

عنوان نشست: مدیریت دانش و اطلاعات - ابزارهای فنی	کد نشست: D1
زمان نشست: چهارشنبه ۱۱ الی ۱۳	

استخراج جملات عمومی از ویکی پدیای فارسی: تلاشی در جهت ساخت آنتولوژی دانش عرفی

مهدی مرادی^۱، بهرام وزیرتژاد^۲، پروانه خسروی زاده^۳ و هادی عبدی قویدل^۴
^۱دانشگاه صنعتی شریف، mehdi_moradi@mehr.sharif.ir
^۲دانشگاه صنعتی شریف، bahram@sharif.ir
^۳دانشگاه صنعتی شریف، khosravizadeh@sharif.ir
^۴دانشگاه صنعتی شریف، hadi_abdighavidel@mehr.sharif.ir

چکیده - متون موجود در اینترنت حاوی انواع مختلفی از دانش هستند که با استفاده از تکنیک های مهندسی دانش می توان این نوع دانش ها را که در حوزه ی وب معنایی کاربردهای متنوع و مختلفی دارند استخراج نمود. دانش عرفی یکی از کاربردی ترین دانش ها در حوزه محاسباتی است که اغلب در قالب جملاتی عمومی مانند "پستانداران بچه زا هستند." بیان می شود. ویژگی بارز این نوع جملات بر خلاف جملات غیر عمومی که اشاره به شخص، موجود یا شی خاص دارد بیان قاعده ای کلی راجع به انواع یا رخدادها است. هدف این تحقیق ارائه ی روشی خودکار جهت استخراج این نوع جملات از ویکی پدیای فارسی و استفاده از آن در ساخت آنتولوژی های دانش عرفی برای زبان فارسی است.

کلمات کلیدی - استخراج جملات عمومی، دانش عرفی، آنتولوژی، تشخیص الگو.