

Hà Nội, ngày 20 tháng 11 năm 2017

THƯ MỜI TÀI TRỢ

Kính gửi: Quý Công ty

Từ năm 2012 tới nay, nhóm nghiên cứu xử lý ngôn ngữ và tiếng nói tiếng Việt (VLSP – Vietnamese Language and Speech Processing) của Viện Công nghệ thông tin, Viện Hàn lâm KH&CN VN đã phối hợp cùng các nhóm nghiên cứu VLSP ở nhiều trường đại học, trung tâm nghiên cứu trong cả nước đứng ra tổ chức [chuỗi hội thảo quốc tế VLSP](#), kết hợp với các hội nghị khoa học quốc tế tổ chức tại Việt Nam. Các hoạt động của hội thảo bao gồm báo cáo khoa học, tọa đàm về phát triển cộng đồng VLSP, trình diễn sản phẩm VLSP và tổ chức các cuộc thi đánh giá hiệu quả của các công cụ xử lý tiếng Việt.

Năm 2018, hội thảo [VLSP](#) sẽ được tổ chức trong khoảng 18-24/03/2018 cùng với hội nghị quốc tế về Ngôn ngữ học tính toán và xử lý văn bản thông minh CICLing 2018 (<https://www.cicling.org/2018>). Hoạt động chuẩn bị cho hội thảo được bắt đầu từ cuối tháng 08/2017, với sự phát động 04 cuộc thi đánh giá các công cụ xử lý văn bản và tiếng nói tiếng Việt. Ngày hội thảo sẽ là buổi công bố kết quả các cuộc thi và các báo cáo kỹ thuật của các đội dự thi, cùng với hoạt động truyền thông là tọa đàm về phát triển VLSP.

Bốn cuộc thi xử lý tiếng Việt được phát động cụ thể là:

1. Nhận dạng thực thể có tên (named entity recognition - NER) trong văn bản báo chí và văn bản trên mạng xã hội (tên người, tên tổ chức, tên địa điểm).
2. Phân tích tình cảm (sentiment analysis - SA): phân cực quan điểm thể hiện trong các bình luận đánh giá trên các trang web du lịch (khách sạn) và ẩm thực (nhà hàng), theo các tiêu chí đánh giá được phân loại trước.
3. Nhận dạng tiếng nói (automatic speech recognition – ASR): nhận dạng tiếng nói dạng tin tức với các giọng nói 3 miền Bắc, Trung, Nam.
4. Tổng hợp tiếng nói (text to speech, TTS): tổng hợp tiếng nói với giọng nói theo một hay nhiều miền (Bắc, Trung, Nam).

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 và đánh giá các công cụ NER và SA theo 2 định hướng nêu trên;
2. Ủng hộ kinh phí cho việc xây dựng dữ liệu chuẩn cho huấn luyện, đánh giá các công cụ NER, SA, ASR và TTS.

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. Vì vậy, việc góp công sức, kinh phí của các cá nhân, các doanh nghiệp, các trường viện là vô cùng có ý nghĩa với sự phát triển của cộng đồng.

Các gói tài trợ cụ thể có thể ở các mức dưới 30 triệu đồng với doanh nghiệp nhỏ (mức 1), 30-50 triệu đồng với các doanh nghiệp lớn hơn (mức 2), trên 50 triệu đồng với các doanh nghiệp lớn (mức 3).

Quyền lợi chung: Các đơn vị đồng hành sẽ được in logo trên trang web và trên banner của hội thảo VLSP 2018. 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 hệ thống nhận dạng thực thể có tên, phân tích tình cảm, nhận dạng tiếng nói, tổng hợp tiếng nói tiếng Việt nếu có nhu cầu.

Quyền lợi riêng: được quảng bá trên trang web cộng đồng VLSP vlsp.org.vn trong vòng 24 tháng (mức 2), 36 tháng (mức 3).

Mọi sự ủng hộ tài chính của các đơn vị xin thông qua đơn vị đại diện nhận tài trợ là Viện Công nghệ thông tin, Viện Hàn lâm Khoa học Công nghệ Việt Nam, người trực tiếp phụ trách là PGS. TS. Lương Chi Mai, Chủ tịch Hội đồng Khoa học Viện. Địa chỉ email: lcmai@ioit.ac.vn.

Xin trân trọng cảm ơn Quý công ty!

Thay mặt BTC chuỗi hội thảo VLSP

PGS. TS. Lương Chi Mai