

به نام خدا



وزارت علوم، تحقیقات، و فناوری

پژوهشگاه علوم و فناوری اطلاعات ایران

(ایرانداک)

خلاصه گزارش طرح پژوهشی سازمانی

عنوان طرح پژوهشی

امکان‌سنجی ساخت پیکره زبانی از داده‌های علمی پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)

تاریخ شروع طرح پژوهشی	۹۷/۴/۲۳	تاریخ پایان طرح پژوهشی	۹۸/۷/۹
-----------------------	---------	------------------------	--------

نام و نام خانوادگی مجری طرح پژوهشی	الهام علایی ابوذر
------------------------------------	-------------------

نام و نام خانوادگی همکاران طرح پژوهشی	
---------------------------------------	--

درباره طرح پژوهشی

پیکره‌های زبانی و واکاوی پیکره‌بنیاد زبان در زبان‌شناسی پیکره‌ای نقشی کلیدی دارند. ساخت پیکره‌ها کاری پیچیده، زمان‌بر، دارای گام‌های گوناگون، و میان‌رشته است. بنابراین برای کارایی و اثربخشی در ساخت آن‌ها نیاز است که پیش از آغاز کار، امکان‌سنجی انجام شود. امکان‌سنجی؛ گام‌ها، هزینه‌ها، نیروی انسانی، حقوق مادی و معنوی، و مانند آن‌ها را برای یک پروژه، بررسی و به مدیریت پروژه کمک می‌کند تا آن را با آمادگی و آینده‌نگری بیشتری به سرانجام برساند. از آنجایی که مدلی برای امکان‌سنجی ساخت پیکره متنی همچون گونه‌ای از پیکره‌ها در پژوهش‌های پیشین نبود، این پژوهش در پی ساخت چنین مدلی انجام شد. برای این کار، پس از بررسی پژوهش‌های پیشین، شاخص‌های کلی ساخت پیکره‌های زبانی استخراج شد. سپس، فرآیند ساخت پیکره به تفصیل مورد بررسی قرار گرفت. پس از بررسی پژوهش‌های پیشین در زمینه امکان‌سنجی، بر پایه اطلاعات به دست آمده از مطالعات امکان-سنجی، بررسی شاخص‌های ساخت پیکره و فرآیند ساخت پیکره، ابعاد و مولفه‌های امکان‌سنجی ساخت پیکره استخراج شد و مدلی کلی برای ساخت پیکره پیشنهاد شد. سپس با روش دلفی و در دو دور اعتباریابی، مدل پایانی به دست آمد. این مدل دارای هفت بُعد فنی، اقتصادی، زمان‌بندی، قانونی حقوقی، عملیاتی، تأمین مالی، و بازاریابی و روی هم رفته ۳۳ مؤلفه است. این مدل و کاربرد آن در پروژه‌های ساخت پیکره متنی می‌تواند کمک بزرگی برای پژوهشگران و سازمان‌ها باشد. سپس در آخرین مرحله از پژوهش، مدل امکان‌سنجی ساخت پیکره در پنلی متشکل از متخصصان موضوعی و اجرایی در ایرانداک به شور گذاشته و اعتباریابی شد و مدل پایانی برای ساخت پیکره از داده‌های علمی ایرانداک به دست آمد و در پایان، امکان‌سنجی ساخت پیکره از داده‌های ایرانداک نیز انجام شد و چگونگی پرداختن به ابعاد و مولفه‌های مدل مشخص شد.

روش پژوهش

- بررسی پژوهش‌های پیشین حوزه ساخت پیکره
- استخراج شاخص‌های کلی ساخت پیکره‌های زبانی
- بررسی فرآیند ساخت پیکره
- بررسی پژوهش‌های پیشین در زمینه امکان‌سنجی
- بررسی شاخص‌های ساخت پیکره و فرآیند ساخت پیکره
- استخراج ابعاد و مولفه‌های امکان‌سنجی ساخت پیکره
- پیشنهاد مدل امکان‌سنجی برای ساخت پیکره
- انجام روش دلفی و دو دور اعتباریابی
- ارائه مدل پایانی امکان‌سنجی ساخت پیکره
- به شور گذاشتن و اعتباریابی مدل امکان‌سنجی ساخت پیکره در ایرانداک
- ارائه مدل پایانی برای ساخت پیکره از داده‌های علمی ایرانداک
- انجام امکان‌سنجی ساخت پیکره از داده‌های ایرانداک

یافته‌های طرح پژوهشی

در این پژوهش، شاخص‌های کلی ساخت پیکره زبانی از بررسی پژوهش‌های مربوطه استخراج شد. فرآیندهای ساخت پیکره‌های متنی و نوشتاری مشخص شد. مدل امکان‌سنجی ساخت پیکره از طریق بررسی پژوهش‌های زمینه امکان‌سنجی و ساخت پیکره ارائه شد. پس از انجام روش دلفی و دو دور اعتباریابی، مدل پایانی امکان‌سنجی ساخت پیکره پیشنهاد شد. در نهایت مدل پایانی در ایرانداک به شور گذاشته شد و به امکان‌سنجی ساخت پیکره از داده‌های علمی ایرانداک پرداخته شد و در نهایت مشخص شد در صورت تصمیم‌گیری برای ساخت پیکره از داده‌های ایرانداک، چه ابعاد، مولفه‌ها و چه میزان هزینه برای ساخت پیکره از داده‌های علمی ایرانداک باید مدنظر قرار گیرد.

نتیجه‌گیری و پیشنهادها

پیشنهادات:

در ساخت پیکره‌های بزرگ برای زبان فارسی (مانند پیکره علمی از داده‌های ایرانداک)، پیشنهاد می‌شود که بر پایه ابعاد و مولفه‌های امکان‌سنجی، نخست پیکره‌ای کوچک ساخته شود؛ سپس اندازه داده‌ها افزایش یابد. می‌توان پروژه ساخت پیکره را ابتدا از مجله منتسب به ایرانداک (پژوهش‌نامه پردازش و مدیریت اطلاعات) شروع کرد. پژوهش‌های کاربردی برای ساخت یا بهبود عملکرد ابزارهایی مانند خطایاب، تشخیص‌دهنده گروه‌های نحوی، پارسر، ابزار نرمال‌سازی متن و.... و انواع حاشیه‌نویسی گفتمانی، کاربردی، سبکی، نحوی و.... برای زبان فارسی از دیگر کارهایی است که نیاز به انجام آن‌ها هست. این پژوهش‌ها می‌توانند هم‌زمان با آغاز پروژه ساخت پیکره انجام پذیرند.

پیشنهاد می‌شود این مدل با روش اقدام‌پژوهی در یک پروژه ساخت پیکره متنی به کار رود تا برای آن، اعتبار تجربی نیز فراهم شود. بررسی دقیق‌تر برخی از مؤلفه‌های مدل نیز می‌تواند جداگانه موضوع پژوهش باشند. برای نمونه «الزام‌های قانونی و حقوقی در ساخت پیکره» در بُعد قانونی حقوقی یا «پذیرش و آمادگی سازمان در کاربرد پیکره» در بُعد عملیاتی، می‌توانند در پژوهش‌هایی جداگانه بررسی و راهکارها و مدل‌های آن‌ها بررسی و تبیین شوند.