

عنوان مقاله:

روشی جهت استخراج اطلاعات از وبسایت های خبری فارسی با استفاده OWL

محل انتشار:

کنگره ملی تحقیقات بنیادین در مهندسی کامپیوتر و فن اوری اطلاعات (سال: 1398)

تعداد صفحات اصل مقاله: 9

نویسندگان:

ویدا مظاهری - گروه مهندسی کامپیوتر، واحد تهران غرب، دانشگاه آزاد اسلامی، تهران، ایران

چنگیز دل آرا - گروه مهندسی کامپیوتر، واحد تهران غرب، دانشگاه آزاد اسلامی، تهران، ایران.

خلاصه مقاله:

امروزه، وب برای بسیاری از افراد یک ابزار ضروری است. شبکه جهانی وب یک رسانه قدرتمند برای برقراری ارتباط بین رایانه ها و دسترسی به اسناد آنلاین است که به ازای هر واژه جستجو، میلیون ها صفحه وب را برای کاربر فراهم می کند. اما این ابزار برای یافتن یا سازماندهی اطلاعات مناسب نیست. ابزارهایی مانند موتورهای جستجو به کاربران در یافتن اطلاعات کمک می کنند. با توجه به اینکه منابع اطلاعاتی موجود در وب، با سرعت زیادی در حال افزایش است، استخراج نتایج مطلوب و مورد نیاز در کمترین زمان ممکن بسیار دشوار خواهد بود. بنابراین در سال های اخیر استفاده از ابزارهای خودکار جهت استخراج اطلاعات مورد نیاز از صفحات وب، بیش از پیش مورد توجه قرار گرفته است. در این مقاله، با ایجاد آنتولوژی دامنه سایت های خبری ورزشی و استفاده از مبدل OWL، روشی جهت بهبود استخراج اخبار مربوطه از دیگر متون خبری ارائه گردیده است که نتایج حاصل از آن نشان می دهد با استفاده از روش مذکور، اخبار مورد نظر با دقت بالایی به کاربر نمای داده می شود.

کلمات کلیدی:

وب کاوی، متون خبری، آنتولوژی OWL

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/924639>

