

Hà Nội, ngày 05 tháng 09 năm 2023

THƯ KÊU GỌI ĐÓNG GÓP CHO QUỸ PHÁT TRIỂN VLSP

Từ năm 2020, Câu lạc bộ VLSP (VLSP – Vietnamese Language and Speech Processing) đã chính thức thành lập như một chi hội của Hội Tin học Việt Nam và tiếp tục các hoạt động phát triển cộng đồng VLSP vốn được khởi xướng từ năm 2005 với sự chủ trì của các nhóm nghiên cứu VLSP tại các trường viện ở Việt Nam và Nhật Bản (JAIST). Hoạt động điển hình là [chuỗi hội thảo quốc tế VLSP](#) được tổ chức thường niên từ năm 2012 đến nay.

Trong những năm gần đây, thông qua chuỗi hội thảo VLSP, CLB VLSP đã có những đóng góp rất tích cực cho cộng đồng thông qua việc cung cấp các bộ dữ liệu có chú giải miễn phí dùng trong nghiên cứu; tổ chức các cuộc thi cho các nhóm nghiên cứu NLP có được đánh giá khách quan về chất lượng các công cụ giải quyết các tác vụ khác nhau trong lĩnh vực NLP, qua đó cộng đồng nghiên cứu cũng được chia sẻ các tiến bộ khoa học công nghệ nhờ các báo cáo khoa học của các đội thi xuất sắc nhất tại hội thảo.

Năm 2023, hội thảo [VLSP 2023](#) đã được phát động với 10 cuộc thi đánh giá các công cụ xử lý văn bản và tiếng nói tiếng Việt:

1. Phân tích cú pháp tiếng Việt thành phần (*Vietnamese Constituency Parsing*): Tự động phân tích các thành phần ngữ pháp trong câu tiếng Việt
2. Dịch máy Việt-Lào (*Vietnamese-Lao Machine Translation*): Dịch tự động các bài viết cho các cặp ngôn ngữ Việt - Lào.
3. Phân tích ý kiến so sánh từ đánh giá sản phẩm tiếng Việt (*Comparative Opinion Mining from Vietnamese Product Reviews*): Phân tích ý kiến so sánh từ đánh giá sản phẩm tiếng Việt trên các nền tảng thương mại điện tử.
4. Nhận diện hàm ý văn bản pháp luật (*Legal Textual Entailment Recognition*): Xác định mối quan hệ ngữ nghĩa giữa một phát biểu được trích dẫn và một đoạn văn bản pháp luật.
5. Đọc hiểu hình ảnh chứa tiếng Việt (*Vietnamese Visual Reading Comprehension*): Trả lời câu hỏi tiếng Việt bằng cách đọc hiểu hình ảnh chứa văn bản tiếng Việt.
6. Mô hình ngôn ngữ lớn tiếng Việt (*Vietnamese Large Language Models*): Xây dựng mô hình ngôn ngữ lớn cho tiếng Việt.
7. Nhận dạng tiếng nói tiếng Việt và nhận dạng cảm xúc (*Vietnamese Automatic Speech Recognition and Speech Emotion Recognition*): Nhận dạng tiếng nói và nhận dạng cảm xúc trong câu nói tiếng Việt.
8. Tổng hợp tiếng nói tiếng Việt có cảm xúc (*Emotional Vietnamese Speech Synthesis*): Tổng hợp tiếng nói tiếng Việt với các cảm xúc khác nhau từ một văn bản đầu vào.
9. Xác minh người nói tiếng Việt có chống giả mạo (*Vietnamese Spoofing-Aware Speaker Verification*): Xác minh người nói tiếng Việt có giải quyết vấn đề giả mạo giọng nói.
10. Phát hiện phát âm sai tiếng Việt (*Vietnamese Mispronunciation Detection*): Phát hiện phát âm sai trong các câu nói tiếng Việt.

Các hoạt động đóng góp cho cộng đồng của VLSP có thể được duy trì là nhờ các hoạt động tình nguyện của từng thành viên. Để có thể tổ chức thành công hội thảo, Ban Tổ chức

chuỗi hội thảo VLSP rất mong nhận được sự ủng hộ của tất cả các doanh nghiệp, các nhóm nghiên cứu, phát triển sản phẩm liên quan tới xử lý tiếng Việt ở trong cũng như ngoài nước với các hoạt động của hội thảo thông qua một hay nhiều hình thức như:

1. Đóng góp dữ liệu đã gán nhãn để làm nguồn cho việc xây dựng dữ liệu huấn luyện phục vụ 10 cuộc thi.
2. Hỗ trợ kinh phí cho việc xây dựng dữ liệu có gán nhãn chuẩn cho mục đích huấn luyện và đánh giá các hệ thống dự thi.
3. Hỗ trợ nhân lực hoặc kinh phí cho nhân lực tham gia đánh giá các hệ thống dự thi.
4. Ủng hộ kinh phí cho các giải thưởng dành cho các đội đạt thứ hạng cao.

Các dữ liệu được xây dựng phục vụ hội thảo VLSP sẽ được chia sẻ chung cho toàn cộng đồng phục vụ việc nghiên cứu, phát triển ứng dụng xử lý ngôn ngữ tự nhiên. Các kết quả thực nghiệm tốt nhất đạt được trong mỗi nội dung thi cũng là căn cứ cho các nhóm nghiên cứu triển khai lựa chọn giải pháp cho các bài toán thực tế. Vì vậy, việc góp công sức, kinh phí của các cá nhân, các doanh nghiệp, các trường viện thực sự có ý nghĩa với sự phát triển của cộng đồng.

Các gói tài trợ tiền mặt có thể ở các mức dưới 30 triệu đồng (tài trợ đồng), từ 30 đến dưới 50 triệu đồng (tài trợ bạc), 50 triệu đồng đến dưới 100 triệu đồng (tài trợ vàng), 100 triệu đồng trở lên (tài trợ kim cương).

Quyền lợi chung: Các đơn vị đồng hành sẽ được in logo trên trang web và trên banner của hội thảo VLSP 2023; được sử dụng tất cả các dữ liệu huấn luyện được xây dựng cho các cuộc thi tại hội thảo; được các chuyên gia tư vấn về kỹ thuật phát triển các hệ thống nói trên nếu có nhu cầu.

Quyền lợi riêng: được quảng bá trên trang web Câu lạc bộ VLSP (vlsp.org.vn) trong vòng 12 tháng (mức đồng), 24 tháng (mức bạc), 36 tháng (mức vàng và kim cương).

Mọi sự ủng hộ tài chính của các đơn vị xin thông qua đơn vị tổ chức là Câu lạc bộ VLSP, chi hội của Hội Tin học Việt Nam, đại diện là TS Nguyễn Thị Minh Huyền, Chủ tịch câu lạc bộ. Địa chỉ email: huyenntm@hus.edu.vn.

Thông tin tài khoản tiếp nhận tài trợ:

Tên chủ tài khoản: Hội Tin học Việt Nam

Tài khoản: 003 70406 0000 979

Tại ngân hàng TMCP Quốc tế Việt Nam, SGD 62-64 Lý Thường Kiệt, Hà Nội.

Xin trân trọng cảm ơn Quý công ty!

CLB XỬ LÝ NGÔN NGỮ VÀ TIẾNG NÓI TIẾNG VIỆT

(TUQ TỔNG THƯ KÍ HỘI TIN HỌC VIỆT NAM)

CHỦ TỊCH



Nguyễn Thị Minh Huyền