

Study on Search Engines' Reaction to the Metadata Records Created Based on Combined Method of Rich Snippets and Linked Data

Seyyed Mahdi Taheri

Assistant Professor at Department of Knowledge and Information Science; Allameh Tabataba'i University;
Corresponding Author taherismster@gmail.com

Reza Nikzad Bahle

M.A. of Knowledge and Information Science; Allameh Tabataba'i University r.nikzad20@gmail.com

Mitra Samiee

Assistant Professor; Department of Knowledge and Information Science; Allameh Tabataba'i University samiei.mitra66@gmail.com

Received: 09, Dec. 2017 Accepted: 19, Dec. 2017

Abstract: The purpose of this research was to find out the reaction of Web Search Engines to Metadata records created based on the combined method of rich snippets and linked data. 200 metadata records in two groups (100 records as the control group with the normal structure and 100 records created based on microdata and implemented in RDF/XML as experimental group) extracted from the information gateway of Iranian Content National Consortium (INCN) were analyzed through an experimental approach. The metadata records of two groups were published on an independent website (www.mikzad.ir), and were introduced to Google and Bing search engines directly. Using a checklist, the status of index ability and visibility of the published metadata records in two search engines was examined. Findings show all the element values of the metadata records of two groups indexed by Google and Bing, and were visible in the search results, but the records were not displayed in the search results semantically like the metadata records created in rich snippets methods only. In other words, the spider-indexer software of Web search engines are not able to identify and index the metadata elements (tags) created based on microdata method and implemented in RDF (as the syntax of linked data method) and the search engines cannot display the records in the search results semantically. So, Web search engines are not adaptable to the standards of the semantic web.

Keywords: Indexing, Visibility, Metadata Records, Rich Snippets, Linked Data, Web Search Engines

**Iranian Journal of
Information
Processing and
Management**

**Iranian Research Institute
for Science and Technology**

ISSN 2251-8223

eISSN 2251-8231

Indexed by SCOPUS, ISC, & LISTA

Vol. 33 | No. 2 | pp. 639-658

Winter 2018



واکنش موتورهای کاوش وب به پیشینه‌های فراداده‌ای مبتنی بر روش ترکیبی داده‌های خرد و روش داده‌های پیوندی

سیدمهدی طاهری

دکتری علم اطلاعات و دانش‌شناسی؛ استادیار؛ گروه
علم اطلاعات و دانش‌شناسی؛ دانشگاه علامه طباطبائی؛
پدیدآور رابط taherismster@gmail.com

رضا نیکزاد بهله

کارشناسی ارشد علم اطلاعات و دانش‌شناسی؛
r.nikzad20@gmail.com

میترا صمیعی

دکتری علم اطلاعات و دانش‌شناسی؛ استادیار؛ گروه
علم اطلاعات و دانش‌شناسی؛ دانشگاه علامه طباطبائی؛
samiei.mitra66@gmail.com



مقاله برای اصلاح به مدت ۱ روز نزد پدیدآوران بوده است.

پذیرش: ۱۳۹۵/۰۹/۲۸

دریافت: ۱۳۹۵/۰۹/۱۸

فصلنامه | علمی پژوهشی
پژوهشگاه علوم و فناوری اطلاعات ایران

شاپا (چاپی) ۸۲۳۳-۲۲۵۱

شاپا (الکترونیکی) ۸۲۳۱-۲۲۵۱

نمایه در SCOPUS و ISI، LISTA

jipm.irandoc.ac.ir

دوره ۳۳ | شماره ۲ | صص ۶۳۹-۶۵۸

زمستان ۱۳۹۶



چکیده: پژوهش حاضر با هدف تبیین واکنش موتورهای کاوش وب به پیشینه‌های فراداده‌ای مبتنی بر روش ترکیبی داده‌های خرد و روش داده‌های پیوندی با رویکرد تجربی انجام شد. جامعه پژوهش را ۲۰۰ پیشینه فراداده‌ای در قالب دو گروه گواه و آزمون با تعداد پیشینه برابر تشکیل می‌دادند. پیشینه‌های یادشده به روش نمونه‌گیری تصادفی و از میان پیشینه‌های فراداده‌ای دروازه اطلاعاتی کنسرسیوم محتوای ملی انتخاب شدند. در گروه آزمون، ۱۰۰ پیشینه فراداده‌ای انتخاب‌شده با رویکرد ترکیبی بر مبنای دو روش داده‌های خرد، به‌عنوان بهترین روش تولید خرده‌های غنی، و داده‌های پیوندی (قالب آردی‌اف/ایکس‌ام‌ال) ایجاد، و در گروه گواه، ۱۰۰ پیشینه بدون تغییر و با قالب اولیه آماده و بر روی وب‌سایت <http://www.mikzad.ir> منتشر شدند. سپس، وب‌سایت مزبور به‌صورت مستقیم و بر اساس روش‌های توصیه‌شده به موتورهای کاوش «گوگل» و «بینگ» معرفی شد. با استفاده از سیاهه و آرسی، وضعیت نمایه‌پذیری و پیدانمایی پیشینه‌های فراداده‌ای هر دو گروه در محیط موتورهای کاوش انتخابی مورد بررسی قرار گرفت. یافته‌های پژوهش حاکی از آن است که موتورهای کاوش «گوگل» و «بینگ» ارزش‌های عناصر پیشینه‌های فراداده‌های

گروه‌های گواه و آزمون را نمایه‌سازی کردند و پیشینه‌های مرتبط را با جست‌وجوی ارزش‌های عناصر فراداده‌ای در نتایج جست‌وجو پدیدار می‌نمایند، اما نمایش آن‌ها به شکل معنادار و همانند پیشینه‌های فراداده‌ای تولیدشده فقط مبتنی بر روش‌های تولید خرده‌های غنی نبود. به عبارت دیگر، نرم‌افزارهای خزنده-نمایه‌ساز موتورهای کاوش وب توانایی شناسایی و تشخیص برچسب‌های (عناصر) توصیه‌شده بر اساس روش داده‌های خرد را که در قالب چارچوب توصیف منبع (آردی‌اف)، به‌عنوان چهارچوب و بستر نحوی روش داده‌های پیوندی پیاده‌سازی شده‌اند، ندارند و به پیروی از آن، این پیشینه‌ها را در نتایج جست‌وجو به صورت معنادار نمایش نمی‌دهند. بنابراین، موتورهای کاوش وب با وجود تلاش برای ایجاد وب معنایی، سازگاری کامل با استانداردهای وب معنایی ندارند.

کلیدواژه‌ها: پیشینه‌های فراداده‌ای، نمایه‌پذیری، پیدانمایی، روش داده‌های خرد، روش داده‌های پیوندی، موتورهای کاوش وب

۱. مقدمه

شبکه وب به‌عنوان مهم‌ترین و بزرگ‌ترین بخش شبکه اینترنت از زمان پیدایش تاکنون، سه نسل گوناگون را تجربه نموده است. در نسل نخست، قابلیت‌های فرارسانه‌ای و فرایوندنمودن محتوا مد نظر قرار گرفت (طاهری ۱۳۹۲، ۵). نسل دیگر، به ایجاد بسترهای تعاملی بین انسان‌ها با استفاده از فناوری‌های وب و در محیط شبکه وب پرداخت (جمالی مهمویی ۱۳۸۲، ۵۱). نسل دو دیگر، بر این نکته تأکید داشت که داده‌های دسترس‌پذیر کنونی در محیط وب به دلیل عدم پیوند با داده‌های دیگر، داده‌های خام به شمار می‌آیند و برای معنادارشدن نیاز به پردازش (ساختارمندنمودن و ایجاد ارتباط میان آن‌ها) دارند (طاهری و همکاران ۱۳۹۴، ۴۶-۴۸). به عبارت دیگر، داده‌های ساختارمندی که با دیگر داده‌های منتشرشده روی وب پیوند یابند، بازیابی دانش را در پی خواهند داشت و افزون بر بهبود تعامل انسان و ماشین، موجب بهبود تعامل ماشین‌ها با یکدیگر می‌شوند. «برنرز-لی» مخترع وب، نام نسل جدید را وب معنایی نهاد و روش داده‌های پیوندی^۱ را برای تحقق آن توصیه نمود (Berners-Lee 2006, 5-7).

از سوی دیگر، برای جست‌وجو و دسترسی به داده‌های منتشرشده بر روی وب، ابزارها

۱. برای کسب اطلاعات بیشتر در خصوص روش داده‌های پیوندی به این مقاله رجوع شود:
سیدمهدی طاهری. ۱۳۹۲. داده‌های پیوندی: مفاهیم و استانداردها، با تأکید بر جامعه میراث فرهنگی.
فصلنامه اطلاع‌رسانی کنسرسیوم محتوای ملی ۱۲: ۴-۱۳.

و نظام‌های گوناگونی طراحی شدند. در میان این ابزارها و نظام‌ها، موتورهای کاوش به دلیل داشتن قابلیت‌های منحصربه‌فرد، بیش از دیگران مورد اقبال کاربران واقع شدند. کاربران با استفاده از کلیدواژه‌ها یا عبارات کلیدی به جست‌وجو در موتورهای کاوش می‌پردازند. این در حالی است که این کلیدواژه‌ها یا عبارات کلیدی نمی‌توانند به شیوه‌ای مناسب دربرگیرنده اطلاعات مورد نظر آن‌ها باشد و کاربر با حجم وسیعی از اطلاعات به‌عنوان نتایج بازیابی روبه‌رو می‌شود، زیرا موتورهای کاوش فقط امکان محدود نمودن محتوای درخواست بر اساس سه یا چهار برچسب را هنگام فرایند جست‌وجو امکان‌پذیر می‌سازند. بنابراین، عدم نمایش ساختارمند اطلاعات بازیابی‌شده، درک و تشخیص ارتباط آن‌ها با نیاز کاربر، و انتخاب گزینه‌های مناسب را دشوار می‌سازد. در صورتی که کاربران در اغلب موارد برای هر درخواستی به سراغ موتورهای کاوش می‌روند، این ویژگی، دسترس‌پذیر کردن اطلاعات از طریق موتورهای کاوش را برای فراهم‌کنندگان اطلاعات پیوسته ضروری می‌سازد (آزاد و شریف ۱۳۸۶، ۱۳۲؛ طاهری ۱۳۹۴). در چنین شرایطی، ابزارها و تمهیداتی مورد نیاز است که بتوانند درخواست کاربر را از نظر معنایی درک کرده و امکان جست‌وجوی معنایی را فراهم نمایند. از این رو، ساختارمند نمودن داده‌ها و فراداده‌ها گام مهمی به شمار می‌آید. با توجه به این که هدف هر موتور کاوش ارائه اطلاعات موجود در پایگاه داده‌ای خود به کاربران است، نمایه‌سازی و همچنین نمایش ساختارمند (معنادار) این اطلاعات توسط موتورهای کاوش برای کاربران، افزون بر بهبود ربط در فرایند بازیابی و افزایش رضایت کاربران، به تولید ارزش افزوده می‌انجامد. در راستای معنابخشی به محتوای قابل دسترس از طریق شبکه وب، تلاشی دیگر از سوی موتورهای کاوش صورت گرفته است. این موتورها با طراحی و توصیه روش‌های تولید خرده‌های غنی، در جهت نمایه‌سازی و پدیدارنمایی معنادار داده‌های وب گام برداشتند. این روش نیز همانند روش داده‌های پیوندی بر ساختارمند نمودن داده‌های ذخیره‌شده و پیوند دادن آن‌ها با داده‌های مرتبط تأکید می‌کند. پیاده‌سازی این روش و تولید پیشینه‌های فراداده‌ای مبتنی بر آن سبب شناخت دقیق موتور کاوش از موجودیت توصیف‌شده در صفحه وبی و نمایش معنادار اطلاعات در نتایج بازیابی می‌شد (Van der Meer et al 2011: www.schema.org).

با وجود این که تاکنون اقدامات گوناگونی برای ایجاد قابلیت معنایی در محیط وب توسط فراهم‌کنندگان اطلاعات و خدمات پیوسته صورت گرفته، اما در این میان، دوره کار

روش داده‌های پیوندی بر اساس ایده وب معنایی و تولید خرده‌های غنی به‌عنوان لایه معنایی نسخه جدید زبان نشانه‌گذاری فرامتن (اچ‌تی‌ام‌ال) که از سوی موتورهای کاوش توصیه شده، بیش از دیگران مورد توجه قرار گرفته است. بدیهی است شیوه مطلوب تولید فراداده‌های معنادار برای موتورهای کاوش، استفاده از روش ترکیبی برای تولید پیشینه‌هاست تا حداکثر بهره‌وری از قابلیت‌های معنایی تلاش‌های جدید صورت گیرد. بدین جهت، استانداردها و ابزارهایی چند از سوی «کنسرسیوم وب جهانی» برای تبدیل پیشینه‌های تولیدشده بر اساس روش‌های تولید خرده‌های غنی به‌روش داده‌های پیوندی طراحی و ارائه شده است. با توجه به این که هر دو راه کار بر معنابخشی به فرایندهای نمایه‌سازی و بازیابی اطلاعات تأکید دارند، مسئله‌ای که مطرح می‌شود این است که واکنش موتورهای کاوش وب به پیشینه‌های فراداده‌ای تولیدشده بر اساس روش ترکیبی (داده‌های پیوندی و خرده‌های غنی) چگونه خواهد بود؟ آیا نرم‌افزارهای خزنده-نمایه‌ساز موتورهای کاوش قادر به تشخیص عناصر فراداده‌ای به‌طور کامل (با نام و ارزش برجسب) و ارائه ساختارمند و معنادار آن‌ها در نتایج خود هستند؟ پژوهش حاضر، برای پاسخگویی به این مسئله طرح‌ریزی شده است. نمایه‌سازی و پیدانمایی^۱ معنادار اشیاء محتوایی توسط موتورهای کاوش، به دلیل معنابخشی و نمایش معنادار اطلاعات مرتبط با یکدیگر موجب افزایش رضایت کاربران نهایی و درک بیشتر آن‌ها از محتوای بازیابی شده می‌گردد. این مهم همواره دغدغه طراحان موتورهای کاوش وب بوده است. طراحی تمهیدات گوناگون و انجام اقدامات گسترده در این خصوص از سوی طراحان یادشده دلیلی بر این مدعاست. از سوی دیگر، وب در حال تجربه نسل جدید خود، یعنی وب معنایی (یا وب ۳.۰) است. برای تحقق وب معنایی، روش داده‌های پیوندی از سوی مخترع وب پیشنهاد شده و مورد استقبال گسترده فراهم‌کنندگان اطلاعات پیوسته واقع شده است. این در حالی است که موتورهای کاوش نیز خرده‌های غنی را به‌عنوان گامی معنابخش توصیه می‌نمایند. استفاده از هر دو روش به‌صورت ترکیبی، زمینه را برای هم‌راست نمودن هر دو تلاش برای تحقق وب معنایی هموار می‌سازد. در عین حال، پاسخ مثبت موتورهای کاوش وب به پیشینه‌های ایجادشده بر پایه روش داده‌های پیوندی، بیانگر اقبال موتورهای کاوش به‌روش داده‌های پیوندی و همسویی و ایفای نقش فعال در تحقق وب معنایی خواهد بود.

1. visibility

۴. هدف و پرسش‌های پژوهش

هدف اصلی این پژوهش، تبیین واکنش موتورهای کاوش وب به پیشینه‌های فراداده‌ای مبتنی بر روش ترکیبی داده‌های خرد و روش داده‌های پیوندی است. به منظور دستیابی به این هدف چهار پرسش طراحی شد:

۱. وضعیت نمایه‌سازی پیشینه‌های فراداده‌ای مبتنی بر روش ترکیبی داده‌های خرد و داده‌های پیوندی توسط موتور کاوش «گوگل» چگونه است؟
۲. وضعیت نمایه‌سازی پیشینه‌های فراداده‌ای مبتنی بر روش ترکیبی داده‌های خرد و داده‌های پیوندی توسط موتور کاوش «بینگ» چگونه است؟
۳. وضعیت پیدانمایی (نمایش ساختارمند) پیشینه‌های فراداده‌ای مبتنی بر روش ترکیبی داده‌های خرد و داده‌های پیوندی توسط موتور کاوش «گوگل» چگونه است؟
۴. وضعیت پیدانمایی (نمایش ساختارمند) پیشینه‌های فراداده‌ای مبتنی بر روش ترکیبی داده‌های خرد و داده‌های پیوندی توسط موتور کاوش «بینگ» چگونه است؟

۳. مرور پیشینه‌های پژوهش

بهبود عملکرد موتورهای کاوش و واکنش آن‌ها نسبت به فراداده‌ها موجب افزایش و سهولت دسترسی به اطلاعات موجود در شبکه وب خواهد شد. به منظور نیل به این هدف، همواره اقبال ویژه‌ای از سوی طراحان موتورهای کاوش و طراحان نظام‌ها و استانداردهای فراداده‌ای به افزایش میان‌کنش‌پذیری این دو ابزار بازیابی محتوای وب وجود داشته است. در همین راستا، تاکنون پژوهش‌های گوناگونی در حوزه میان‌کنش‌پذیری موتورهای کاوش و نظام‌های فراداده‌ای انجام شده است. هدف مشترک تمامی این پژوهش‌ها تبیین وضعیت نمایه‌سازی و پیدانمایی پیشینه‌های فراداده‌ای توسط موتورهای کاوش و ارائه راه‌کارهایی برای بهبود وضعیت موجود بوده است. با مرور این پژوهش‌ها، می‌توان آن‌ها را به سه گروه کلی تقسیم نمود. مبنای این گروه‌بندی، عناصر (برچسب‌های) مورد استفاده برای توصیف و بازنمون خصایص اشیاء محتوایی هستند، زیرا تعامل موتورهای کاوش با پیشینه‌های فراداده‌ای و به پیروی از آن اشیاء محتوایی مرتبط از طریق این عناصر صورت می‌گیرد. گروه نخست، پژوهش‌هایی را دربرمی‌گیرد که به مطالعه واکنش موتورهای کاوش به برچسب‌ها و فرابرجسب‌های زبان نشانه‌گذاری فرامتن (اچ‌تی‌ام‌ال) پرداخته‌اند. گروه دیگر، تعامل موتورهای کاوش با نظام‌های فراداده‌ای دارای پیشینه‌های ایجادشده بر

مبنای استانداردهای فراداده‌ای را مورد مطالعه قرار داده‌اند و گروه پایانی، تأثیر استفاده از روش‌های تولید خرده‌های غنی بر بهبود نمایه‌پذیری و پیدانمایی پیشینه‌های فراداده‌ای توسط موتورهای کاوش را ارزیابی نموده است.

در گروه نخست که بستر نحوی عناصر فراداده‌ای آن‌ها زبان نشانه‌گذاری فرامتن (اچ‌تی‌ام‌ال) بود، صفحات یا اشیاء محتوایی جامعه پژوهش، برای توصیف یا سازماندهی محتوای خود از همان برچسب‌ها و فرابرچسب‌های بستر نحوی استفاده نمودند. زبان نشانه‌گذاری فرامتن به‌عنوان بهترین قالب نمایش صفحات وب که از اقبال ویژه‌ای در بین فراهم‌کنندگان اطلاعات پیوسته برخوردار است، افزون بر برچسب‌هایی برای ذخیره (بیشتر برای قابلیت‌های نمایشی) محتوا، فرابرچسب‌هایی برای توصیف محتوای ذخیره‌شده در هر صفحه ارائه می‌کند. پژوهش‌های این گروه به مطالعه تأثیر به‌کارگیری برچسب‌ها و فرابرچسب‌های اچ‌تی‌ام‌ال بر بهبود نمایه‌پذیری و بازیابی‌پذیری اشیاء محتوایی پرداختند. برخی از این پژوهش‌ها مانند پژوهش‌های (Quevedo-Torrero, and Ubaldo (2004); Turmer and Brackbill (1998) و Zhang, and Dimitroff (2005a) بر اثربخشی برچسب‌ها و فرابرچسب‌های زبان نشانه‌گذاری یادشده بر بهبود بازیابی و پیدانمایی اشیاء محتوایی وب متمرکز بودند. نتایج این پژوهش‌ها واکنش مثبت موتورهای کاوش به برچسب‌ها و فرابرچسب‌های زبان نشانه‌گذاری فرامتن را نشان می‌داد و بهبود بازیابی و پیدانمایی اشیاء محتوایی مورد مطالعه قابل مشاهده بود.

گروه دیگر پژوهش‌هایی بودند که کارآمدی استفاده از عناصر یا پیشینه‌های فراداده‌ای مبتنی بر طرح‌های فراداده‌ای استاندارد مانند طرح فراداده‌ای هسته «دوبلین» طرح فراداده‌ای توصیف‌شده و قالب فراداده‌ای مارک در بهبود بازیابی‌پذیری و رتبه‌بندی اشیاء محتوایی مرتبط در موتورهای کاوش را مورد مطالعه قرار دادند. گرایش به بررسی تأثیر عناصر فراداده‌ای استاندارد، جامعیت و بستر معناشناختی استانداردهای فراداده‌ای برای سازماندهی اشیاء محتوایی وب در مقابل محدودیت‌های برچسب‌ها و فرابرچسب‌های اچ‌تی‌ام‌ال بود. بستر نحوی پیشینه‌های فراداده‌ای مورد مطالعه در این گروه زبان‌های نشانه‌گذاری فرامتن و گسترش‌پذیر (ایکس‌ام‌ال) بودند. یافته‌های این پژوهش‌ها به‌دلیل استفاده از رویکردهای گوناگون در ایجاد پیشینه‌های فراداده‌ای مورد بررسی متفاوت بود. برخی به مقایسه برچسب‌ها و فرابرچسب‌های اچ‌تی‌ام‌ال با عناصر فراداده‌ای استاندارد توجه داشتند. نتایج این پژوهش‌ها تأثیر بهتر برچسب‌ها و فرابرچسب‌های اچ‌تی‌ام‌ال نسبت

به عناصر فراداده‌ای استاندارد بر بهبود بازیابی اشیاء محتوایی توسط موتورهای کاوش را نشان می‌دادند (Henshaw and Valauskas 2001؛ Zhang, and Dimitroff (2004)؛ Zhang, and Mohamed 2006؛ Dimitroff (2005b)؛ شریف ۱۳۸۶). برخی دیگر، بررسی بازیابی‌پذیری و رتبه‌بندی اشیاء محتوایی حاوی عناصر فراداده‌ای استاندارد مربوط به طرح فراداده‌ای هسته «دوبلین» مبتنی بر زبان نشانه‌گذاری فرامتن را انجام دادند. این پژوهش‌ها اذعان داشتند که بهره‌گیری از عناصر فراداده‌ای استاندارد جاسازی‌شده در اشیاء محتوایی مبتنی بر اچ‌تی‌ام‌ال، تأثیر چندانی بر بهبود بازیابی و رتبه‌بندی در محیط موتورهای کاوش وب ندارد (Sokvine 2000؛ Safari 2005). دسته‌پایانی در این گروه، پژوهش‌هایی بودند که بر پیاده‌سازی پیشنهادهاى فراداده‌ای مورد مطالعه در بستر نحوی اصلی طرح‌های فراداده‌ای استاندارد، یعنی زبان نشانه‌گذاری گسترش‌پذیر (ایکس‌ام‌ال) متمرکز بودند. باور این پژوهش‌ها بر آن بود که بررسی درست واکنش موتورهای کاوش به عناصر و پیشنهادهاى فراداده‌ای باید نسبت به بستر نحوی منتخب استانداردهای فراداده‌ای مورد ارزیابی قرار گیرد. نتایج این پژوهش‌ها بیانگر بهبود نمایه‌پذیری عناصر و پیشنهادهاى فراداده‌ای مبتنی بر ایکس‌ام‌ال توسط نرم‌افزارهای خزنده-نمایه‌سازی موتورهای کاوش و پیدانمایی آنها در نتایج جست‌وجو بود (طاهری، حریری و فتاحی، ۱۳۸۸؛ طباطبائی امیری و همکاران ۱۳۹۱؛ باب‌الحوائجی، طاهری و آقاعابدی ۱۳۹۴). به‌ویژه پژوهش‌های «طاهری» و همکاران نشان داد که پیاده‌سازی پیشنهادهاى فراداده‌ای در بستر نحوی زبان نشانه‌گذاری گسترش‌پذیر (بر اساس ساختاری خاص)، موجب نمایه‌پذیری نام‌های برچسب عناصر فراداده‌ای و امکان جست‌وجوی اشیاء محتوایی از طریق آنها (جست‌وجوی عنصر-پایه) فراهم خواهد شد (Taheri, Hariri, and Fattahi 2014a; 2014b).

مطالعه‌ کارایی و اثربخشی ایجاد پیشنهادهاى فراداده‌ای بر اساس جدیدترین روش‌های توصیه‌شده از سوی طراحان موتورهای کاوش، رویکرد پژوهش‌های گروه پایانی بوده است. در این گروه از پژوهش‌ها پژوهشگران قصد داشتند وضعیت نمایه‌پذیری و نمایش معنادار برچسب‌های (عناصر) فراداده‌ای مبتنی بر روش‌های تولید خرده‌های غنی را ارزیابی نمایند. روش‌های ایجاد خرده‌های غنی عبارت‌اند از داده‌های خرد، قالب‌های خرد، و چارچوب توصیف منبع (خصایص) که در بین آنها روش داده‌های خرد بیش از دیگر روش‌ها توصیه شده است. برچسب‌ها (عناصر)، خصایص، و ارزش‌های آنها در روش‌های تولید خرده‌های غنی با نمونه‌های معادل خود در طرح‌های فراداده‌ای استاندارد

بافت میراث فرهنگی (کتابخانه، موزه و آرشیو) متفاوت هستند و در اصل، گسترشی از جنبه معنانشاخی نسخه جدید (نسخه ۵/۰) زبان نشانه‌گذاری فرامتن به شمار می‌آیند. با وجود این، از کارکردهای مشابه طرح‌های فراداده‌ای استاندارد پشتیبانی می‌کنند و می‌توان آن‌ها را استانداردهای فراداده‌ای جدید دانست. نتایج پژوهش‌های این گروه حاکی از آن است که درج برچسب‌های (عناصر) فراداده‌ای مبتنی بر روش‌های داده‌های خرد، قالب‌های خرد، و چارچوب توصیف منبع (خصایص) افزون بر بهبود نمایه‌پذیری آن‌ها توسط موتورهای کاوش، آن‌هم به صورت معنایی، نمایش معنادار پیشینه‌های فراداده‌ای و به پیروی از آن اشیاء محتوایی را در پی خواهد داشت. به عبارت دیگر، به دلیل آن‌که این روش‌ها از سوی موتورهای کاوش وب توصیه می‌شوند و برخی مانند روش داده‌های خرد توسط توسعه‌دهندگان موتورهای کاوش طراحی شده‌اند، از اقبال بیشتری از سوی این موتورها برخوردارند (Mixer 2013؛ Isaksen 2011؛ طاهری، ذوالقدر و حریری ۱۳۹۵). البته، پژوهش‌های «طاهری» و همکاران اعمال تغییراتی محدود با رعایت جنبه‌های فنی استاندارد بر روی پیشینه‌های فراداده‌ای استاندارد مبتنی بر زبان نشانه‌گذاری گسترش‌پذیر را نسبت به پیشینه‌های تولیدشده بر پایه روش‌های تولید خرده‌های غنی در بهبود بازیابی‌پذیری اشیاء محتوایی در محیط وب کارآمدتر بیان می‌دارد (Taheri, Hariri, and Fattahi 2014a; 2014b).

مرور پیشینه‌های مربوط به حوزه میان‌کنش‌پذیری نظام‌های فراداده‌ای و موتورهای کاوش نشانگر آن است که به‌طور کلی موتورهای کاوش وب با وجود مزایای عناصر یا پیشینه‌های فراداده‌ای ایجادشده بر اساس طرح‌های فراداده‌ای استاندارد، که حاصل تلاش‌ها و تجربه بیش از نیم قرن بافت میراث فرهنگی است، واکنش بهتری به برچسب‌های (عناصر) یا پیشینه‌های فراداده‌ای مبتنی بر روش‌های تولید فراداده یا توصیه‌شده توسط طراحان موتورهای کاوش نشان می‌دهند؛ اگرچه پیاده‌سازی عناصر یا پیشینه‌های فراداده‌ای استاندارد که بستر نحوی ایکس‌ام‌ال را برای پیاده‌سازی برگزیدند، پاسخ مناسب‌تری نسبت به اچ‌تی‌ام‌ال از سوی نرم‌افزارهای خزنده-نمایه‌ساز موتورهای کاوش، به‌ویژه وقتی ساختاری خاص بر روی آن پیشینه‌ها اعمال می‌شود، دریافت نمودند. از سوی دیگر، مخترع وب، «تیم برنرز-لی» و «کنسرسیوم وب جهانی» روشی (راهبردی) به نام داده‌های پیوندی را برای تحقق وب معنایی توصیه می‌کند که بیشتر به بستر نحوی و ساختار پیشینه‌های فراداده‌ای استاندارد نزدیک است. نیز «کنسرسیوم وب جهانی» ابزارهایی

برای تبدیل پیشینه‌های مبتنی بر روش‌های تولید خرده‌های غنی به بستر نحوی روش داده‌های پیوندی، یعنی آردی اف طراحی نموده است. بدیهی است اگر موتورهای کاوش تمهیدات مناسبی برای نمایه‌سازی و پیدانمایی معنادار پیشینه‌های تولیدشده مبتنی بر هر دو روش تولید خرده‌های غنی و روش داده‌های پیوندی بیندیشند، گام مؤثر و بزرگی در تحقق وب معنایی برداشته‌اند. اما تاکنون پژوهشی برای بررسی وجود چنین تمهیداتی صورت نگرفته است و گمان می‌رود انجام چنین پژوهشی رویکرد کنونی موتورهای کاوش نسبت به استانداردهای وب معنایی را بازنمایاند.

۴. روش پژوهش

پژوهش حاضر از نظر نوع، یک پژوهش کاربردی محسوب می‌شود، زیرا به حل مسئله ایجاد پیشینه‌های فراداده‌ای مبتنی بر رویکرد ترکیبی روش تولید خرده‌های غنی و داده‌های خرد با قابلیت نمایه‌پذیری و پیدانمایی ساختارمند (معنادار) در محیط موتورهای کاوش وب پرداخته است. برای انجام این پژوهش از روش تجربی استفاده شده است. جامعه پژوهش دارای ۲۰۰ پیشینه فراداده‌ای مربوط به موجودیت‌های کتابی موجود در دروازه اطلاعاتی کنسرسیوم محتوای ملی بود که از کتابخانه دیجیتال «دانشگاه علامه طباطبائی (ره)»، کتابخانه دیجیتالی «دانشگاه تهران»، «سازمان کتابخانه‌ها، موزه‌ها و مرکز اسناد آستان قدس رضوی» و کتابخانه دیجیتالی «دفتر تبلیغات اسلامی حوزه علمیه قم» در دامنه‌های موضوعی «علم اطلاعات و دانش‌شناسی»، «روان‌شناسی»، «پزشکی»، و «حقوق» بنا بر سلیقه پژوهشگران انتخاب شد. این پیشینه‌ها به دو گروه گواه و آزمون با تعداد پیشینه‌های یکسان تقسیم گردیدند. پیشینه‌های گروه نخست (گواه)، بر اساس روش داده‌های خرد ایجاد شدند. اما پیشینه‌های گروه آزمون بر مبنای رویکرد ترکیبی، دو روش داده‌های خرد و داده‌های پیوندی تولید گردیدند. بدین صورت که نخست، پیشینه‌های فراداده‌ای انتخاب‌شده بر مبنای توصیه‌ها و ساختارهای الگوی داده‌های خرد، به پیشینه‌های مبتنی بر این روش تبدیل شدند (نمونه ۱)، و سپس، با استفاده از ابزار بیان‌شده در زیر، به قالب پیشنهادی روش داده‌های پیوندی (قالب آردی اف. / ایکس ام ال) تغییر بستر نحوی دادند. بنابراین، پیشینه‌های جدید، بستر نحوی و قالب داده‌های خرد را رعایت نمودند. همچنین، برای بررسی واکنش موتورهای کاوش وب به پیشینه‌های ایجادشده، دو موتور کاوش «گوگل» و «بینگ» به دلیل گستردگی کاربرد و جایگاه آن‌ها در میان دیگر موتورهای

کاوش وب انتخاب شدند.

به منظور اعتبارسنجی تمامی پیشینه‌های فراداده‌ای ایجادشده از بُعد رعایت قابلیت‌های فنی روش داده‌های خرد، از ابزار^۱ Structured Data Testing Tools بخش tools webmaster موتور کاوش «گوگل» استفاده شد. پس از اطمینان از صحت و درستی هر یک از پیشینه‌های مورد مطالعه، برای تبدیل پیشینه‌های فراداده‌ای گروه آزمون از روش داده‌های خرد به روش داده‌های پیوندی (قالب آردی.اف./ایکس.ام.ال)، از ابزار Microdata to RDF Distiller توسعه‌یافته و دسترس‌پذیر در وب‌سایت «کنسرسیوم جهانی وب» استفاده شد. سپس، پیشینه‌های هر دو گروه به صورت مستقیم و بر اساس بهترین روش‌های توصیه‌شده (XML sitemap) به موتورهای کاوش وب برگزیده معرفی شدند. نمونه‌هایی از پیشینه‌های مربوط به هر دو گروه آزمون و گواه در زیر ارائه شده است:

Little Water Cantina - Eastlake - Seattle, WA
www.yelp.com | Restaurants | Mexican
★★★★☆ 90 reviews | Price range: \$\$
90 Reviews of **Little Water Cantina**: "Three things are on my list when I eat out: great food, atmosphere, and..."

Vegetarian Vegan Pizza No Cheese) Recipe - Food.com - 248865
www.food.com/recipe/vegetarian-vegan-pizza-no-c...
★★★★☆ 2 reviews - 1 hr 32 mins - 242.9 cal
Aug 26, 2007 - This is from my dad, who developed some **vegan recipes**... doesn't have any cheese, and you

Leonard Cohen - Free listening, videos, concerts, stats, & pictures at ...
www.last.fm/music/Leonard+Cohen
Watch videos & listen to **Leonard Cohen**: Suzanne, Hallelujah & more, plus 132 pictures. **Leonard Cohen**, (born September 21, 1934 in Montréal, Québec, ...)

Track	Duration
Suzanne	3:48
The Darkness	4:29
Going Home	3:51
Hallelujah	4:12

تصویر ۱. تصویری از نمایش ساختارمند پیشینه‌های فراداده‌ای توسط موتور کاوش «گوگل» به‌عنوان هدف تولید و انتشار پیشینه‌های فراداده‌ای مورد مطالعه در پژوهش حاضر

۱. ابزاری برای بررسی و اعتبارسنجی داده‌های ساختارمند تولیدشده بر اساس الگوی داده‌های خرد که توسط موتور کاوش «گوگل» طراحی شده است. برای اطلاعات بیشتر در مورد این امکان، به نشانی اینترنتی زیر مراجعه شود:

http://www.w3.org/2012/pyMicrodata

```
<div itemscope itemtype="http://schema.org/Book"><meta itemprop="bookFormat"
content="الکترونیکی"/><d1><dt><dd itemprop="name">اخلاق پزشکی
<dt><dd itemprop="author">منصور اشرفی</dd><dt><dd
itemprop="translator"></dt><dt><dd itemprop="description">
اخلاق پزشکی</dd><dt><dd><span itemprop="number of pages">328</span>pages/<
dd><dt><dd itemprop="ISBN"></dd>
<dt><dd itemsprop="publisher">دانشگاه آزاد اسلامی تبریز</dd>
<dt><dd itemprop="copyrightyear">1367</dd><dt><dd>صاحب محتوا:
<dd itemprop="copyrightholder">سازمان کتابخانه‌ها، موزه‌ها و مرکز اسناد آستان قدس
</dd>
<dt><dd itemprop="creator">تاریخ ایجاد پیشینه: نیکزاد</dd><dt><dd
itemprop="datecreated">1394/3/20</dd><dt><dd
itemprop="inLanguage">فارسی</dd><dt><dd itemprop="genre">کتاب</dd>
</d1></div>
```

نمونه ۱. نمونه‌ای از یک پیشینه فراداده‌ای مبتنی بر الگوی داده‌های خرد

```
<?xml version="1.0" encoding="utf-8"?><rdf:RDF xmlns:rdf="http://www.
w3.org/1999/02/22-rdf-syntax-ns#" xmlns:schema="http://schema.org/"> <schema:Book>
<schema:copyrightyear>1367</schema:copyrightyear><schema:author>منصور اشرفی</
schema:author><schema:ISBN></schema:ISBN> <schema:bookFormat>الکترونیکی</
schema:bookFormat><schema:description>اخلاق پزشکی</
schema:description><schema:datecreated>1394/3/20</schema:datecreated>
<schema:copyrightholder>سازمان کتابخانه‌ها، موزه‌ها و مرکز اسناد آستان قدس رضوی</
schema:copyrightholder><schema:genre>کتاب</schema:genre> <schema:name>اخلاق</
schema:name><schema:creator>نیکزاد</schema:creator>
<schema:inLanguage>فارسی</schema:inLanguage><schema:translator>
</schema:translator></schema:Book></rdf:RDF>
```

نمونه ۲. نمونه‌ای از یک پیشینه فراداده‌ای مبتنی بر روش ترکیبی داده‌های خرد و داده‌های پیوندی
(آردی‌اف / ایکس‌ام‌ال)

در مرحله اجرای پژوهش، روش گردآوری داده‌ها مشاهده طراحی شده^۱ و ابزار آن سیاهه واریسی بود که بر پایه پرسش‌های پژوهش و توسط پژوهشگران طراحی گردید. گردآوری داده‌ها در خرداد ماه ۱۳۹۵ صورت گرفت. به منظور استخراج داده‌ها از راهبرد جست‌وجوی "keyphrase" در site:rnikzad.ir در موتورهای کاوش مورد مطالعه استفاده شد. برای تجزیه و تحلیل داده‌های گردآوری شده نیز بر مبنای آمار توصیفی، از نرم‌افزار «اس‌پی‌اس‌اس» بهره‌گیری شد و برای نمایش آن‌ها با توجه به پرسش‌های پژوهش، جدول‌هایی متناسب با داده‌های درج شده در سیاهه واریسی استخراج شد.

۵. یافته‌های پژوهش

پژوهش حاضر برای نیل به اهداف خود، چهار پرسش مطرح نمود. برای پاسخ به پرسش‌های اول و دوم، جدول شماره ۱، و برای پاسخ به پرسش‌های سوم و چهارم، جدول شماره ۲، طراحی شد. داده‌های گردآوری شده برای پاسخ به پرسش‌های پژوهش تصادفی بودند. پیشینه‌های فراداده‌ای هر دو گروه گواه و آزمون به مدت شش ماه روی وبسایت ایجاد شده دسترس‌پذیر گردیدند، و فرایند گردآوری داده‌ها برای اطمینان از صحت آن‌ها یک ماه به طول انجامید.

پرسش‌های اول و دوم پژوهش:

- ◇ وضعیت نمایه‌سازی پیشینه‌های فراداده‌ای مبتنی بر روش ترکیبی داده‌های خرد و داده‌های پیوندی توسط موتور کاوش «گوگل» چگونه است؟
- ◇ وضعیت نمایه‌سازی پیشینه‌های فراداده‌ای مبتنی بر روش ترکیبی داده‌های خرد و داده‌های پیوندی توسط موتور کاوش «بینگ» چگونه است؟

جدول ۱. وضعیت نمایه‌سازی پیشینه‌های فراداده‌ای گروه‌های گواه و آزمون توسط موتورهای کاوش «گوگل» و «بینگ»

گروه مورد مطالعه	تعداد کل پیشینه‌ها	تعداد		درصد	
		گوگل	بینگ	گوگل	بینگ
گواه	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰
آزمون	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰

چنان‌که داده‌های درج‌شده در جدول ۱، نشان می‌دهد، تمامی پیشینه‌های فراداده‌ای هر دو گروه گواه و آزمون توسط موتورهای کاوش «گوگل» و «بینگ» نمایه‌سازی شده‌اند. به عبارت دیگر، نرم‌افزارهای خزنده-نمایه‌ساز دو موتور کاوش مورد مطالعه، نسبت به عناصر (برچسب‌های) پیشینه‌های فراداده‌ای گروه‌های گواه و آزمون واکنش مثبت نشان داده‌اند. بنابراین، پیاده‌سازی پیشینه‌های فراداده‌ای مبتنی بر روش داده‌های خرد در بستر نحوی آردی‌اف، به‌عنوان چارچوب توصیه‌شده در روش داده‌های پیوندی، از ارزش نمایه‌ای برابری در مقایسه با پیشینه‌های فقط مبتنی بر روش داده‌های خرد برخوردار هستند، اگرچه ویژگی نمایه‌شدن معنادار پیشینه‌های فراداده‌ای مورد مطالعه بدون بازیابی آن‌ها و نمایش ساختارمند در نتایج جست‌وجو امکان‌پذیر نیست. بنابراین، داده‌ها بیانگر نمایه‌شدن ارزش‌های عناصر پیشینه‌های فراداده‌ای گروه آزمون هستند و درک معنادار عناصر و صفات^۱ آن‌ها توسط موتورهای کاوش مشخص نیست.

پرسش‌های سوم و چهارم پژوهش:

- ◇ وضعیت پیدانمایی (نمایش ساختارمند) پیشینه‌های فراداده‌ای مبتنی بر روش ترکیبی داده‌های خرد و داده‌های پیوندی توسط موتور کاوش «گوگل» چگونه است؟
- ◇ وضعیت پیدانمایی (نمایش ساختارمند) پیشینه‌های فراداده‌ای مبتنی بر روش ترکیبی داده‌های خرد و داده‌های پیوندی توسط موتور کاوش «بینگ» چگونه است؟

1. properties

جدول ۲. وضعیت پیدانمایی (نمایش ساختارمند) پیشنهادی گروه‌های گواه و آزمون توسط موتورهای کاوش «گوگل» و «بینگ»

گروه مورد مطالعه	تعداد کل پیشنهادها	تعداد		درصد	
		گوگل	بینگ	گوگل	بینگ
گواه	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰
آزمون	۱۰۰	۰	۰	۰	۰

بر اساس داده‌های جدول ۲، وضعیت پیدانمایی پیشنهادی گروه‌های فراداده‌ای گروه آزمون با پیشنهادی‌های فراداده‌ای گروه گواه کاملاً متفاوت است. موتورهای کاوش «گوگل» و «بینگ» پیشنهادی‌های فراداده‌ای گروه گواه را به صورت ساختارمند و معنایی در نتایج جست‌وجو نشان می‌دهند، اما پیشنهادی‌های فراداده‌ای گروه آزمون را با وجود نمایه‌سازی کامل، به صورت معمولی در فهرست نتایج پدیدار می‌نمایند. بدین معنا که واکنش موتورهای کاوش «گوگل» و «بینگ» نسبت به دو گروه از پیشنهادها یکسان نیست و استفاده از استانداردهای روش داده‌های پیوندی بر روی نمایش ساختارمند پیشنهادی‌های فراداده‌ای مبتنی بر خرده‌های غنی تأثیر منفی می‌گذارد. به عبارت دیگر، نرم‌افزارهای خزنده-نمایه‌ساز موتورهای کاوش «گوگل» و «بینگ» فقط عناصر و صفات پیشنهادی‌های مبتنی بر روش داده‌های خرد را که از سوی این موتورهای طراحی و توصیه شده است، به صورت معنادار تشخیص داده و نمایه‌سازی نموده‌اند، اما قادر به درک عناصر و صفات پیشنهادی‌های گروه آزمون نبوده‌اند.

۶. نتیجه‌گیری

همان‌طور که «برنرزی» بیان داشت، داده‌های منتشرشده کنونی بر روی وب، داده‌های خام هستند و صرفاً توصیف ویژگی‌های آن‌ها بدون برقراری ارتباط میان داده‌ها پردازش کاملی نخواهد بود (Berners-Lee 2006). پیوند میان داده‌های مرتبط موجب ایجاد بافت، و به دنبال آن تولید معنا خواهد شد. برای تغییر وب کنونی به وب معنایی تلاش‌های گوناگونی از سوی طراحان نظام‌های مبتنی بر وب انجام شده است.

افزون بر روش داده‌های پیوندی که توسط «تیم برنرز-لی» پیشنهاد گردید، توسعه نسخه ۵/۰ زبان نشانه‌گذاری فرامتن، و به‌ویژه روش‌های تولید خرده‌های غنی نیز حرکتی در راستای تولید معنا در محیط وب به شمار می‌آیند. بدیهی است تحقق وب معنایی وابسته به سازگاری این روش‌ها با یکدیگر و کوشش در این راستا خواهد بود. یافته‌های پژوهش حاضر که به‌منظور تبیین واکنش پرکاربردترین ابزارهای جست‌وجو و دسترسی به داده‌های وب، یعنی موتورهای کاوش نسبت به پیشینه‌های فراداده‌ای ایجادشده بر اساس روش ترکیبی داده‌های خرد و داده‌های پیوندی انجام شد، حاکی از آن است که موتورهای کاوش علاقه بیشتری به نمایه‌سازی و پیدانمایی معنادار پیشینه‌های فراداده‌ای مبتنی بر داده‌های خرد، که از سوی طراحان آن‌ها گسترش و توصیه شده، دارند و الگوریتم‌های مناسبی برای شناسایی عناصر و صفات پیشینه‌های فراداده‌ای به‌صورت معنادار، و به پیروی از آن نمایه‌سازی و نمایش ساختارمند آن‌ها در نتایج جست‌وجو طراحی نموده‌اند. این مهم در پیدانمایی ساختارمند پیشینه‌های فراداده‌ای تولیدشده بر پایه روش داده‌های خرد (گروه گواه) نمایان گردید. چنین واکنش مثبتی پیش‌تر در یافته‌های پژوهش‌های Turner (1998) and Brackbill (2004)، Quevedo-Torrero and Ubaldo (2004)، Zhang and Dimitroff (2005b)، Zhang, and Dimitroff (2005a)، Henshaw and Valauskas (2001)، Zhang and Dimitroff (2004)، Mohamed (2006) و «شریف» (۱۳۸۶) نسبت به روش‌های (یا فرابرچسب‌های) توصیه‌شده از سوی موتورهای کاوش ذکر گردیده بود. اما در خصوص پیشینه‌های فراداده‌ای تولیدشده بر پایه روش ترکیبی، وضعیت نمایه‌سازی و پیدانمایی توسط موتورهای کاوش متفاوت بود. با وجود این که نمایه‌پذیری پیشینه‌های فراداده‌ای مبتنی بر روش ترکیبی به‌دلیل استفاده از بستر نحوی زبان نشانه‌گذاری گسترش‌پذیر (ایکس‌ام‌ال) در سطح نمایه‌پذیری ارزش‌های عناصر فراداده‌ای شبیه پیشینه‌های فراداده‌ای بر اساس روش داده‌های خرد بود، و پژوهش‌های «طاهری، حریری و فتاحی» (۱۳۸۸)، طباطبائی امیری و همکاران (۱۳۹۱)، «باب‌الحوائجی، طاهری و آقاعبدی» (۱۳۹۴)، و Taheri, Hariri and Fattahi (2014a; 2014b) نیز یافته‌های یکسانی داشتند، اما این پیشینه‌ها در فهرست نتایج جست‌وجو به‌صورت ساختارمند و معنادار پدیدار نشدند. به‌عبارت دیگر، استفاده از روش داده‌های خرد برای ایجاد عناصر یا پیشینه‌های فراداده‌ای در بستر استانداردهای روش داده‌های پیوندی باعث تولید خرده‌های غنی در نتایج جست‌وجوی موتورهای کاوش نخواهد شد و موتورهای یادشده این پیشینه‌ها را به‌صورت معنادار نمایش نمی‌دهند. این یافته بیانگر عدم سازگاری

دو روش مهم تولید معنا در محیط وب است. بدین معنا که نوعی واگرایی در تلاش‌های صورت گرفته از سوی طراحان اصلی استانداردها و خدمات وب به چشم می‌خورد. این واگرایی سردرگمی توسعه‌دهندگان اطلاعات پیوسته در وب را موجب می‌شود. در حال حاضر بسیاری از این طراحان در حال پیاده‌سازی روش داده‌های پیوندی که از سوی «کنسرسیون وب» توصیه شده، در نظام‌های اطلاعاتی خود هستند. اگر داده‌های این نظام‌ها حتی با بهره‌گیری از روش‌های تولید خرده‌های غنی به سطح پدیدارنمایی معنایی در نتایج جست‌وجوی موتورهای کاوش نرسد، چالشی بزرگ در دسترس‌پذیری محتوای وب ایجاد خواهد شد. زیرا عدم سازگاری میان روش‌های توصیه شده از سوی «کنسرسیون وب» به‌عنوان مرجع اصلی طراحی و تدوین استانداردهای وب و الگوها و روش‌های توصیه شده از سوی موتورهای کاوش، به‌عنوان پرکاربردترین و محبوب‌ترین ابزارهای کاوش وب را نشان می‌دهد. به نظر می‌رسد طراحی الگوریتم‌های مناسب برای شناسایی استاندارد روش داده‌های پیوندی یعنی چارچوب توصیف منبع، به‌ویژه در بستر نحوی ایکس‌ام‌ال از سوی طراحان موتورهای کاوش، برای نمایش ساختارمند پیشینه‌های فراداده‌ای نمایه شده راه کار مناسبی در جهت همراستایی تلاش‌های توسعه وب معنایی خواهد بود. بر اساس نتایج پژوهش حاضر، انجام پژوهش‌هایی بر روی تبیین واکنش موتورهای کاوش به داده‌های مستند جاسازی شده در پیشینه‌های مبتنی بر روش‌های تولید خرده‌های غنی، و بررسی استفاده از الگوهای مفهومی و استانداردهای مدیریت اطلاعات و دانش بافت میراث فرهنگی برای بهبود ساختار و اثربخشی روش‌های تولید خرده‌های غنی در طراحی پیشینه‌های فراداده‌ای با نمایه‌پذیری و پیدانمایی معنادار در محیط موتورهای کاوش وب پیشنهاد می‌گردد.

فهرست منابع

- آزاد، اسدالله، و عاطفه شریف. ۱۳۸۶. وب معنایی در پیوند با سایبرنتیک. مجله مطالعات تربیتی و روان‌شناسی ۸ (۳): ۱۳۱-۱۴۷.
- باب‌الحوائجی، فهمیه، سیدمهدی طاهری، و زهرا آقاعابدی. ۱۳۹۴. بررسی تطبیقی کیفیت نمایه‌سازی و رتبه‌بندی پیشینه‌های فراداده‌ای هسته دوبلین و مارک ۲۱ توسط موتورهای کاوش عمومی. فصلنامه مطالعات دانش‌شناسی ۱ (۳): ۴۳-۵۹.
- جمالی مهمویی، حمیدرضا. ۱۳۸۲. وب معنایی: شیوه‌ای رو به تکامل برای ذخیره و بازیابی کارآمدتر اطلاعات روی اینترنت. فصلنامه اطلاع‌شناسی زمستان ۱۳۸۲ (۲): ۴۷-۶۶.

- شریف، عاطفه. ۱۳۸۶. بررسی میزان اثربخشی عناصر ابرداده‌ای بر رتبه‌بندی صفحات وب توسط موتورهای کاوش. فصلنامه کتابداری و اطلاع‌رسانی. ۱۰ (۲): ۲۴۱-۲۵۸.
- طاهری، سیدمهدی. ۱۳۹۲. داده‌های پیوندی: مفاهیم و استانداردها، با تأکید بر جامعه میراث فرهنگی. فصلنامه داخلی کنسرسیوم محتوای ملی ۱۲: ۴-۱۳.
- _____. ۱۳۹۴. فراداده‌ها، موتورهای کاوش و میانکنش‌پذیری آن‌ها. تهران: کتابدار.
- _____. سارا ذوالقدر، و نجلا حریری. ۱۳۹۵. بررسی تطبیقی نمایه‌سازی و پیدانمایی پیشینه‌های فراداده‌ای مبتنی بر الگوی داده‌های خرد توسط موتورهای کاوش وب. (مقاله در دست داوری).
- طاهری، سیدمهدی، ناهید گویلی، مریم شکفته، و مریم کازرانی. ۱۳۹۴. رویکرد بافت-مدار در طراحی پروفایل کاربردی فراداده‌ای و توسعه وب معنایی. فصلنامه مطالعات دانش‌شناسی ۲ (۵): ۴۵-۵۷.
- طاهری، سیدمهدی، نجلا حریری، و سیدرحمت‌الله فتاحی. ۱۳۸۸. بررسی تطبیقی کیفیت نمایه‌سازی و رتبه‌بندی اشیای محتوایی حاوی عناصر فراداده‌ای هسته دوبلین و مارک ۲۱ توسط موتورهای کاوش عمومی. فصلنامه کتابداری و اطلاع‌رسانی ۱۲ (۴): ۱۴۱-۱۶۲.
- طباطبایی امیری، فائزه‌السادات، سیدمهدی طاهری، عبدالحسین فرج‌پهلوی، فریده عصاره، و عبدالحمید معرف‌زاده. ۱۳۹۱. موتورهای کاوش وب و نمایه‌سازی و رتبه‌بندی اشیای محتوایی حاوی عناصر فراداده‌ای دسترس‌پذیر در محیط‌های اطلاعاتی پیوسته پویا. فصلنامه پردازش و مدیریت اطلاعات ۲۷ (۴): ۹۰۷-۹۲۰.
- Berners-Lee, Tim. 2006. Linked Data. Retrieved from <http://www.w3.org/DesignIssues/LinkedData.html>. (accessed Feb. 5, 2016).
- Henshaw, Robin, and Edward J Valauskas. 2001. Metadata as a Catalyst: Experiments with Metadata and Search Engines in the Internet Journal. First Monday. www.librijournal.org/pdf/1999-3pp125-131.pdf (accessed July 14, 2016).
- Isaksen, Leif. 2011. Archaeology and the Semantic Web. http://eprints.soton.ac.uk/196571/1.hasCoversheetVersion/y_gao_PhD_thesis_0111.pdf (accessed July 14, 2016).
- Mixer, Jeffrey. 2013. Linked Data in VRA Core 4.0: Converting VRA XML Records into RDF/XML. <http://jmixter.s3-website-us-east-1.amazonaws.com/thesis/LinkedDataInVRACore4.pdf> (accessed July 14, 2016).
- Mohamed, Khaled A. F. 2006. The impact of metadata in web resources discovering. *Online Information Review* 30 (2): 155-167.
- Quevedo-Torrero, and Jesus Ubaldo. 2004. Improving web retrieval by mining the html tags for keywords and exploring the hyperlink structures web pages. Ph. D. Dissertation. Department of Computer Science, University of Houston. <http://www.lib.umi.com/dissertations/fullcit/3156028>. (accessed July 14, 2016).
- Safari, Mehdi. 2005. Search Engine and Resource Discovery on the Web: Is Dublin Core an Impact Factor. www.webology.ir/2005/v2n2/a13.html. (accessed July 14, 2016).
- Sokvine, Lloyd. 2000. An Evaluation of the Effectiveness of Current Dublin Core Metadata for Retrieval. www.vala.org.au/vala2000/2000pdf/Sokvitne.PDF (accessed July 14, 2016).
- Taheri, S. M., Nadjla Hariri, and S. R. Fattahi. 2014a. Does discarding XML Declarations and Changing File Extensions improve the Indexability and Visibility of Metadata Tag Names in Web Search Engines? *Journal of Information Science* 40 (6): 796-805.

- _____. 2014b. Using Data Island Method for Creating Metadata Records with Indexability and Visibility of Element Tag Names in the Web Search Engines. *Library High Tech*. 32 (1): 83-97.
- Turner, Thomas P, and Lise Brackbill. 1998. Rising to the Top: Evaluating the Use of the HTML META Tag to Improve Retrieval of World Wide Web Documents through Internet Search Engines. <http://cat.inist.fr/?aModele=afficheN&cpsid=1748620> (accessed July 14, 2016).
- Van der Meer, J., F. Boon, F. Hogenboom, F. Frasinca, and U. Kaymak. 2011. A framework for automatic annotation of web pages using the Google rich snippets vocabulary. In Proceedings of the 2011 ACM Symposium on Applied Computing, TaiChung, Taiwan, pages: 765-772. ACM.
- Zhang, Jin, and Alexandra Dimitroff. 2004. Internet search engine's response to metadata Dublin Core implementation. <http://portal.acm.org/citation.cfm?id=1142111> (accessed July 14, 2016).
- _____. 2005a. The impact of metadata implementation on Webpage visibility in search engine result (Part II). http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6VC8-4BHCBX4-2&_user=10&_rdoc=1&_fmt=&_orig=search&_sort=d&view=c&_acct=C000050221&_version=1&_urlVersion=0&_userid=10&md5=a853d410a866732d3f8ab5dd3217d412 (accessed July 14, 2016).
- _____. 2005b. The impact of Webpage content characteristics on webpage visibility in search engine result (Part I). http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6VC8-4BHCBX4-1&_user=10&_rdoc=1&_fmt=&_orig=search&_sort=d&view=c&_acct=C000050221&_version=1&_urlVersion=0&_userid=10&md5=33927751b92200b392f8c79b950dcdb1 (accessed July 14, 2016).

سیدمهدی طاهری

متولد سال ۱۳۵۷، دارای مدرک تحصیلی دکتری در رشته علم اطلاعات و دانش‌شناسی از دانشگاه آزاد واحد علوم و تحقیقات است. ایشان هم‌اکنون استادیار گروه علم اطلاعات و دانش‌شناسی در دانشگاه علامه طباطبائی است.

فرداده، سیستم‌ها و وب معنایی، موتورهای کاوش، و کتابخانه‌های دیجیتال از جمله علایق پژوهشی وی است.



رضا نیکزاد بهله

متولد سال ۱۳۶۹، دارای مدرک تحصیلی کارشناسی ارشد در رشته علم اطلاعات و دانش‌شناسی از دانشگاه علامه طباطبائی است.

فرداده و موتورهای کاوش علایق پژوهشی وی است.



میترا صمیعی

متولد سال ۱۳۵۰، دارای مدرک تحصیلی دکتری در رشته علم اطلاعات و دانش‌شناسی از دانشگاه آزاد واحد علوم و تحقیقات است. ایشان هم‌اکنون استادیار گروه علم اطلاعات و دانش‌شناسی در دانشگاه علامه طباطبائی است. استانداردهای فراداده‌ای، کتابخانه‌های دیجیتالی، حفاظت دیجیتالی، و سازماندهی اطلاعات از جمله علایق پژوهشی وی است.

