

# **Text Transcription of Technical Video Lectures and Creation of Searchable Video Index, Metadata and Online Quizzes**

**A project for Enhancing the Reach of the National Programme on  
Technology Enhanced Learning (NPTEL) in India with Standardized  
Metadata Creation**

**Principal Investigator: Dr. Mangala Sunder Krishnan**, Professor, Department  
of Chemistry, Indian Institute of Technology (IIT) Madras, Chennai 600036, India

**Co-Investigator: Dr. Hema A. Murthy**, Professor, Department of Computer  
Science and Engineering, Indian Institute of Technology Madras, Chennai  
600036

**Research area keywords:** Information retrieval, extraction and organization,  
multi-media search, education innovation

## **Objective:**

- **To create accurate text transcriptions of NPTEL video lectures in engineering sciences from Phase I and other metadata for video indexing and searching**
- **To use the transcribed text and the video to set up a tutored online programme of quizzes for self-evaluation by users**
- **To create a suitable database for AI algorithms in speech recognition (Indian English)**

**Source for the Project:** Video archive of 120 courses developed under the  
National Programme on Technology Enhanced Learning

**Total Funding required: Rs. 3,15,00,000 (Rupees three crores and fifteen  
lakhs)**

**Duration of the Project:** 24 months

### **Expected Outcomes (Deliverables):**

1. Accurately transcribed and certified text files with video images of all lectures from 5,000 hours of video lectures. Approximately 70,000 print pages (A4) will be made available for online access. **The text files will be certified by the faculty who developed the video courses.** This will enable viewers to browse through authenticated text contents of 5,000 hours of engineering lessons in video and search for specific topics with the help of powerful search engines.
2. A programme for online quiz and self-evaluation by students in these courses through the Internet.
3. Creation of a thesaurus of Indian pronunciation of technical terms as a suitable database for future research in speech-to-text translation using AI and other search algorithms.
4. Extensive indexing of videos for enabling search tools to search through the video.
5. The creation of text files to act as catalysts to the design and development of digital and online text books in engineering by the faculty.

### **Technical Description**

#### **Introduction:**

The National Programme on Technology Enhanced Learning (NPTEL), proposed for the first time in 1999 by Professor M. S. Ananth, Director, IIT Madras and funded by the Ministry of Human Resource Development, Government of India, has developed curriculum based course contents for more than 240 engineering courses in five disciplines at the undergraduate level. The programme has been executed jointly by all seven Indian Institutes of Technology (Bombay, Delhi, Guwahati, Kanpur, Kharagpur, Madras and Roorkee) and the Indian Institute of Science, Bangalore. More than 350 faculty