

Số: 3327/QĐ-BKHCN

Hà Nội, ngày 03 tháng 11 năm 2016

QUYẾT ĐỊNH

Về việc phê duyệt danh mục nhiệm vụ khoa học và công nghệ cấp quốc gia đặt hàng để tuyển chọn thực hiện trong kế hoạch năm 2017

**BỘ TRƯỞNG
BỘ KHOA HỌC VÀ CÔNG NGHỆ**

Căn cứ Nghị định số 08/2014/NĐ-CP ngày 27 tháng 01 năm 2014 của Chính phủ quy định chi tiết và hướng dẫn thi hành một số điều của Luật Khoa học và Công nghệ;

Căn cứ Nghị định số 20/2013/NĐ-CP ngày 26 tháng 02 năm 2013 của Chính phủ quy định chức năng, nhiệm vụ, quyền hạn và cơ cấu tổ chức của Bộ Khoa học và Công nghệ;

Căn cứ Thông tư số 07/2014/TT-BKHCN ngày 26 tháng 5 năm 2014 của Bộ trưởng Bộ Khoa học và Công nghệ quy định trình tự, thủ tục xác định nhiệm vụ khoa học và công nghệ cấp quốc gia sử dụng ngân sách nhà nước;

Trên cơ sở kết quả làm việc và kiến nghị của Hội đồng tư vấn xác định nhiệm vụ khoa học và công nghệ độc lập cấp quốc gia năm 2016;

Theo đề nghị của Vụ trưởng Vụ Kế hoạch - Tổng hợp, Vụ trưởng Vụ Công nghệ cao,

QUYẾT ĐỊNH:

Điều 1. Phê duyệt danh mục nhiệm vụ khoa học và công nghệ cấp quốc gia đặt hàng “Nghiên cứu xây dựng hệ dịch tự động văn bản tiếng Việt ra tiếng nói tiếng Mường, hướng đến áp dụng cho các ngôn ngữ dân tộc thiểu số chưa có chữ viết ở Việt Nam” để tuyển chọn thực hiện trong kế hoạch năm 2017 (Phụ lục kèm theo).

Điều 2. Giao Vụ trưởng Vụ Công nghệ cao và Vụ trưởng Vụ Kế hoạch - Tổng hợp:

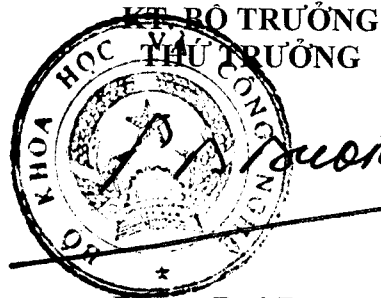
- Thông báo danh mục nhiệm vụ nêu tại Điều 1 trên cổng thông tin điện tử của Bộ Khoa học và Công nghệ theo quy định để các tổ chức, cá nhân biết và đăng ký tham gia tuyển chọn.

- Tổ chức Hội đồng khoa học và công nghệ đánh giá hồ sơ nhiệm vụ đăng ký tham gia tuyển chọn theo quy định hiện hành và báo cáo Bộ trưởng Bộ Khoa học và Công nghệ về kết quả thực hiện.

Điều 3. Vụ trưởng Vụ Công nghệ cao, Vụ trưởng Vụ Kế hoạch - Tổng hợp, Giám đốc Văn phòng các Chương trình trọng điểm cấp nhà nước và Thủ trưởng các đơn vị có liên quan chịu trách nhiệm thi hành Quyết định này./.

Nơi nhận:

- Như Điều 3;
- Lưu: VT, Vụ KHTH.



Phạm Đại Dương



**DANH MỤC NHIỆM VỤ KHOA HỌC VÀ CÔNG NGHỆ CẤP QUỐC GIA
ĐẠI HẠNG ĐỂ TUYỂN CHỌN THỰC HIỆN TRONG KẾ HOẠCH NĂM 2017**

(Kèm theo Quyết định số 27/QĐ-BKHCN ngày 03 tháng 11 năm 2016 của Bộ trưởng Bộ Khoa học và Công nghệ)

TT	Tên nhiệm vụ	Định hướng mục tiêu	Yêu cầu đối với kết quả	Phương thức tổ chức thực hiện
1	Nghiên cứu xây dựng hệ dịch tự động văn bản tiếng Việt ra tiếng nói tiếng Mường, hướng đến áp dụng cho các ngôn ngữ dân tộc thiểu số chưa có chữ viết ở Việt Nam	<p>1. Nghiên cứu, làm chủ công nghệ dịch máy từ văn bản tiếng Việt ra tiếng nói tiếng dân tộc thiểu số không có chữ viết ở Việt Nam.</p> <p>2. Xây dựng được cơ sở dữ liệu song ngữ văn bản tiếng Việt - tiếng nói tiếng Mường.</p> <p>3. Xây dựng hệ thống dịch máy từ văn bản tiếng Việt ra tiếng nói tiếng Mường, với đặc điểm chưa có chữ viết và sự khác biệt về phương/thổ ngữ.</p> <p>4. Phát triển một số sản phẩm sử dụng hệ thống dịch tự động Việt-Mường.</p>	<p>1. Tài liệu:</p> <ul style="list-style-type: none"> - Bộ tài liệu công nghệ dịch từ văn bản tiếng Việt ra tiếng nói ngôn ngữ không có chữ viết: + Đầu vào: văn bản tiếng Việt; + Đầu ra: tiếng nói tiếng dân tộc không có chữ viết; + Độc lập ngôn ngữ: có khả năng áp dụng với các ngôn ngữ thiểu số không chữ viết khác. - Bộ tài liệu thiết kế phần mềm và hướng dẫn sử dụng bộ công cụ phục vụ thu thập và xử lý dữ liệu tiếng nói tiếng dân tộc thiểu số chưa có chữ viết ở Việt Nam; - Bộ tài liệu quy trình xây dựng và mô tả cơ sở dữ liệu song ngữ văn bản tiếng Việt-tiếng nói tiếng Mường; - Bộ tài liệu thiết kế; hướng dẫn sử dụng, đánh giá chất lượng và đánh giá kết quả hệ thống dịch máy từ văn bản tiếng Việt sang tiếng nói tiếng Mường có định hướng lĩnh vực; - Bộ tài liệu thiết kế phần mềm, hướng dẫn cài đặt và sử dụng phần mềm trên thiết bị cầm tay. <p>2. Thiết bị:</p> <ul style="list-style-type: none"> - 01 Hệ thống dịch máy từ văn bản tiếng Việt sang tiếng nói tiếng Mường có định hướng lĩnh vực; + Ứng dụng cho tối thiểu 02 tỉnh có tiếng nói tiếng 	Tuyển chọn



TT	Tên nhiệm vụ	Định hướng mục tiêu	Yêu cầu đối với kết quả	Phương thức tổ chức thực hiện
			<p>Mường.</p> <ul style="list-style-type: none"> + Chất lượng dịch tự động trong lĩnh vực lựa chọn được đánh giá bởi người dùng, theo phương pháp được đề xuất bởi Hiệp hội dữ liệu ngôn ngữ học (LDC) và Viện Quốc gia về Chuẩn và Công nghệ Hoa Kỳ (NIST), gồm có: <ul style="list-style-type: none"> • Tính trôi chảy (fluency): Đạt mức 4/5; • Tính đầy đủ (adequacy): Đạt mức 4/5. + Chất lượng tiếng Mường đầu ra được đánh giá theo tiêu chuẩn do ITU-TP.85 quy định, đạt độ tự nhiên $\geq 3.5/5$ với thang điểm MOS, đạt độ dễ hiểu (Intelligibility) $\geq 95\%$. + Cho phép sử dụng thông qua: <ul style="list-style-type: none"> • Trình duyệt Web; • Ứng dụng trên thiết bị di động. + Thời gian phản hồi của hệ thống dịch: <ul style="list-style-type: none"> • ≤ 2 giây với câu đầu vào tiếng Việt dưới 20 từ đơn. • ≤ 10 giây với văn bản dưới 200 từ đơn. + Tính tiện dụng trong tương tác người - máy của sản phẩm được đánh giá tuân theo chuẩn ISO 9241. - 01 Bộ công cụ phục vụ thu thập và xử lý dữ liệu tiếng nói tiếng dân tộc thiểu số chưa có chữ viết ở Việt Nam: + Không phụ thuộc vào kiểu ngôn ngữ. + Hỗ trợ thu âm, phân tách câu, gán nhãn, quản lý và đóng gói dữ liệu theo chuẩn OLAC. + Hỗ trợ phiên âm tự động cho ngôn ngữ không có chữ viết, với độ chính xác phiên âm mức âm vị $\geq 50\%$ 	

TT	Tên nhiệm vụ	Định hướng mục tiêu	Yêu cầu đối với kết quả	Phương thức tổ chức thực hiện
			<p>cho tiếng Mường.</p> <ul style="list-style-type: none"> + Hỗ trợ trích rút, phân tích và thống kê tự động các đặc trưng âm học. - 01 cơ sở dữ liệu (CSDL) song ngữ văn bản tiếng Việt - tiếng nói tiếng Mường: + CSDL câu song ngữ dung lượng ≥ 20.000 cặp câu văn tiếng Việt-Mường thuộc 02 tỉnh và CSDL từ điển Việt-Mường có dung lượng ≥ 5.000 mục từ tiếng Việt; + Được đóng gói theo chuẩn OLAC. - 01 Phần mềm trên thiết bị cầm tay cho phép tra cứu các từ và câu văn bản tiếng Việt ra tiếng nói tiếng Mường: + Gồm ít nhất: 1.000 từ thuần Mường và 100 câu cơ bản trong giao tiếp; + Ứng dụng cho tối thiểu 02 tỉnh có tiếng nói tiếng Mường; + Có cấu trúc mô-đun; + Có khả năng bổ sung thêm các từ và câu; + Tính tiện dụng trong tương tác người - máy của sản phẩm được đánh giá tuân theo chuẩn ISO 9241. - Sản phẩm phải được thử nghiệm, đánh giá và kiểm định bởi cơ quan có thẩm quyền. 	



