

An Indian-Australian research partnership

Project Title:

**Multilingual Sentiment Analysis and Summarization
along with Ontology Development**

Project Number

IMURA0321

Monash Supervisor(s)

Prof. Mark Carman

Full names and titles

Monash Primary Contact:

mark.carman@monash.edu

Email, phone

**Monash Head of
Department:**

Prof. Kim Marriott

Full name, email

Monash Department:

Clayton School of IT

Full name

Monash ADRT:

Prof Kai-Ming Ting

Full name, email

IITB Supervisor(s)

Prof. Pushpak Bhattacharyya

Full names and titles

IITB Primary Contact:

pb@cse.iitb.ac.in, 91-9819708718

Email, phone

IITB Head of Department:

Prof. Amitabha Sanyal

Full names and titles

IITB Department:

Computer Science and Engineering

Email, phone

Research Academy Themes:**Highlight which of the Academy's Theme(s) this project will address?**

(Feel free to nominate more than one. For more information, see www.iitbmonash.org)

1. **Advanced computational engineering, simulation and manufacture**
2. Infrastructure Engineering
3. Clean Energy
4. Water
5. Nanotechnology
6. Biotechnology and Stem Cell Research

The research problem

Sentiment Analysis and Summarization are heavily machine learning dependent tasks today. However, after a stage the accuracy in these tasks taper off, unless knowledge is injected from the domain and/or from linguistics. Multilingual sentiment analysis and summarization have become extremely important problems today from the point of view of both politics and economics.

In the current proposal we would like to investigate cross lingual and monolingual sentiment detection and summarization assisted by domain specific ontology like tourism. Corpora of such domains will be mined to create a hierarchy of concepts. Such a hierarchy will then be put to use for better parameter learning in machine learning based summary generation and sentiment detection. Many low level NLP tools that will be required for the task are available as in-house capability at IIT Bombay.

Project aims

As above

Expected outcomes

The outcomes will be Natural Language Processing systems in many languages for sentiment analysis and summarization. The techniques will be language independent at the fundamental level with language specific layers built on top.

How will the project address the Goals of the above Themes?

The proposed project directly benefits the area of advanced computing of which language technology is definitely a part. The perspective is that of Artificial Intelligence or Intelligent Computation.

Capabilities and Degrees Required

The student should be proficient in programming with very good grasp of data structures and algorithms. Exposure to linguistics and/or NLP will be an advantage. Good foundation of statistics and probability is required.