

---

# BioCreative VI Track 3 – BEL Track

Extraction of Causal Network Information in Biological Expression Language (BEL)

Next steps & discussion

by Sumit Madan

---

Sumit Madan,  
Jens Dörpinghaus,  
Juliane Fluck

Bioinformatics Department  
Fraunhofer SCAI



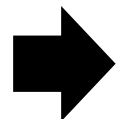
Justyna Szostak  
Julia Hoeng

Philip Morris International R&D



# Challenges of Current BEL Corpus

- One sentence is a very limited context
  - Disambiguation of named entities is context dependent
  - One sentence often does not offer sufficient context, in particular for species identification
- No negative set and few examples for some categories
- Several levels of analysis: entities, functions, relations
- Multiple / large namespaces (HGNC, MGI, EntrezGene, ChEBI, MeSH Diseases, GO Biological Processes)



Still far from using **real world data** to **tackle real world use cases**

# Outlook

- Usage of additional data generated by the curation groups
  - **Extraction of context information**
- How to increase participation?
  - Complexity of the task is an obstacle (it takes time to develop an integrative approach, adapt or create tools)
  - Only few groups might be interested in an integrative approach
  - **Coordinate with other groups to reduce parallel text mining tasks**
  - **Need to provide more supporting resources (parsers, converters, state-of-the-art recognition tools etc.)**
- Investment made so far will pay off in future editions
  - **Evaluation framework, documentation, and data sets can be reused (as for this year)**

# Summary

- BEL challenge: Usage of real world data to train systems
- Competitive evaluations: provide rigorous evaluation in a controlled environment
- New technologies have been used in the challenge
- Steps defined to tackle such complex tasks and attract more groups in the future
- Workshop raised attention to BEL in the text mining (and in the biocuration) community

---

# BioCreative VI Track 3 – BEL Track

Extraction of Causal Network Information in Biological Expression Language (BEL)

Thank you!

---

Sumit Madan,  
Jens Dörpinghaus,  
Juliane Fluck

Bioinformatics Department  
Fraunhofer SCAI



Justyna Szostak  
Julia Hoeng

Philip Morris International R&D

