Expression of interest for Tamil Computing Projects

Tamil Virtual Academy a Government of Tamil Nadu Organization intends to develop a comprehensive and collaborative repository of Tamil Computing Research and Application Tools. To achieve this objective, expression of interest is invited from Educational and Research Institutions, reputed firms, non-profit organizations and individuals who are engaged in the field of Tamil Computing to undertake various projects in the following fields.

1. **Image & Speech Recognition for Tamil**

   (To go for an artificial intelligence, Image & Speech recognition are must. This Image Recognition software is to identify the objects, places, people, writing and actions available in the images etc., either in line drawing or pictures, photographs and the Speech Recognition or speech-to-text program is to identify the spoken words and convert them into text dynamically or machine readable text.)

2. **Tamil Syntactic Parser (like: Stanford parser for English)**

   (Syntactic parser deals with syntactic structure of a sentence. The syntactic parser is to find and provide the grammatical arrangement of words / immediate constituent in a sentence and their relationship with each other with line drawing.)
3. **POS & Semantic role of labeling (200 million words Tamil Corpus Bank)**

(To find out the meaning in a sentence, POS (Parts of Speech) and Semantic role of labelling are must.

This tool is to be assigned the Tamil grammatical categories (like: noun-பெயர்ச்சொல், verb-வினைச்சொல், adjective-பெயரனை, adverb-வினையனை, relative/adjectival participle-பெயபரச் ம், conditional participle-நிெந்தனை, verbal participle-அிடப்புரச் ம், verbal noun-பதொழில் பெயர், infinitive-சுனைபயச் ம் etc.) and the semantic role of labelling is to be provided (like: agent-எழுவொய், predicate-ெயைினை, object-பய்பெடுபெொருள், goal, result, and the verbs are to be classified like: procedural verb e.g.: கைந்து, and solution verb நீங்கும், வடியும் in medical domain) for given sentences, paragraphs or whole chapters and It should find and give the clues of the meaning of the sentence.)

4. **Sentiment Analysis for Tamil**

(Sentiment analysis software is to be identified the opinions, appraisals, emotions, or attitudes towards a topic, person or entity available in the text.)

5. **Chatbot and Question – Answer system for Tamil**

(Chatbot is a piece of software designed to perform an automated task. And a chatbot is to be made a conversation with a human being using text or voice. )

6. **Natural Language Generation tools**

(The 5 Natural Language Generation tools will be used in artificial intelligence field to generate language for various purposes.)
i. Content analysis
(Content analysis tool is used in NLP and it is to be involved to break the sentences into n-grams and noun phrases etc. to extract the themes within a collection of unstructured text documents.)

ii. Data understanding
(For Machine Learning purposes, Data Understanding tool is to be developed. Data Understanding tool is to interpret the Data and identify the patterns in the context of the sentence.)

iii. Document structuring
(Document structure tool should find the arrangement of a document into graphical constituents like chapters, sections, paragraphs, sentences, bulleted points, lists, and figures and it also to cover some features within sentences, including quotation, sayings, proverbs etc.)

iv. Sentence aggregation
(Sentence aggregation tool is to summarize the topic accurately and bring the relevant sentences.)

v. Grammatical structuring
(This tool is to assume the syntactical structure of the sentence and get the information to rewrite the sentence in a grammatically correct manner.)

7. Machine Translation (Tamil – English vice-versa)
(This Machine Translation tool is to be translated accurately Tamil/English sentences, paragraphs or whole chapters are available in modern Tamil / English vice-versa. Example: Policy Note in Government offices)

8. Any other Tamil Computing tools

Format for submission of project proposals can be downloaded from Tamil Virtual Academy website (www.tamilvu.org) Expression of interest should be sent to The Director, Tamil Virtual Academy in a sealed cover or by mail (tva@tn.gov.in) on or before 11.05.2022 along with the detailed description of the experience of the co-ordinators in the field, work done, method of accomplishing the project, funds requirements, time line etc.

The proposals will be evaluated and screened by an Expert committee.