

Format for submitting projects under Project Varanasi. (Do not change the serial no and headings of the items. If the headings are not relevant to your project then say so. The proposal document must not exceed 5 pages.)

1	Project type (Strike off those not applicable, refer to the policy document for project types)	(i) Technology Development or Prototype Development. (ii) Faculty Projects (Innovation and application Projects) (iii) Project (Student Nurturing)
2	Title of the Project	Machine Readable Dictionary of Bhojpuri, Magahi and Maithili
3	Duration of the project	5 -7 months
4	Total Cost	
5	Name address and phone numbers of PIs and Co-PI's	PI-- Prof. Rajeev Sangal Director, IIT(BHU) Varanasi-221005 CO-PI – (1) Dr. Swasti Mishra Deptt. of Computer Science & Engineering, IIT(BHU), Varanasi-5, Mo.no.—9389156777 CO-PI – (2) Dr. Anil Singh Deptt. of Computer Science & Engineering, IIT(BHU), Varanasi-5, Mo.no.—9460009846

6. General Description of the project:

This project consists of the task of making concept based machine readable dictionary of Bhojpuri, Magahi and Maithili. Machine readable multilingual dictionary is a part of machine translation system developed in India for Indian languages. This concept based dictionary is in synset format and it supports translation between multiple Indian languages. But in most of the Indian languages there is still no facility of machine translation. Therefore an intensive linguistic research work is required in order to achieve the desired goal.

In this project we propose to compile the machine readable dictionary for Bhojpuri, Magahi and Maithili languages. This will enhance the possibility of making machine translation tools available in these Indian languages of eastern group, which are lagging behind the other major Indian languages in this regard.

7. General Description of experience/ expertise of team on such/ similar projects:

Part of the above mentioned project is already carried out as a summer project.

8. Deliverables (The deliverables are to be described in each section. If there is no deliverable in a particular section then say the same clearly.):

(a) Prototype -nil-

(b) Process Prototype -nil-

(c) Design/ Technical Document **-yes-**

(d) Software **-yes-**

(e) Document (audio, visual, write ups web sites etc) -nil-

(f) Any other -nil-

9. Method/ Technology to reach the deliverable. (A detailed description of method or technology may be described):

It is a Machine Readable Dictionary which is concept based. This MRD is in synset format. MRD module which will be included here is a part of Machine Translation system for Indian languages developed at IIIT Hyderabad.

In this project three teams of 12 project fellows (4 for each language) with good knowledge of Linguistics along with good knowledge of specific language and Hindi will be formed. These three language groups (Bhojpuri, Magahi and Maithili) will be working in parallel to achieve the desired task.

10. Time line / mile stones for achieving the deliverables:- 5-7 months.

(Prof. Rajeev Sangal)
Principal Investigator

(Dr. Swasti Mishra)
Co-PI

(Dr. Anil Singh)
Co-PI

(Prof. K. K. Shukla)
Head, Deptt. of Computer Science & Engineering
IIT (BHU)