, etc)
  - Find information about product development and formulations

What are the Critical Business Needs that Text Mining Could Address?

- Identify and prioritize targets
  - Be aware of “all” relevant information about target
  - Associate this with target class and pathway information
  - Associate with chemical class modulator information
- Finding “all” safety information about a topic
  - Internal
  - external
- Support general reporting needs

What Questions do Users Want to Ask ?

- What is known about ‘X’?
- What reports have been published about ‘X’?
- What protocols have been run on ‘X’?
- Give me all information related to the filing for ‘X’.
- Give me all known internal and external safety information about ‘X’.
- What are all known safety issues for this class of compound?
- Give me everything ‘Y’ has said about ‘X’.
- Who is knowledgeable about ‘X’?
- ‘X’ associated with which pathways, and where is it critical?

Text Mining Applications/Components

Function	Current	Future
Manual TM Methods	Verity, Documentum, Microsoft search	
3rd Party curation services		
Categorization & Clustering	Verity	
NLP or OBIIE		
Full text access		
Recall and precision metrics		
Visualisation		

Rowan Gardner:

Do you agree with this view of Text Mining? If not Change IT!

Current Text Mining Applications in Bristol-Myers Squibb?

- Documentum
- Verity
- Microsoft Search
- internal human-curated database

What Future Applications are Planned?

- Nothing definitive

Conclusions

- We have a lot of work to do