

بررسی ویژگی‌های شبکه‌ استنادی در محیط وب: مطالعه کیفی در حوزه‌های علوم و علوم اجتماعی

دکتر کیوان کوشا

عضو هیات علمی دانشگاه تهران

چکیده

هدف: تحقیقات نسبتاً زیادی درباره ویژگی‌های منابع استنادی و مورد استناد مقالات علمی مبتنی بر داده‌های پایگاه آی.اس.آی صورت گرفته است، با این حال دانش کمی در باب خصوصیات شبکه استنادی در محیط وب در دسترس است. هدف از تحقیق حاضر بررسی پنج ویژگی برگرفته از منابع استنادی در محیط وب شامل نوع قالب فایل، حوزه اینترنتی، زبان، تاریخ انتشار و فرایبندی بودن است تا بدین ترتیب بتوان خصوصیات نشر آثار علمی در محیط وب را بررسی و شناسایی کرد.

روش: تحقیق حاضر ۵۴۵ منبع استنادی ویی نمونه‌گیری شده در چهار رشته از علوم و چهار رشته از علوم اجتماعی حاوی ۱۶۵۰ مقاله علمی منتشر شده در ۱۰۸ مجله الکترونیکی (تمام متن داوری شده) را مورد استناد قرار داده و از طریق روش تحلیل محتوا، بررسی شده است.

یافته‌ها: نتایج تحقیق نمایانگر آن است که حدود ۶۱ درصد از منابع استنادی (یا استنادکننده) در محیط وب در قالب فایل‌های پی.دی.اف، ۳۰ درصد از حوزه‌های اینترنتی دانشگاهی edu و ac، ۱۵ درصد به زبان‌های غیرانگلیسی و ۷۱ درصد از استنادها به صورت غیرفرامتنی‌اند. به عبارت دیگر، در رشته‌های مورد مطالعه، شبکه استنادی در محیط وب تحت تاثیر استندهای متنی گنجانده شده در آثار پی.دی.اف به زبان انگلیسی قرار گرفته است. نتایج به دست آمده می‌تواند به‌طور کاربردی در طراحی و توسعه ابزارهای جستجو و بازیابی اطلاعات علمی به‌ویژه نمایه‌های استنادی خودکار یا خود محور.

واژه‌های کلیدی: وب‌سنجی. شبکه استنادی. وب. ویژگی‌های نشر آثار علمی. تحلیل محتوا.

۱. مقدمه

هم اکنون اینترنت به عنوان یک منبع اطلاعاتی مهم و کاربردی در بسیاری از زمینه‌های علمی، تجاری و تفریحی و غیره مورد استفاده قرار می‌گیرد. افزایش روز افزون انواع منابع اطلاعاتی در اینترنت به عنوان یکی از مهمترین چالش‌های علم اطلاع‌رسانی به منظور بازیابی اطلاعات مرتبط به‌شمار می‌آید (بادم، هولثمن، کورتنی^۱، ۱۹۹۹). در حال حاضر، از یک سو با افزایش سریع تعداد مولفان، مجلات و سازمان‌هایی که تمایل به انتشار آزاد یافته‌ها و دستاوردهای تحقیقاتی خود در محیط وب دارند، مواجه هستیم (سوان و بران، ۲۰۰۳؛ سوان و بران) و از سوی دیگر مشکلات بازیابی اطلاعات در ابزارهای کاوش اینترنت همچنان باقی است. (موشوویتچ و کاواگوچی^۲، ۲۰۰۵). از این رو، از سال‌های نخست ظهور شبکه جهانی وب، محققان اطلاع‌رسانی در جستجوی روشی نوین برای رویارویی با مشکل بازیابی اطلاعات مرتبط در این محیط بودند.

اگرچه، تلاش‌هایی توسط کتابداران به منظور فهرست‌نویسی موضوعی منابع اطلاعاتی علمی در محیط وب صورت گرفته است (اینفوماین، ۲۰۰۷)، اما این روش به دلیل انتشار روزافزون منابع اطلاعاتی علمی در وب بسیار زمان‌بر و هزینه‌بر بوده و از جامعیت و روزآمدی لازم برخوردار نمی‌باشد. به این ترتیب، برخی از محققان مبتنی بر نظرات خلاقانه گارفیلد^۳ در رابطه با نمایه‌سازی استنادی (گارفیلد، ۱۹۵۵) به فکر طراحی و توسعه نمایه استنادی خودکار/ خودم‌محور^۴ در محیط وب بودند تا به این ترتیب با صرف هزینه، زمان و نیروی انسانی بسیار کمتری نسبت به فرایند نمایه‌سازی موضوعی، بتوان به ذخیره و بازیابی منابع اطلاعاتی در قالب شبکه استنادی میان آثار علمی مبادرت ورزید. برای مثال، گمرون^۵ در سال ۱۹۹۷ روزی را متصور شده است که در آن تمامی اطلاعات کتاب‌شناختی و داده‌های استنادی مرتبط با آثار علمی را بتوان مبتنی بر شبکه میان منابع استنادی و مورد استناد به یکدیگر پیوند داد و بازیابی کرد (گمرون، ۱۹۹۷). پس از مدت اندکی، رویای گمرون محقق شد، به طوری که امروزه ما شاهد ظهور و توسعه

پایگاه‌های استنادی تحت وب نظیر سایت سیر^۶، سایت بیس^۷ و اخیراً گوگل اسکالار^۸ هستیم. این پایگاه‌های استنادی ماشینی توانسته‌اند مبتنی بر الگوریتم‌های از پیش تعریف شده، میان آثار علمی، شبکه استنادی برقرار کنند و بدین ترتیب به کاربران امکان ردگیری استنادی میان منابع منتشر شده در محیط وب را می‌دهند. با این حال، روش گردآوری داده‌های استنادی در نمایه‌های استنادی خودکار با آنچه در نمایه‌های استنادی سنتی نظیر آی.اس.آی^۹ صورت می‌گیرد، متفاوت است. در نمایه‌های استنادی سنتی نظیر آی.اس.آی که هم اکنون تحت عنوان پایگاه وب آو ساینس^{۱۰} شناخته می‌شوند، اطلاعات استنادی محدود به طیفی از مجلات گزینش شده با اثرگذاری استنادی بالا^{۱۱} هستند که بر مبنای نظام ارزش گذاری کیفی از پیش تعریف شده، انتخاب و نمایه‌سازی می‌شوند (فرایند انتخاب مجلات در آی.اس.آی^{۱۲}، ۲۰۰۶). حال آنکه در نظام‌های نمایه‌سازی استنادی خودکار یا خودمحور موجود در محیط وب، طیف گسترده‌تری از منابع اطلاعاتی داوری شده یا داوری نشده نمایه می‌شوند و الگوریتم نمایه‌سازی در این گونه پایگاه‌ها مبتنی بر ویژگی‌ها و خصوصیات غالب نشر آثار علمی در محیط وب نمی‌باشد (بول لاکر^{۱۳} و دیگران، ۱۹۸۸؛ لارنس و دیگران، ۱۹۹۹؛ هیتچکوک^{۱۴} و دیگران، ۲۰۰۲؛ گوگل اسکالار، ۲۰۰۶). برای مثال، هنوز مشخص نیست که چه میزان از منابع اطلاعاتی علمی در قالب پی.دی.اف^{۱۵} یا به زبان انگلیسی بوده یا بر گرفته از سایت‌های دانشگاهی هستند. بنابراین، بررسی ویژگی‌های ساختاری یا محتوای شبکه استنادی در محیط وب می‌تواند به عنوان شاخصی مهم در الگوریتم نمایه‌سازی منابع اطلاعاتی علمی توسط پایگاه‌های استنادی در محیط وب مورد استفاده قرار گیرد. به عبارت دیگر، انجام مطالعات کیفی روی ماهیت منابع استنادی در محیط وب می‌تواند به شناسایی عناصری منجر گردد که بیشترین نقش را در شبکه استنادی در محیط وب دارند.

با وجود این، در رابطه با بررسی ویژگی‌های ساختاری یا محتوای شبکه استنادی در محیط وب مطالعات کمی صورت گرفته است و مطالعات پیشین نیز طیف بسیار محدودی از منابع و رشته‌های علمی را پوشش می‌دهند. از این رو،

هدف اصلی مطالعه حاضر بررسی پنج ویژگی شامل نوع قالب فایل^{۱۶}، حوزه اینترنتی^{۱۷}، زبان، تاریخ انتشار و فرایبندی بودن^{۱۸} منابع استنادی^{۱۹} (یا استنادکننده) در چهار رشته از علوم و چهار رشته از علوم اجتماعی است. نتایج تحقیق حاضر می‌تواند به‌طور کاربردی در تعیین الگوریتم نمایه‌سازی ابزارهای کاوش علمی نظیر نمایه‌های استنادی تحت وب مورد استفاده قرار گیرد. برای مثال، اینکه اغلب منابع علمی در وب که مقالات مجلات داوری شده در حوزه‌های علوم و علوم اجتماعی را مورد استناد قرار داده‌اند، در چه قالب فایلی منتشر می‌شوند (نظیر HTML، PDF، WORD، Postscript)، چه میزان از آنها از حوزه‌های اینترنتی دانشگاهی (نظیر edu, ac) هستند، زبان‌های غیرانگلیسی و استندهای فرامتنی/متنی^{۲۰} چه نقشی را در شبکه استنادی در محیط وب دارند، از جمله مسائلی هستند که تحقیق حاضر به دنبال یافتن شواهد علمی درباره آنها است. به علاوه، هنوز مشخص نیست که آیا تفاوت‌های میان رشته‌ای در علوم و علوم اجتماعی عامل مهمی در ویژگی‌های ساختاری یا محتوای شبکه استنادی در محیط وب است؟ برای مثال، آیا در رشته‌های علوم بیشتر از علوم اجتماعی نشر آثار علمی در محیط وب تحت تاثیر قالب پی.دی.اف، زبان انگلیسی، حوزه‌های اینترنتی دانشگاهی و استندهای فرایبندی قرار گرفته است و یا بالعکس؟

۲. پیشینه پژوهش

به‌طور کلی تحقیقات کمی در ارتباط با بررسی ویژگی‌های ساختاری یا محتوای نشر آثار علمی و شبکه استنادی میان آنها در محیط وب صورت گرفته است. یکی از اولین مطالعات استنادی مبتنی بر نمایه‌های استنادی خودکار در محیط وب، توسط گودرام^{۲۱} و دیگران انجام شد. این تحقیق، با هدف بررسی الگوی استنادی به مقالات کامپیوتر در قالب پی.دی.اف و پست‌اسکرپت^{۲۲} که در نمایه استنادی تحت وب سایت سیر وارد شده‌اند، صورت پذیرفت. نتایج به‌دست آمده از آن نشان می‌دهد که مقالات کنفرانس‌ها در رشته کامپیوتر بیشتر از مقالات مجلات در محیط وب مورد استناد قرار می‌گیرند و اغلب آنها در قالب

پی.دی.اف در اینترنت منتشر می‌شوند. این پژوهش همچنین نشان می‌دهد که ۱۵ درصد پیوندهای وبی به مقالات پیوسته در رشته کامپیوتر، مرتبط با مقالات ارائه شده در کنفرانس‌ها است و تنها ۳ درصد از پیوندهای وبی از سوی مجلات نمایه شده در پایگاه آی.اس.آی ایجاد شده‌اند (گودرام و دیگران، ۲۰۰۱). اهمیت این تحقیق از آنجا ناشی می‌شود که به جای استفاده از داده‌های استنادی در پایگاه آی.اس.آی که در اغلب مطالعات علم‌سنجی از آن استفاده می‌گردد، داده‌های استنادی از منبع دیگری یعنی نمایه استنادی خود محور تحت وب استخراج شده است که قابلیت شناسایی و نمایه‌سازی استنادی مقالات کنفرانس‌ها را نیز دارد.

چپسین و دیگران مبتنی بر تحلیل محتوای ۶۰۰ سایت وب بازیابی شده در زمینه زیست‌شناسی گیاهی از طریق موتورهای کاوش تجاری، به مطالعه ویژگی‌های انتشارات علمی در محیط وب پرداختند. نتایج به دست آمده از این تحقیق نشان می‌دهد که میان نوع فایل‌های پی.دی.اف و نشر آثار علمی (نظیر مقالات تحقیقاتی) ارتباط مستقیم و معنی‌داری وجود دارد. بنابراین، می‌توان گفت که فایل‌های پی.دی.اف نقش نسبتاً مهمی در ارتباطات علمی رسمی در محیط وب در رشته مذکور دارند (چپسین و دیگران، ۲۰۰۴). لازم به ذکر است در تحقیق حاضر تنها یک موضوع (زیست‌شناسی گیاهی) مورد بررسی قرار گرفته است و به سایر ویژگی‌های مورد نظر در تحقیق حاضر (شامل زبان، تاریخ انتشار، حوزه اینترنتی، فرایبندی بودن استنادها) توجه نشده است.

وُترز و دی ویریس^{۲۳} به بررسی ماهیت پیوندهای فرامتنی در فهرست مآخذ مقالات علمی در ۳۸ مجله مرتبط با رشته جامعه‌شناسی، کتابداری و اطلاع‌رسانی، زیست‌شیمی و بیوتکنولوژی، نروساینس^{۲۴} و ریاضیات پرداختند. هدف از این تحقیق بررسی ماهیت پیوندهای فرامتنی در ارتباطات علمی رسمی و انتشار آثار علمی بود. در این مطالعه برخی از ویژگی‌های ساختاری پیوندهای فرامتنی در فهرست مآخذ مقالات مورد مطالعه قرار گرفت. نتایج این تحقیق حاکی از آن است که مولفان از طریق استنادهای فرامتنی در فهرست مآخذ مقالات خود، به‌طور متوسط به آثار جدیدتری استناد می‌کنند. این تحقیق همچنین نشان می‌دهد

که ذکر استناد به منابع اطلاعاتی به صورت فرامتنی در مقالات پی.دی.اف بسیار کمتر از مقالات اِچ.اِتی.ام.ال است. از این رو، قالب فایل عامل مهمی در ایجاد استندهای فرامتنی در مقالات علمی است. ایجاد استندهای فرامتنی در مقالات به عوامل مختلف فنی، رفتاری و انتشاراتی بستگی دارد. به این ترتیب، ناشران مجلات نقش مهمی در فعال کردن استندهای فرامتنی در مقالات علمی دارند (وترز و دی ویریس، ۲۰۰۴).

۳. پرسش‌های پژوهش

۳-۱. در هر یک از رشته‌های علوم و علوم اجتماعی، نوع و درصد ویژگی‌های منابع استندهای برگرفته از موتور کاوش گوگل که مقالات مورد مطالعه را هدف قرار داده‌اند چیست؟ این ویژگی‌ها شامل:

۳-۱-۱. قالب فایل: منابع استنادی در وب در چه قالب‌های فایلی (شامل HTML، PDF، DOC، PostScript) منتشر شده‌اند؟ این پرسش نمایانگر نقش انواع قالب فایل در نشر آثار علمی و شبکه استنادی میان آنها در محیط وب و در رشته‌های مورد مطالعه است.

۳-۱-۲. حوزه اینترنتی: حوزه اینترنتی منابع استنادی در وب (شامل edu, ac, org, com) چیست؟ این پرسش در پی شناخت اهمیت حوزه‌های اینترنتی به ویژه حوزه‌های اینترنتی دانشگاهی edu و ac^{۲۵} در ارتباطات علمی رسمی و ایجاد شبکه استنادی تحت وب است.

۳-۱-۳. استندهای فرایوندی: چه نسبتی از استندهای وبی در بخش فهرست مآخذ منابع استنادی در قالب فرایوندی یا متن منابع اطلاعاتی در محیط وب را هدف قرار داده‌اند؟ این سؤال در پی بررسی اهمیت ساختار فرایوندی در شبکه استنادی در محیط وب است.

۳-۱-۴. زبان: زبان‌های غیرانگلیسی چه نقشی را در ارتباطات علمی رسمی در محیط وب دارند؟ طرح این سؤال می‌تواند اطلاعات بیشتری را در باب نقش زبان‌های غیرانگلیسی در شبکه استنادی میان آثار علمی در محیط وب ارائه دهد.

۳-۱-۵. تاریخ انتشار: تاریخ انتشار منابع استنادی در وب چیست (۲۰۰۱-۲۰۰۶)؟
این سؤال به ویژه در پی کشف میزان استنادهای فوری^{۲۶} (۲۰۰۱-۲۰۰۲) به منابع اطلاعاتی است. به این ترتیب، می‌توان فرهنگ استنادی محققان به آثار علمی در حوزه‌های علوم و علوم اجتماعی را مورد تجزیه و تحلیل کمی قرار داد.

۴. روش تحقیق

به‌طور خلاصه، در تحقیق حاضر پنج ویژگی از طریق تحلیل محتوای ۵۴۵ منبع استنادی نمونه‌گیری شده در محیط وب که ۱۶۵۰ مقاله علمی منتشر شده در ۱۰۸ مجله الکترونیکی (تمام متن داوری شده) در چهار رشته از علوم (شیمی، فیزیک، زیست‌شناسی و کامپیوتر) و چهار رشته از علوم اجتماعی (جامعه‌شناسی، روان‌شناسی، اقتصاد و علوم تربیتی) را در بخش فهرست مآخذ خود مورد استناد قرار داده‌اند، مورد بررسی قرار گرفته است. شایان ذکر است، برای اینکه بتوان تحقیق حاضر را در مدت زمان مناسب به اتمام رساند، تنها دو حوزه علوم و علوم اجتماعی و هشت رشته مرتبط با آنها مورد مطالعه قرار گرفته است. بنابراین تحقیقات آتی می‌توانند حوزه‌های علوم انسانی و پزشکی و طیف گسترده‌تری از رشته‌ها را به منظور تجزیه و تحلیل ویژگی‌های شبکه استنادی و نشر آثار علمی مورد مطالعه قرار دهند.

۴-۱. انتخاب مجلات و مقالات

برای انتخاب مجلات الکترونیکی تمام متن در چهار رشته از علوم و چهار رشته از علوم اجتماعی از راهنمای مجلات الکترونیکی با دسترسی آزاد^{۲۷} استفاده شده است. این راهنما یکی از جامع‌ترین و شناخته شده‌ترین راهنماهای موضوعی است که مجلات الکترونیکی تمام متن علمی) را نمایه می‌کند. با این وجود، از منابع دیگری نیز نظیر گزارش موسسه اطلاعات علمی (آی.اس.آی) درباره اثرگذاری استنادی مجلات الکترونیکی نمایه شده در پایگاه وب آو ساینس (گزارش آی.اس.آی، ۲۰۰۴). مجلات انتخاب شده به زبان انگلیسی و

در سال ۲۰۰۱ منتشر شده بودند و دارای نوعی نظام کنترل کیفی محتوا بودند (داوری شده^{۲۸} یا بررسی شده توسط هیات تحریریه^{۲۹}). سال ۲۰۰۱، به این دلیل انتخاب شده است تا مقالات از مدت زمان کافی برای دریافت استناد از سوی منابع اطلاعاتی در محیط وب برخوردار باشند.

برای انتخاب مقالات در هر یک از هشت رشته مورد مطالعه، از روش نمونه‌گیری طبقه‌ای^{۳۰} استفاده شده است. این روش نمونه‌گیری امکان انتخاب نمونه آماری را به‌طور نسبی بر اساس هر مجله می‌دهد و از این لحاظ شیوه مناسب‌تری نسبت به نمونه‌گیری تصادفی ساده^{۳۱} است. به این ترتیب، مجلاتی که دارای مقالات منتشر شده بیشتری بودند، به همان میزان از نمونه آماری بیشتری نیز برخوردار شدند و برعکس. شایان ذکر است در تحقیق حاضر تنها مقالات تحقیقی مورد بررسی قرار گرفته است و از بررسی یادداشت‌های سردبیری، گزارش‌ها، معرفی و نقد مقالات، خدمات و کتاب‌ها، و مقالات کوتاه^{۳۲} به دلیل بحث انگیز بودن ماهیت علمی آنها و تعمیم نتایج به‌دست آمده از آنها به ارتباطات علمی و ماهیت شبکه استنادی در محیط وب صرف نظر شده است.

۲-۴. روش جستجو و بازیابی منابع استنادی

به منظور اینکه بتوان منابع استنادکننده به مقالات مجلات نمونه‌گیری شده را در محیط وب بازیابی کرد و سپس ویژگی‌های مورد نظر آنها را مورد بررسی قرار داد، از موتور کاوش گوگل به عنوان مناسب‌ترین ابزار گردآوری استنادهای وبی استفاده شده است. موتور کاوش گوگل دارای بیشترین میزان نتایج پایدار در طول زمان‌های مختلف است (واگان، ۲۰۰۴ بی) و جامعیت بیشتری نیز نسبت به سایر موتورهای کاوش تجاری دارد (بار-ایلان، ۲۰۰۴ بی). یکی دیگر از قابلیت‌های برجسته این ابزار کاوش پوشش نسبتاً جامع آن از منابع اینترنتی در قالب انواع مختلفی از فایل‌های الکترونیکی (نظیر PDF, RTF, DOC, PPT, XLS, PS) است که بنا بر اهداف تحقیق استفاده از آن را توجیه می‌کند. برای گردآوری و شناسایی منابع استنادی در وب از روشی استفاده شده است

که در واقع ترکیبی از دو روش پیشین، برای بازیابی دامنه گسترده‌تری از داده‌ها است. روش به کار رفته در این تحقیق که تحت عنوان *Web/URL citation extraction* نام گذاری شده است، ترکیبی از دو روش جست و جوی استنادهای وبی از طریق عنوان مقالات (*Web citation*) یا نشانی اینترنتی (*URL citation*) آنها است. روش اول توسط *و اگان و شا* (واگان و شا، ۲۰۰۳) و روش دوم توسط *کوشا و ثلوال* (کوشا و ثلوال، ۲۰۰۶) برای بازیابی استنادهای وبی به مقالات مجلات بکار رفته است. در روش نوینی که در تحقیق حاضر به کار رفته است، از عملگر *OR* برای ترکیب عنوان مقالات یا نشانی اینترنتی مقالات مجلات الکترونیکی استفاده شده است. در مثال زیر، نمونه جستجوی بکار رفته برای یک عنوان مقاله در قالب پی.دی.اف نشان داده شده است. این جستجو منجر به بازیابی منابع استنادی (منابع استنادکننده) می‌شود که در آنها: الف. عنوان یا نشانی اینترنتی مقاله در قالب فرامتن در صفحات وب ظاهر شده باشد. ب. عنوان یا نشانی اینترنتی مقاله در قالب متن ایستا (حتی اگر فرامتنی نباشد) در صفحات وب درج شده باشد. همان طوری که ملاحظه می‌شود، از قابلیت جستجوی *site-* نیز برای نادیده گرفتن استنادهای وبی احتمالی از فضای سایت خود مجله استفاده شده است:

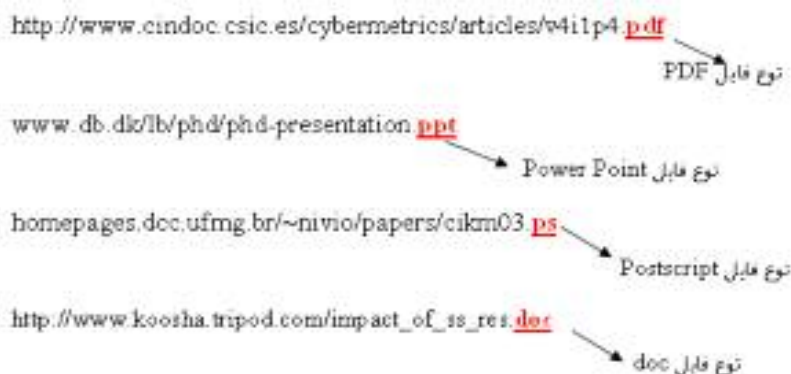
"Enantioselective solvent-free Robinson annulation reactions" OR
www.ias.ac.in/chemscup/pdf-Jun2001/Pc3049.pdf -site:ias.ac.in/

بدیهی است انجام جستجوهای مشابه در موتور کاوش گوگل، به بازیابی سایت‌ها یا صفحاتی نیز منجر می‌شود که به دلیل استناد رسمی، مقالات مجلات مورد مطالعه را هدف قرار نداده‌اند. در واقع، بسیاری از استنادهای وبی ممکن است به دلایلی غیر از استناد رسمی در اسناد وبی ظاهر شده باشند نظیر راهبری، آموزش، اشاعه اطلاعات و غیره. بنابراین، در تحقیق حاضر تنها آن دسته از منابع استنادی مورد بررسی قرار گرفته است که استنادهای وبی به‌طور مشخص در بخش فهرست منابع و مآخذ آنها ظاهر شده باشد. به این ترتیب، از میان نتایج بازیابی شده از گوگل ۵۴۵ منبع استنادی که در بخش فهرست مآخذ خود مقالات

مجلات مورد مطالعه در هشت رشته از علوم و علوم اجتماعی را مورد استناد قرار داده بوده‌اند، انتخاب شدند و ویژگی‌های مورد نظر نیز برگرفته از آنها بوده است. شایان ذکر است، منابع استنادی مورد مطالعه، طیف گسترده‌ای از انواع منابع اطلاعاتی علمی را نظیر مقاله مجله، مقاله کنفرانس، پایان نامه، گزارش تحقیقاتی و غیره را در بر دارد.

۳-۴. روش تحلیل ویژگی‌های منابع استنادی وی

به منظور شناسایی و تحلیل ویژگی‌های منابع استنادی یا به عبارت دیگر استنادکننده برگرفته از جستجوهای گوگل که با انگیزه استناد رسمی، مقالات الکترونیکی را هدف قرار داده‌اند، پنج ویژگی زیر مرتبط با هریک از منابع استنادی به صورت مشاهده عینی مورد مطالعه و ثبت گردیده است، این موارد شامل: ۳-۴-۱. نوع قالب فایل: بررسی نوع قالب فایل منابع استنادی در وب (شامل HTML, PDF, DOC, PostScript) تا حدود بسیار زیادی می‌تواند نمایانگر نقش ساختار فایل‌های الکترونیکی در ارتباطات علمی و شبکه استنادی در محیط وب باشد. روش شناسایی نوع فایل‌های منابع استنادی از طریق مشاهده عینی محتوا یا بررسی اختصارات بکار رفته در انتهای نشانی اینترنتی آنها بوده است. در زیر نمونه‌هایی از قالب‌های فایل مرتبط با منابع استنادی در وب نشان داده شده است که برای تقسیم‌بندی آنها در این تحقیق از آنها استفاده شده است:



۲-۳-۴. حوزه اینترنتی: تقسیم‌بندی حوزه اینترنتی منابع استنادی براساس حوزه‌های دانشگاهی (edu., ac.)، سازمان‌های غیر انتفاعی (org.)، تجاری (com.) و سایر حوزه‌ها صورت گرفته است. این فرایند، از طریق مشاهده عینی نشانی اینترنتی منابع استنادی دنبال شده است. به این ترتیب، منابع استنادی وبی با حوزه‌های اینترنتی edu (شامل دانشگاه‌ها و مراکز آموزش عالی آمریکا و برخی دیگر از کشورهای جهان) و ac (بسیاری از دانشگاه‌های جهان نظیر ایران، انگلستان، هند) به عنوان منابع استنادی برگرفته از حوزه‌های دانشگاهی؛ حوزه اینترنتی org به عنوان حوزه سازمان‌های غیرانتفاعی؛ و com به عنوان حوزه تجاری در نظر گرفته شده است. شایان ذکر است، حوزه اینترنتی برخی از دانشگاه‌های دنیا نظیر کانادا و برخی از کشورهای اروپایی از طریق edu و ac قابل ردگیری نیستند. بنابراین، حوزه‌های edu و ac تنها نماینده بخشی از منابع استنادی برگرفته از فضای وب سایت‌های دانشگاهی هستند و نتایج به دست آمده را نمی‌توان به فضای اینترنتی کل دانشگاه‌های دنیا تعمیم داد. در زیر نمونه‌هایی از منابع استنادی از حوزه‌های اینترنتی مختلف نشان داده شده که مبنای تقسیم‌بندی آنها در تحقیق حاضر بوده است:



۲-۳-۴. تاریخ انتشار: این ویژگی از طریق مشاهده عینی تاریخ انتشار منابع استنادی در وب به دست آمده است. روش شناسایی تاریخ انتشار منابع استنادی اغلب از طریق مشاهده تاریخ انتشار آنها در بخش سرصفحه یا پاصفحه آثار

علمی است. با این وجود، در برخی از موارد از طریق روش‌های مختلف امکان شناسایی تاریخ انتشار منابع استنادی وجود نداشت که در این صورت تاریخ انتشار آنها «نامشخص» محسوب شده است. هدف از بررسی تاریخ انتشار منابع استنادی وبی، به دست آوردن شواهدی علمی مبتنی بر مدت زمانی است که مقالات توانسته‌اند بیشترین استنادهای رسمی را در محیط وب دریافت کنند. در این تحقیق، از این ویژگی تحت عنوان استناد سریع^{۳۳} نام برده شده است که دلالت بر درصد استنادهای رسمی دارد که دو سال (۲۰۰۱-۲۰۰۲) پس از انتشار مقالات به دست آمده است. این موضوع خود در بررسی الگوهای استنادی در محیط وب از اهمیت بسزایی برخوردار است و تقریباً مشابه کاربردی است که نمایه استنادهای سریع^{۳۴} موسسه اطلاعات علمی (آی.اس.آی) دارد.

۴-۳-۴. زبان: به منظور بررسی ویژگی‌های زبانی منابع استنادی در هشت رشته مورد مطالعه، زبان انگلیسی و زبان‌های غیرانگلیسی مدنظر بوده است. هدف این بخش از تحقیق، بررسی میزان اهمیت زبان‌های غیرانگلیسی در شبکه استنادی میان منابع اطلاعات علمی و مشخص ساختن نقش تفاوت‌های میان رشته‌ای است.

۴-۳-۵. فرایبندی بودن استنادها: اینکه استنادهای وبی از چه پویایی برخوردارند، ویژگی است که تنها از طریق مشاهده عینی آنها در منابع استنادی بازیابی شده قابل بررسی است. جستجوی عنوان یا نشانی اینترنتی مقالات الکترونیکی در موتور کاوش گوگل ممکن است منجر به بازیابی استنادهایی شود که یا به صورت فرایبندی یا متنی در بخش فهرست مآخذ آنها درج شده باشند. منظور از استنادهای وبی فرایبندی این است که آنها به صورت پویا ایجاد شده باشند و امکان انتخاب آنها از طریق دکمه «ماوس» وجود داشته باشد. حال آن که استنادهای وبی متنی به صورت متن ایستا بوده و امکان انتخاب آنها برای مشاهده منبع استناد کننده از طریق دکمه «ماوس» وجود ندارد. اینکه کدام نوع از استنادهای وبی بیشتر از دیگری ارتباطات علمی رسمی در محیط وب را تحت تأثیر خود قرار داده است، از طریق نسبت حضور آنها در منابع استنادی قابل

شناسایی است. در زیر به ترتیب نمونه یک استناد وبی فرایوندی و یک استناد وبی بر گرفته از بخش فهرست مآخذ اسناد وبی نشان داده شده است:

استناد وبی بر گرفته از بخش فهرست مآخذ اسناد وبی نشان داده شده است:



۵. یافته های پژوهش

۵-۱. ویژگی‌های منابع استنادی در حوزه علوم

همان طوری که در جدول ۱ مشاهده می‌شود، نتایج تحقیق نشان می‌دهد که ۵۵ درصد از استنادهای وبی که مقالات مجلات مورد مطالعه در چهار رشته از علوم را در بخش فهرست مآخذ خود مورد استناد قرار داده‌اند، در قالب پی.دی.اف در محیط وب منتشر شده‌اند و ۶۴ درصد از استنادها به صورت متنی (غیر فرامتنی) در بخش فهرست مآخذ آنها درج شده است. بنابراین می‌توان گفت که ارتباطات علمی رسمی و شبکه استنادی در محیط وب در رشته‌های مورد مطالعه، تحت تاثیر استنادهای متنی از اسناد پی.دی.اف قرار گرفته است. در حقیقت، حجم بسیار زیادی از آثار علمی منتشر شده در وب نظیر مقالات مجلات الکترونیکی، مقالات کنفرانس‌ها، گزارش‌های تحقیقاتی، پایان‌نامه‌های دانشگاهی و غیره در قالب پی.دی.اف منتشر می‌شوند، و احتمالاً به همین دلیل ساختار استنادهای وبی در بخش فهرست مآخذ این‌گونه آثار به صورت غیر فرامتنی است. به این ترتیب، یکی از مهمترین یافته‌های این بخش از تحقیق این است که روش جستجوی استنادهای وبی به مقالات مجلات در مقایسه با جستجوی پیوندهای وبی (Link Search) که در بسیاری از تحقیقات پیشین به کار گرفته شده است، روش

بهتری برای سنجش اثرگذاری مقالات و مجلات علمی در محیط وب است. زیرا از طریق روش جستجوی پیوندهای وبی در صورتی می توان منابع استنادی را بازیابی کرد که به صورت فرایبندی ظاهر شده باشند. حال آنکه بخش قابل ملاحظه‌ای از استندهای وبی در قالب غیر فرامتنی به ویژه در اسناد پی.دی.اف. درج می شوند که از طریق روش فوق غیر قابل ردگیری هستند.

همچنین، از طریق جدول ۱ می توان دریافت که ۲۶ درصد از منابع استنادی از حوزه‌های اینترنتی دانشگاهی شامل edu و ac نشأت گرفته است. اگرچه بسیاری از دانشگاه‌ها از حوزه‌های فوق استفاده نمی کنند (نظیر دانشگاه‌های کانادا و برخی از دانشگاه‌های کشورهای اروپایی) و این خود محدودیت تحقیق حاضر در شناسایی این گونه حوزه‌های اینترنتی به شمار می آید، اما می توان گفت که حداقل حدود یک سوم منابع استنادی در محیط وب که مقالات مجلات را از طریق درج اطلاعات کتاب‌شناختی آنها در فهرست مآخذ خود هدف قرار داده‌اند، مرتبط با سایت‌های دانشگاهی است. این یافته اهمیت منابع علمی ذخیره شده روی سایت دانشگاه‌ها در ارتباطات علمی رسمی را نشان می‌دهد که می‌تواند در طراحی و توسعه ابزارهای کاوش اطلاعات علمی مورد توجه قرار گیرد.

در رابطه با سال انتشار منابع استنادی و بررسی اینکه چه مدت زمان طول کشیده است که مقالات مجلات مورد مطالعه در این تحقیق به صورت رسمی توسط سایر اسناد وبی مورد استناد قرار گیرند، نتایج تحقیق نشان می‌دهد که ۲۸ درصد از منابع استنادی در خلال سال‌های ۲۰۰۱ و ۲۰۰۲ منتشر شده‌اند. اگرچه این میزان در فیزیک بیشتر (۴۶ درصد) و در زیست‌شناسی (۱۷ درصد) کمتر بوده است. یافته‌های این بخش از پژوهش، نتایج تحقیقات پیشین را مورد تایید قرار می‌دهد که در رشته فیزیک فرهنگ استنادی سریع‌تری حاکم است و فاصله زمانی میان ارسال مقالات به آرشیوهای مقالات E-print و استناد به آنها نسبت به سایر رشته‌ها بسیار کمتر است (برادی، کار و هارنارد، ۲۰۰۲).

همچنین جدول ۱ نشان می‌دهد که تقریباً ۸۰ درصد از منابع استنادی در چهار رشته از علوم به زبان انگلیسی و ۲۰ درصد به سایر زبان‌ها است. نتایج این بخش

از تحقیق نشان می‌دهد که از طریق محیط وب می‌توان به حجم قابل ملاحظه‌ای از منابع استنادی به سایر زبان‌های غیرانگلیسی دست یافت که ممکن است در تجزیه و تحلیل استنادی و بررسی رویه‌های ارتباطات علمی از اهمیت به سزایی برخوردار باشند، اما در گذشته به دلیل محدودیت پایگاه استنادی آی.اس.آی در پوشش منابع اطلاعاتی غیر انگلیسی زبان، این مسئله همواره در مطالعات علم‌سنجی کمتر مورد توجه و بررسی بوده است.

به علاوه از طریق مراجعه به جدول ۱ می‌توان به تفاوت‌های میان رشته‌ای در چهار رشته از علوم نیز پی برد و تفاسیر مختلفی مرتبط با آنها ارائه داد. برای مثال، در رشته شیمی تنها ۱۲ درصد از منابع استنادی به زبان‌های غیر از انگلیسی بودند و این خود احتمالاً وابستگی ارتباطات علمی رسمی به زبان انگلیسی را در این رشته نشان می‌دهد. در نقطه مقابل، در رشته کامپیوتر، تعداد نسبتاً بیشتری از منابع استنادی به زبان‌های غیر از انگلیسی شناسایی شده است که احتمالاً نقش زبان‌های دیگر را در نشر آثار علمی در این حوزه نشان می‌دهد. همچنین، در رشته شیمی ۳۷ درصد از استنادهای وبی از دو حوزه اینترنتی دانشگاهی edu و ac نشأت گرفته است که این خود نقش دانشگاه‌ها را در نشر آثار علمی در رشته شیمی و در محیط وب نشان می‌دهد.

همچنین، در رشته‌های فیزیک (۶۶ درصد) و کامپیوتر (۶۸ درصد) نسبت به شیمی (۵۰ درصد) و زیست‌شناسی (۴۰ درصد) درصد بیشتری از منابع استنادی در قالب پی.دی.اف در محیط وب منتشر شده است. از آنجا که امروزه قالب فایل پی.دی.اف اغلب به عنوان رایج‌ترین نوع قالب انتشار آثار علمی در محیط وب شناخته می‌شود، می‌توان این طور استنباط کرد که محققان در رشته‌های کامپیوتر و فیزیک بیشتر از ویژگی‌های نشر آثار علمی خود در قالب پی.دی.اف آگاهی دارند.

| ویژگی منابع استنادی وب | تقسیم بندی | شیمی | زیست شناسی | فیزیک | کامپیوتر | جمع % |
|---------------------------|--------------|-------------|-------------|-------------|-------------|--------------|
| زبان | زبان انگلیسی | ۲۸ (%.۸۵/۵) | ۸۲ (%.۸۰/۴) | ۵۰ (%.۸۲) | ۷۰ (%.۷۷/۸) | ۲۳۰ (%.۸۰/۷) |
| | سایر زبان‌ها | ۴ (%.۱۲/۵) | ۲۰ (%.۱۹/۶) | ۱۱ (%.۱۸) | ۲۰ (%.۲۲/۲) | ۵۵ (%.۱۹/۳) |
| حوزه اینترنتی منابع | edu. | ۹ (%.۲۸/۱) | ۱۲ (%.۱۱/۸) | ۶ (%.۹/۸) | ۱۷ (%.۱۸/۹) | ۴۴ (%.۱۵/۴) |
| | ac. | ۳ (%.۹/۴) | ۶ (%.۵/۹) | ۱۰ (%.۱۶/۴) | ۱۱ (%.۱۲/۲) | ۳۰ (%.۱۰/۵) |
| | org | ۷ (%.۲۱/۹) | ۴۲ (%.۴۱/۲) | ۹ (%.۱۴/۸) | ۱۹ (%.۲۱/۱) | ۷۷ (%.۲۷) |
| | com | ۳ (%.۹/۴) | ۲۰ (%.۱۹/۶) | ۱۷ (%.۲۷/۹) | ۲۲ (%.۲۴/۴) | ۶۲ (%.۲۱/۸) |
| | سایر حوزه ها | ۱۰ (%.۳۱/۳) | ۲۲ (%.۲۱/۶) | ۱۹ (%.۳۱/۱) | ۲۱ (%.۲۳/۳) | ۷۲ (%.۲۵/۳) |
| | فرامتنی | ۷ (%.۲۱/۹) | ۵۷ (%.۵۵/۹) | ۲۰ (%.۳۲/۸) | ۱۹ (%.۲۱/۱) | ۱۰۳ (%.۳۶/۱) |
| استادها | متنی | ۲۵ (%.۷۸/۱) | ۴۵ (%.۴۴/۱) | ۴۱ (%.۶۷/۲) | ۷۱ (%.۷۸/۹) | ۱۸۲ (%.۶۳/۹) |
| تاریخ انتشار | ۲۰۰۱ | ۶ (%.۱۸/۸) | ۳ (%.۲/۹) | ۱۶ (%.۲۶/۲) | ۸ (%.۸/۹) | ۳۳ (%.۱۱/۶) |
| | ۲۰۰۲ | ۵ (%.۱۵/۶) | ۱۴ (%.۱۳/۷) | ۱۲ (%.۱۹/۷) | ۱۶ (%.۱۷/۸) | ۴۷ (%.۱۶/۵) |
| | ۲۰۰۳ | ۸ (%.۲۵) | ۲۵ (%.۲۴/۵) | ۹ (%.۱۴/۸) | ۱۹ (%.۲۱/۱) | ۶۱ (%.۲۱/۴) |
| | ۲۰۰۴ | ۸ (%.۲۵) | ۳۳ (%.۳۲/۴) | ۸ (%.۱۳/۱) | ۲۰ (%.۲۲/۲) | ۶۹ (%.۲۴/۲) |
| | ۲۰۰۵ | ۱ (%.۳/۱) | ۱۸ (%.۱۷/۶) | ۷ (%.۱۱/۵) | ۱۰ (%.۱۱/۱) | ۳۶ (%.۱۲/۶) |
| | نامشخص | ۴ (%.۱۲/۵) | ۹ (%.۸/۸) | ۹ (%.۱۴/۸) | ۱۷ (%.۱۸/۹) | ۳۹ (%.۱۳/۷) |
| قالب فایل | PDF | ۱۶ (%.۵۰) | ۴۱ (%.۴۰/۲) | ۴۰ (%.۶۵/۶) | ۶۱ (%.۶۷/۸) | ۱۵۸ (%.۵۵/۴) |
| | HTML | ۱۴ (%.۴۳/۸) | ۵۹ (%.۵۷/۸) | ۱۸ (%.۲۹/۵) | ۱۸ (%.۲۰) | ۱۰۹ (%.۳۸/۲) |
| | DOC | ۱ (%.۳/۱) | ۲ (%.۲) | ۲ (%.۳/۳) | ۷ (%.۷/۸) | ۱۲ (%.۴/۲) |
| | PS | ۱ (%.۳/۱) | ۰ (%.۰) | ۱ (%.۱/۶) | ۴ (%.۴/۴) | ۶ (%.۲/۱) |

جدول ۱. ویژگی‌های منابع استنادی وبی در چهار رشته از علوم

۲-۵. ویژگی‌های منابع استنادی در حوزه علوم اجتماعی

در جدول ۲ ویژگی‌های منابع استنادی در چهار رشته از علوم اجتماعی نشان داده شده است.

| ویژگی منابع استنادی وب | تقسیم بندی | روانشناسی | جامعه شناسی | علوم تربیتی | اقتصاد | جمع % |
|------------------------|--------------|-------------|-------------|-------------|-------------|--------------|
| زبان | زبان انگلیسی | ۲۸ (%/۸۴/۵) | ۵۷ (%/۹۳/۴) | ۶۳ (%/۸۸/۷) | ۸۲ (%/۸۶/۳) | ۲۳۰ (%/۸۸/۵) |
| | سایر زبان‌ها | ۵ (%/۱۵/۲) | ۴ (%/۶/۶) | ۸ (%/۱۱/۳) | ۱۳ (%/۱۳/۷) | ۳۰ (%/۱۱/۵) |
| حوزه اینترنتی منابع | edu. | ۶ (%/۱۸/۲) | ۱۲ (%/۱۹/۷) | ۱۶ (%/۲۲/۵) | ۲۲ (%/۲۳/۲) | ۵۶ (%/۲۱/۵) |
| | ac. | ۵ (%/۱۵/۲) | ۷ (%/۱۱/۵) | ۷ (%/۹/۹) | ۱۲ (%/۱۲/۶) | ۳۱ (%/۱۱/۹) |
| | org | ۳ (%/۹/۱) | ۱۴ (%/۲۳) | ۱۳ (%/۱۸/۳) | ۲۵ (%/۲۶/۳) | ۵۵ (%/۲۱/۲) |
| | com | ۵ (%/۱۵/۲) | ۷ (%/۱۱/۵) | ۶ (%/۸/۵) | ۸ (%/۸/۴) | ۲۶ (%/۱۰) |
| | سایر حوزه ها | ۱۴ (%/۴۲/۴) | ۲۱ (%/۳۴/۴) | ۲۹ (%/۴۰/۴) | ۲۸ (%/۲۹/۵) | ۹۲ (%/۳۵/۴) |
| | فرامتنی | ۴ (%/۱۲/۱) | ۴ (%/۲۳) | ۲۲ (%/۳۱) | ۱۶ (%/۱۶/۸) | ۵۶ (%/۲۱/۵) |
| استنادها | متنی | ۲۹ (%/۸۷/۹) | ۴۷ (%/۷۷) | ۴۹ (%/۶۹) | ۷۹ (%/۸۳/۲) | ۲۰۴ (%/۷۸/۵) |
| | ۲۰۰۱ | ۴ (%/۱۲/۱) | ۵ (%/۸/۲) | ۶ (%/۸/۵) | ۱۲ (%/۱۲/۶) | ۲۷ (%/۱۰/۴) |
| تاریخ انتشار | ۲۰۰۲ | ۶ (%/۱۸/۲) | ۸ (%/۱۳/۱) | ۱۴ (%/۱۹/۷) | ۱۸ (%/۱۸/۹) | ۴۶ (%/۱۷/۷) |
| | ۲۰۰۳ | ۶ (%/۱۸/۲) | ۱۳ (%/۲۱/۳) | ۱۵ (%/۲۱/۱) | ۲۲ (%/۲۳/۲) | ۵۶ (%/۲۱/۵) |
| | ۲۰۰۴ | ۸ (%/۲۴/۲) | ۱۵ (%/۲۵/۶) | ۱۷ (%/۲۳/۹) | ۱۸ (%/۱۸/۹) | ۵۸ (%/۲۲/۳) |
| | ۲۰۰۵ | ۵ (%/۱۵/۲) | ۸ (%/۱۳/۱) | ۱۰ (%/۱۴/۱) | ۱۴ (%/۱۴/۷) | ۳۷ (%/۱۴/۲) |
| | نامشخص | ۴ (%/۱۲/۱) | ۱۲ (%/۱۹/۷) | ۹ (%/۱۲/۷) | ۱۱ (%/۱۱/۶) | ۳۶ (%/۱۳/۸) |
| قالب فایل | PDF | ۱۸ (%/۵۴/۵) | ۴۶ (%/۷۵/۴) | ۴۱ (%/۵۷/۷) | ۶۹ (%/۷۲/۶) | ۱۷۴ (%/۶۶/۹) |
| | HTML | ۱۲ (%/۳۶/۴) | ۹ (%/۱۴/۸) | ۲۴ (%/۳۳/۸) | ۱۸ (%/۱۸/۹) | ۶۳ (%/۲۴/۲) |
| | DOC | ۳ (%/۹/۱) | ۵ (%/۸/۲) | ۵ (%/۷) | ۷ (%/۷/۴) | ۲۰ (%/۷/۷) |
| | PS | ۰ (%/۰) | ۱ (%/۱/۶) | ۱ (%/۱/۴) | ۱ (%/۱/۱) | ۳ (%/۱/۲) |

جدول ۲. ویژگی‌های منابع استنادی وبی به مقالات مجلات

در چهار رشته از علوم اجتماعی

همان طور که مشاهده می‌شود ۶۷ درصد از منابع استنادی وبی در قالب فایل پی.دی.اف در محیط وب قرار گرفته‌اند که این میزان بیش از حوزه علوم (۵۵ درصد) است. به عبارت دیگر، می‌توان این طور استنباط کرد که در رشته‌های علوم اجتماعی حتی بیشتر از علوم، نشر آثار علمی در قالب فایل‌های پی.دی.اف صورت می‌گیرد. همچنین، ۷۸ درصد از استنادهای وبی به صورت متنی (غیر فرامتنی) در بخش فهرست مآخذ منابع اینترنتی درج شده است که این میزان نیز بیشتر از حوزه علوم (۶۳ درصد) است. بنابراین، نتایج تحقیق نشان می‌دهد که به‌طور کلی قالب فایل پی.دی.اف و استنادهای وبی متنی، از ویژگی‌های غالب نشر

آثار علمی و شبکه استنادی در محیط وب به شمار می‌آیند. جدول ۲ همچنین نشان می‌دهد که ۳۳ درصد از منابع استنادی مربوط به حوزه‌های اینترنتی دانشگاهی شامل edu و ac می‌باشند که این میزان نیز بیش از حوزه علوم (۲۶ درصد) است. بنابراین، می‌توان گفت در رشته‌های علوم اجتماعی نسبت به علوم، دانشگاه‌ها نقش بیشتری در اشاعه اطلاعات علمی با قابلیت دسترسی آزاد ایفا نموده‌اند.

بررسی سال نشر منابع استنادی در علوم اجتماعی نیز نشان می‌دهد که ۲۸ درصد از آنها مربوط به سال ۲۰۰۱ و ۲۰۰۲ هستند که این میزان مشابه حوزه علوم است. در واقع، تفاوتی میان زمان انتشار مقالات مجلات الکترونیکی مورد مطالعه در این تحقیق و استناد به آنها توسط سایر منابع اطلاعاتی دیده نمی‌شود. در خصوص زبان منابع استنادی وبی در حوزه علوم اجتماعی نیز نتایج تحقیق نشان می‌دهد که ۸۸ درصد از آنها به زبان انگلیسی هستند که این میزان به‌طور غیرمنتظره‌ای بیشتر از حوزه علوم است. در واقع، این انتظار وجود داشت که در رشته‌های علوم اجتماعی به دلیل اینکه اغلب نتایج تحقیقات مرتبط با حوزه‌های جغرافیایی ملی یا منطقه‌ای است (به ویژه در رشته‌های روان‌شناسی و جامعه‌شناسی)، تفاوت‌های زبانی بیشتری از نتایج تحقیق حاصل شود. برای مثال، میزان قابل توجهی از نتایج تحقیقات در رشته‌های علوم اجتماعی در مجلات بومی (نظیر مجلات غیر انگلیسی زبان) منتشر می‌شود (نگاه کنید به ندرهاف، ۲۰۰۶)، اما نتایج پژوهش، شواهد علمی در این رابطه را نشان نمی‌دهد. چه بسا، یکی از دلایل این امر را بتوان مشارکت کمتر کشورهای غیرانگلیسی زبان در انتشار متن کامل آثار علمی نظیر مقالات مجلات، مقالات کنفرانس‌های، مقالات الکترونیکی پیش چاپ یا پس چاپ، پایان‌نامه‌های دانشگاهی، طرح‌ها و گزارش‌های تحقیقاتی و از این قبیل در محیط و اینترنت جستجو کرد. به عبارت دیگر، این احتمال وجود دارد که نتایج تحقیقات در علوم اجتماعی بیشتر از علوم به زبان‌های غیرانگلیسی منتشر شوند، اما به دلیل این که از طریق ابتکارات ملی و بین‌المللی یا شخصی در محیط وب منتشر و در دسترس قرار نگرفته‌اند، شواهد علمی در ارتباط با این موضوع به‌دست نیامده باشد.

با این حال، تحقیقات آتی می‌تواند این موضوع را مورد بررسی قرار دهد.

بحث و نتیجه‌گیری

در پاسخ به سؤال اول تحقیق (۱-۱-۳)، نتایج به‌دست آمده از بررسی ویژگی‌های منابع استنادی در محیط وب در چهار رشته از علوم و چهار رشته از علوم اجتماعی نشان می‌دهد که فایل‌های پی‌دی‌اف (۶۱ درصد)، اچ‌اِی‌ام‌اِل (۳۱ درصد)، میکروسافت ورد (۶ درصد) و پُست‌اسکرپت (۲ درصد) به ترتیب بیشترین نوع قالب انتشار منابع علمی در شبکه استنادی در محیط وب را تشکیل می‌دهند. بنابراین، به نظر می‌رسد در طراحی و توسعه ابزارهای ذخیره و بازیابی اطلاعات علمی و به‌ویژه نمایه‌های استنادی خودکار یا خود محور در محیط وب، شناسایی و نمایه‌سازی آثار علمی در قالب فایل‌های پی‌دی‌اف از اهمیت بسزایی برخوردار است. در صورتی که این‌گونه ابزارها امکان شناسایی و نمایه‌سازی خودکار فایل‌های پی‌دی‌اف را نداشته باشند، احتمالاً از جامعیت و کارایی لازم برای ردگیری استنادی برخوردار نخواهند بود.

همچنین، نتایج تحقیق حاضر نشان می‌دهد که حدود یک سوم (۳۰ درصد) از منابع استنادی که مقالات مجلات مورد مطالعه در این تحقیق را هدف قرار داده‌اند، برگرفته از حوزه‌های اینترنتی دانشگاهی edu و ac هستند. [این تحقیق از لحاظ روش گردآوری اطلاعات برای دسترسی به منابع استنادی ذخیره شده بر روی سایت دانشگاه‌هایی که از حوزه‌های فوق استفاده نکرده‌اند (نظیر دانشگاه‌های کانادا و برخی از دانشگاه‌های اروپایی) با محدودیت روبرو بوده است. با این حال انتظار می‌رود درصد بیشتری از منابع استنادی را بتوان از سایت دانشگاه‌ها در سراسر جهان بازیابی کرد و سایت‌های دانشگاهی نقش مهمی نسبت به سایر حوزه‌های اینترنتی در شبکه استنادی در محیط وب ایفا نمایند]. [بدین ترتیب، چه بسا از طریق هدف‌گیری برنامه‌های کامپیوتری تحت عنوان «روبات» در ابزارهای جستجو و بازیابی اطلاعات علمی و به‌ویژه نمایه‌های استنادی خودکار به سمت سایت دانشگاه‌ها بتوان به نتایج بهتری به منظور ایجاد شبکه استنادی میان منابع اطلاعاتی علمی در

محیط وب دست یافت].

در پاسخ به یکی دیگر از سئوالات تحقیق می‌توان گفت که استندهای وبی به صورت متنی (غیر فرامتنی) نقش مهمتری نسبت به استندهای فرامتنی در ارتباطات علمی رسمی و شبکه استنادی در محیط وب در رشته‌های مورد مطالعه دارند. [در برخی از تحقیقات پیشین (نگاه کنید به ثلوال، واگان و بجمبوربون، ۲۰۰۵) از روش جستجوی پیوندهای وبی به منظور بررسی رویه‌های ارتباطات علمی استفاده شده است. نتایج این تحقیق حاکی از آن است که روش فوق احتمالاً از جامعیت لازم برای کشف جنبه‌های مختلف ارتباطات علمی برخوردار نیست، در واقع حجم بسیار زیادی از نشانی‌های اینترنتی در بخش فهرست مآخذ منابع اطلاعاتی ممکن است در قالب متن ایستا درج شده باشند که بدین ترتیب در نتایج تحقیق پوشش داده نخواهند شد]. بنابراین، پیشنهاد می‌شود در طراحی و توسعه نظام‌های بازیابی اطلاعات علمی نظیر نمایه‌های استنادی خودکار یا نیمه خودکار تحت وب به این گونه ویژگی‌های ساختاری و محتوایی حاکم بر شبکه استنادی توجه شود.

بررسی زبان منابع استنادی نشان داد که حدود ۱۵ درصد از این گونه منابع به زبان‌های غیر از انگلیسی در محیط وب در دسترس هستند. این خود قابلیت محیط وب را به منظور دسترسی به طیف گسترده‌تری از منابع اطلاعاتی علمی غیرانگلیسی نشان می‌دهد که در گذشته امکان ردگیری آنها از طریق پایگاه استنادی آی.اس.آی که اغلب مجلات انگلیسی زبان را پوشش می‌دهد، وجود نداشت. به نظر می‌رسد رشد و گسترش نشر آثار علمی در محیط وب (نظیر راه‌اندازی سایت مجلات علمی) توسط کشورهای غیرانگلیسی زبان می‌تواند نقش به‌سزایی در شبکه استنادی در محیط وب ایفا نماید و گذشت زمان بر روی نتایج تحقیق حاضر تاثیرگذار باشد.

[از آنجا که شرط لازم یک نمایه استنادی پویا در محیط وب، دسترسی به منابع استنادکننده و استناد شوند است، پیشنهاد می‌گردد مجلات علمی در صورتی که امکان انتشار متن کامل مقالات را به‌طور رایگان در محیط وب ندارند، حداقل

اطلاعات کتاب‌شناختی و فهرست مآخذ مقالات را در محیط وب منتشر سازند. در پیش گرفتن چنین سیاستی، احتمالاً باعث می‌گردد تا ابزارهای کاوش تجاری و نمایه‌های استنادی خودکار تحت وب بتوانند به اطلاعات کتاب‌شناختی و استنادی آثار علمی دسترسی پیدا کرده و محتوای مقالات را در پایگاه خود ذخیره سازند. بدین ترتیب، بسیاری از آثار علمی که به زبان غیرانگلیسی منتشر شده‌اند، فرصت این را خواهند داشت که در شبکه استنادی تحت وب حضور یافته و احتمالاً در سنجش اثرگذاری تحقیقات در سطح ملی و بین‌المللی محاسبه گردند. در نهایت می‌توان با رویکرد مشابهی شبکه استنادی میان منابع اطلاعاتی علمی در کشور را مورد مطالعه قرار داد تا بدین ترتیب بتوان ویژگی‌ها و ماهیت نشر آثار علمی در وب سایت‌های ایرانی «نظیر دانشگاه تهران» را مورد بررسی قرار داد. نتایج چنین تحقیقاتی می‌تواند به‌طور کاربردی به منظور طراحی و توسعه نمایه استنادی تحت وب فارسی در کشور مورد استفاده قرار گیرد.

پی‌نوشت‌ها

1. Bawden, Holtham & Courtney
2. Mowshowitz & Kawaguchi
3. Garfield
4. automatic/autonomous citation index
5. Cameron
6. CiteSeer نام نمایه‌ایی استنادی و خود محور در رشته کامپیوتر است.
7. Citebase
8. Google Scholar
9. ISI (Institute for Scientific information)
10. Web of Science
11. high citation impact
12. ISI journal selection process
13. Bollacker
14. Hitchcock
15. PDF (Portable Document Format)
16. file format
17. Internet domains
18. hyperlinking
19. citing sources
20. hyperlinked/non-hyperlinked citations
21. Goodrum

- 22. PostScript
- 23. Wouters & de Vries
- 24. neuroscience

۲۵. شایان ذکر است، حوزه اینترنتی برخی از دانشگاه‌های دنیا نظیر کانادا و اغلب کشورهای اروپایی از طریق روش فوق قابل ردگیری نیستند که این خود محدودیت روش تحقیق مذکور است.

- 26. rapid citation
 - 27. Directory of Open Access Journals (<http://www.doaj.org>)
 - 28. peer reviewed
 - 29. editorial reviewed
 - 30. proportional stratified sampling
 - 31. simple random sampling
۳۲. از این مقالات اغلب تحت عنوان Short Communication یا Communication Paper نام برده می شود.

- 33. rapid citation
- 34. Immediacy Index

منابع

- Bar-Ilan, J. (2004b). The use of Web search engines in information science research. *Annual Review of Information Science and Technology*, 38, 231-288.
- Bawden, D., Holtham, C. & Courtney, N. (1999). Perspectives on information overload. *Aslib Proceedings*, 51(8), 249-255.
- Bollacker, K. D., Lawrence, S. & Lee, C. (1998). CiteSeer: An Autonomous Web Agent for Automatic Retrieval and Identification of Interesting. In: *Proceedings of 2nd International ACM Conference on Autonomous Agents*, ACM Press, 1998: 116-123, Retrieved May 10, 2006, from <http://maya.cs.depaul.edu/~classes/ds575/papers/citeseer.pdf>
- Brody, T., Carr, L. & Harnad, S. (2002). Evidence of hypertext in the scholarly archive. *Proceedings of ACM Hypertext 2002*, Retrieved June 10, 2006, from <http://opcit.eprints.org/ht02-short/archiveht-ht02.pdf>
- Cameron, R.D. (1997). A Universal Citation Catalyst for Reform in Scholarly Communication. *First Monday*, 2 (4), Retrieved May 9, 2006, from http://www.firstmonday.dk/issues/issue2_4/cameron/index.html
- Garfield, E. (1955). Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas. *Science*, (122), 3159: 108-111. Retrieved May 3, 2006, from <http://www.garfield.library.upenn.edu/essays/v6p468y1983.pdf>
- Goodrum, A., McCain, K., Lawrence, S. & Giles, C.L. (2001). Scholarly publishing in the Internet age: a citation analysis of computer science literature. *Information Processing & Management*, 37(5), 661-676.
- Google Scholar (2006). Retrieved Feb 12, 2006, from <http://scholar.google.com/scholar/about.html>
- Hitchcock, S., Bergmark, D., Brody, T., Gutteridge, C., Carr, L., Hall, W., Lagoze, C., & Harnad, S. (2002). Open Citation Linking: The Way Forward. *D Lib Magazine*, 8, (10), Retrieved Jan 10, 2006, from <http://www.dlib.org/dlib/october02/hitchcock/10hitchcock.html>
- Infomine (2007). INFOMINE: Scholarly Internet Resource Collections. Retrieved Jan 1, 2007, from <http://infomine.ucr.edu/>
- ISI journal selection process (2004). Retrieved May 10, 2006, from <http://scientific.thomson.com/free/essays/selectionofmaterial/journalselection>
- ISI press release essay on the impact of open access journals: A citation study from Thomson ISI (2004). Retrieved November 13, 2004, from <http://www.isinet.com/oaj>
- Jepsen E., Seiden P., Ingwersen P., Björneborn L., Borlund P. (2004). Characteristics of scientific Web publications: preliminary data gathering and analysis. *Journal of the American Society for Information Science and Technology*, 55(14), 1239-1249.
- Kousha, K. & Thelwall, M. (2006a). Motivations for URL citations to open access

- library and information science articles. *Scientometrics*, 68(3), 501-517.
- Kousha, K. & Thelwall, M. (2007a). Google Scholar citations and Google Web/URL citations: A multi-discipline exploratory analysis, *Journal of the American Society for Information Science and Technology*, 58(7), 1055-1065. Preprint available at: <http://www.scit.wlv.ac.uk/%7Ecm1993/papers/GoogleScholarGoogleWeb-URLcitations.doc>
- Lawrence, S., Giles, C. L. & Bollacker, K. (1999), Digital libraries and autonomous citation indexing. *IEEE Computer*, 32 (6): 67-71, Retrieved April 1, 2006, from <http://csdl.computer.org/dl/mags/co/1999/06/r6067.pdf>
- Mowshowitz, A. & Kawaguchi, A. (2005). Measuring search engine bias. *Information Processing and Management*. 41(5), 1193-1205.
- Nederhof A. J. (2006). Bibliometric monitoring of research performance in the social sciences and the humanities: A review. *Scientometrics*, 66 (1), 81-100.
- Swan, A. & Brown, S. (2004). Report of the JISC/OSI open access journal authors survey, 1-76. Retrieved April 20, 2006, from http://www.jisc.ac.uk/uploaded_documents/JISCOAreport1.pdf
- Swan, A. & Brown, S. (2005). Open access self-archiving: an author study, 1-97. Retrieved April 20, 2006, from <http://eprints.ecs.soton.ac.uk/10999/01/jisc2.pdf>
- Thelwall, M., Vaughan, L., & Björneborn, L. (2005). Webometrics. *Annual Review of Information Science and Technology*, 39, Medford, NJ: Information Today Inc. 81-135.
- Vaughan, L. (2004b). New measurements for search engine evaluation proposed and tested. *Information Processing & Management*, 40(4), 677-691.
- Vaughan, L. & Shaw, D. (2003). Bibliographic and Web citations: What is the difference? *Journal of the American Society for Information Science and Technology*, 54(4), 1313-1324.
- Wouters, P. & de Vries, R. (2004). Formally citing the web. *Journal of the American Society for Information Science and Technology*, 55(14), 1250-1260.