

# بررسی تطبیقی کیفیت نمایه‌سازی و رتبه‌بندی اشیای محتوایی حاوی عناصر فراداده‌ای هسته‌دوبلین و مارک ۲۱ توسط موتورهای کاوش عمومی

سید مهدی طاهری<sup>۱</sup>

دکتر نجلا حریری<sup>۲</sup>

دکتر سید رحمت‌الله فتاحی<sup>۳</sup>

## چکیده

پژوهش حاضر با هدف تبیین کیفیت نمایه‌سازی و رتبه‌بندی اشیای محتوایی حاوی عناصر فراداده‌ای هسته‌دوبلین و مارک ۲۱ مبتنی بر زبان نشانه‌گذاری گسترش‌پذیر (XML) توسط موتورهای کاوش عمومی، و مقایسه آنها با یکدیگر به روش تجربی انجام شد. جامعه پژوهش را صد شیء محتوایی مبتنی بر XML در قالب دو گروه تشکیل می‌داد؛ گروه اول حاوی پیشینه‌های فراداده‌ای هسته‌دوبلین، و گروه دوم حاوی پیشینه‌های فراداده‌ای مارک ۲۱ که بر روی دو حوزه فرعی <http://www.demixml.islamicdoc.com> و <http://www.marxml.islamicdoc.com> انتشار یافته بود. حوزه‌های فرعی مورد مطالعه به دو موتور کاوش گوگل و یاهو معرفی شدند. کیفیت نمایه‌سازی پیشینه‌های فراداده‌ای درج شده در اشیای محتوایی و تفاوت نمایه‌سازی و رتبه‌بندی آنها بررسی گردید. تمامی عناصر پیشینه‌های فراداده‌ای مورد مطالعه توسط موتورهای کاوش گوگل و یاهو نمایه‌سازی شدند. بر اساس نتایج آزمون غیر پارامتری یومان - ویتنی، تفاوتی میان کیفیت نمایه‌سازی پیشینه‌های فراداده‌ای هسته‌دوبلین و مارک ۲۱ مبتنی بر XML توسط موتورهای کاوش گوگل و یاهو مشاهده نگردید. میزان معناداری تفاوت رتبه‌بندی اشیای محتوایی حاوی عناصر فراداده‌ای هسته‌دوبلین و مارک ۲۱ در دو موتور کاوش گوگل و یاهو نیز با استفاده

۱. دانشجوی ترم آخر کارشناسی ارشد کتابداری و اطلاع‌رسانی دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران.

taherismster@gmail.com

۲. عضو هیئت علمی دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران.

۳. استاد گروه کتابداری و اطلاع‌رسانی دانشگاه فردوسی مشهد.

از آزمون «یومان - ویتنی» محاسبه گردید. بین رتبه‌بندی اشیای محتوایی حاوی دو طرح فراداده‌ای در موتور کاوش گوگل تفاوت معناداری مشاهده نشد، اما وضعیت رتبه‌بندی اشیای محتوایی حاوی دو طرح فراداده‌ای در موتور کاوش یاهو، تفاوت معناداری داشت. یافته‌های پژوهش نشان داد هیچ یک از دو طرح فراداده‌ای هستهٔ دوبلین و مارک ۲۱ مبتنی بر XML، از لحاظ دسترس‌پذیری توسط موتورهای کاوش عمومی، نسبت به یکدیگر برتری ندارند. کلیدواژه‌ها: طرح فراداده‌ای هسته دوبلین، قالب فراداده‌ای مارک ۲۱، موتورهای کاوش عمومی، کیفیت نمایه‌سازی عناصر فراداده‌ای، رتبه‌بندی نتایج کاوش، زبان نشانه‌گذاری گسترش‌پذیر (XML).

### مقدمه

با گسترش وب، به عنوان مهمترین فناوری و رسانه شبکه اینترنت، که از قابلیت‌های منحصر به فردی بهره‌مند بود، بسیاری از سازمانها، ناشران، مراکز اطلاع‌رسانی و پدیدآورندگان به انتشار و ارائه اطلاعات و خدمات خود بر روی این رسانه تمایل پیدا کردند. این تمایل که با رشد تصاعدی همراه بود، به یکباره حجم وسیعی از اطلاعات را بر روی شبکهٔ وب دسترس‌پذیر نمود. همانند سایر محیطهای اطلاعاتی، طراحی و ایجاد ابزارها، فنون، و استانداردهایی به منظور تسریع و تسهیل دسترسی به این حجم گسترده اشیای محتوایی، ضرورت داشت. در همین راستا، دو جنبش نرم افزاری در سطح بین‌المللی به وجود آمد. در جنبش نخست، برخی اشخاص و سازمانهای انتفاعی و غیرانتفاعی به تولید ابزارهای کاوش مبادرت ورزیدند. محصول این تلاشها، موتورهای کاوش وب بود. از «واندکس»<sup>۱</sup>، «Aliweb» و «جامپ استیشن»<sup>۲</sup> گرفته تا نسل اول، موتورهای کاوش پیمایش‌گر - پایه<sup>۳</sup> و تمام متن مانند وب کراولر<sup>۴</sup> و موتورهای کاوش امروزی همچون گوگل و یاهو - با نرم‌افزارها و الگوریتم‌ها بسیار پیشرفته برای نمایه‌سازی و رتبه‌بندی همگی بر مبنای جنبش نخست به وجود آمدند. تقریباً به موازات این جنبش، جنبش دوم با دستاوردی به نام فراداده<sup>۵</sup> شکل گرفت. اساس این جنبش، همان نظامهای سازماندهی دانش بود که همواره در طول تاریخ همگام با تحولات حوزه فناوریهای اطلاعاتی گسترش یافته‌اند. این جنبش دو رویکرد را شامل می‌شد: سازگاری نظامهای سنتی با محیط و

1. Wandex.
2. Jump station.
3. Crawler-based search engines.
4. WebCrawler.
5. Metadata.

رسانه‌های جدید و ایجاد نظامهایی نوین برای حداکثر بهره‌وری از امکانات و قابلیت‌های محیط جدید (طاهری، ۱۳۸۷).

موتورهای کاوش، با وجود نقاط ضعف و کاستیهایی که نسبت به نظامهای فراداده‌ای در دسترس پذیر نمودن اشیای محتوایی وب دارند، به دلیل سادگی استفاده، غنای پایگاه‌های داده‌ای و عواملی دیگر، ۸۸٪ از زمانی را که کاربران وب صرف جستجوی اطلاعات می‌کنند، به خود اختصاص داده‌اند (بیفت<sup>۱</sup> و کاستیلو<sup>۲</sup>، ۲۰۰۵). گرایش قابل ملاحظه کاربران به استفاده از موتورهای کاوش از یک سو، و قابلیت‌ها و توانایی‌های بالقوه نظامهای فراداده‌ای از سوی دیگر، ضرورت میان کنش پذیری دستاوردهای دو جنبش یاد شده را بدیهی ساخت. بدین منظور، خط مشی‌هایی در موتورهای کاوش به منظور نمایه‌سازی برچسب‌ها و عناصر فراداده‌ای تدوین، و تمهیداتی (از جمله تغییر بستر نحوی پیشینه‌ها) در نظامهای فراداده‌ای برای افزایش بازیافت پذیری<sup>۳</sup> آنها به وسیله موتورهای کاوش<sup>۴</sup> در نظر گرفته شد.

پژوهشهایی چند در ایران (صفری، ۲۰۰۵؛ رجبعلی بگلو، ۱۳۸۶؛ و شریف، ۱۳۸۶) و خارج از ایران (ترنر<sup>۵</sup> و بریکبیل<sup>۶</sup>، ۱۹۹۸؛ سوکویتن<sup>۷</sup>، ۲۰۰۰؛ هنشاو<sup>۸</sup> و والوسکاس<sup>۹</sup>، ۲۰۰۱؛ ژانگ<sup>۱۰</sup> و دیگران<sup>۱۱</sup>، ۲۰۰۴؛ توررو<sup>۱۲</sup>؛ ژانگ<sup>۱۳</sup>، ۲۰۰۵ a، و دیمیتروف<sup>۱۴</sup>، ۲۰۰۵ b) با هدف بررسی این میان کنش پذیری انجام گرفته است. عناصر فراداده‌ای مورد مطالعه در این پژوهشها، فرا برچسبهای «زبان نشانه‌گذاری فرامتن

- .....
1. Bifet.
  2. Castillo.
  3. Retrievability.

۴. این امر اکنون به عنوان یکی از شاخصهای ارزیابی طرح‌های فراداده‌ای محسوب می‌گردد.

5. Turner, Thomas P.
6. Brackbill
7. Sokvitne, Lloyd
8. Henshaw, Robin
9. Valauskas
10. Zhang, Jin
11. Quevedo-Torrero, Jesus Ubaldo
12. Dimitroff

(HTML) «(کلیدواژه، و توصیف) و برخی از عناصر طرح فراداده‌ای هسته دابلین، و بستر نحوی تمامی آنها HTML بوده است.

در این پژوهش، با رویکردی مقایسه‌ای، واکنش موتورهای کاوش به بر چسب «فراداده»<sup>۱</sup> حاوی تمامی عناصر هسته دابلین و مارک ۲۱ مربوط به اشیای محتوایی<sup>۲</sup> جامعه پژوهش در بستر نحوی «زبان نشانه‌گذاری گسترش‌پذیر» (XML)، مورد مطالعه و بررسی قرار گرفت.

### پرسشها و فرضیه‌های پژوهش

به طور کلی، این پژوهش دارای پنج پرسش و چهار فرضیه بود:

#### الف) پرسشها

۱. کیفیت نمایه سازی اشیای محتوایی حاوی عناصر فراداده‌ای هسته دابلین مبتنی بر XML توسط موتورهای کاوش عمومی (گوگل و یاهو) چگونه است؟
۲. کیفیت نمایه سازی اشیای محتوایی حاوی عناصر فراداده‌ای مارک ۲۱ مبتنی بر XML توسط موتورهای کاوش عمومی (گوگل و یاهو) چگونه است؟
۳. واکنش موتورهای کاوش عمومی (گوگل و یاهو) نسبت به اشیای محتوایی مبتنی بر XML حاوی عناصر فراداده‌ای با ساختار مسطح و سلسله مراتبی چگونه است؟ (با توجه به اینکه پیشینه‌های فراداده‌ای مارک دارای ساختار سلسله مراتبی و پیشینه‌های فراداده‌ای هسته دابلین دارای ساختار مسطح هستند).

#### 1. <metadata></metadata>

۲. Digital objects موجودیتی که داده‌ها یا اطلاعات را در بر می‌گیرد. یک شیء محتوایی خود می‌تواند از اشیای محتوایی دیگر ساخته شود. برای مثال، یک مجله یک شیء محتوایی است که خود از مجموعه‌ای مقاله‌ها که هریک از آنها نیز یک شیء محتوایی هستند، ساخته شده است. متن، تصاویر و اشکالی که در یک مقاله هستند نیز می‌توانند یک شیء محتوایی باشند. نقشه‌ها، نقاشیها، عکسها، و دیگر اشیای غیر متنی نیز شیء محتوایی محسوب می‌شوند. حتی فراداده‌های مربوط به یک شیء محتوایی نیز خود اشیای محتوایی می‌باشند (سازمان ملی استانداردهای اطلاعاتی، ۲۰۰۵).

۴. واکنش موتورهای کاوش عمومی (گوگل و یاهو) نسبت به طرحهای فراداده‌ای با برجسبهای عناصر زبان- پایه (هسته دابلین) و بدون برجسبهای عناصر زبان- پایه چگونه است؟

۵. کدام یک از طرحهای فراداده‌ای هسته دابلین و قالب مارک ۲۱ برای سازماندهی اشیای محتوایی شبکه جهانی وب مبتنی بر XML از لحاظ دسترس پذیری (کیفیت نمایه سازی و رتبه بندی) توسط موتورهای کاوش عمومی (گوگل و یاهو) مناسب تر است؟

### ب) فرضیه‌ها

۱. میان کیفیت نمایه سازی سه عنصر اصلی (عنوان، پدیدآورنده، و موضوع) اشیای محتوایی حاوی عناصر فراداده‌ای هسته دابلین و مارک ۲۱ مبتنی بر XML توسط موتور کاوش گوگل، تفاوت معناداری وجود دارد.

۲. میان کیفیت نمایه سازی سه عنصر اصلی (عنوان، پدیدآورنده، و موضوع) اشیای محتوایی حاوی عناصر فراداده‌ای هسته دابلین و مارک ۲۱ مبتنی بر XML توسط موتور کاوش یاهو تفاوت معناداری وجود دارد.

۳. میان رتبه‌بندی اشیای محتوایی حاوی عناصر فراداده‌ای هسته دابلین و مارک ۲۱ مبتنی بر XML توسط موتور کاوش گوگل تفاوت معناداری وجود دارد.

۴. میان رتبه‌بندی اشیای محتوایی حاوی عناصر فراداده‌ای هسته دابلین و مارک ۲۱ مبتنی بر XML توسط موتور کاوش یاهو تفاوت معناداری وجود دارد.

### روش‌شناسی

پژوهش حاضر از نظر نوع کاربردی<sup>۱</sup> است، زیرا بر حل مسئله سازماندهی اشیای محتوایی وب و دسترس‌پذیری آنها از طریق ابزارهای کاوش عمومی وب در شرایط واقعی و عملی تأکید دارد و از لحاظ روش پژوهش، یک پژوهش مبتنی بر فناوری<sup>۲</sup> و تجربی<sup>۳</sup> محسوب می‌گردد. مراحل پژوهش شامل موارد زیر است:

1. Applied or action research.
2. Technology – based Research.
3. Experimental research.

## ۱. طراحی و انتشار حوزه‌های فرعی (subdomains) مورد مطالعه

جامعه پژوهش عبارت بود از ۱۰۰ شیء محتوایی (کتاب الکترونیکی) در قالب دو گروه ۵۰ پیشینه‌ای<sup>۱</sup> در بستر «زبان نشانه‌گذاری گسترش‌پذیری (XML)». هر پیشینه از گروه اول حاوی<sup>۲</sup> یک پیشینه فراداده‌ای هسته دابلین (مبتنی بر XML) است که به عنوان برجسب فراداده<sup>۳</sup> پس از عنصر ریشه<sup>۴</sup> جاسازی<sup>۵</sup> (درج) شده است و بر روی یک حوزه فرعی با آدرس: <http://www.demixml.islamicdoc.com> بر روی وب سایت مرکز اطلاعات و مدارک اسلامی انتشار یافته است. هر یک از اشیای محتوایی گروه دوم حاوی پیشینه‌های فراداده‌ای قالب مارک ۲۱ (مبتنی بر XML) است که همانند پیشینه‌های فراداده‌ای گروه اول عمل شده، و بر روی حوزه فرعی دیگری با آدرس: <http://www.marxml.islamicdoc.com> بر روی همان وب سایت منتشر شده است.<sup>۶</sup>

اشیای محتوایی یاد شده از وب سایت <http://www.archive.org> و از مجموعه منابع کتابخانه دیجیتال کالیفرنیا<sup>۷</sup> و از یک رده موضوعی (Knowledge, theory of)<sup>۸</sup> انتخاب گردید. روش گزینش به این صورت بود که با استفاده از امکان مرور موضوعی<sup>۹</sup>

---

← (۳۴مه از صفحه قبل) شایان ذکر است، پژوهشهایی که به تعیین میزان اثر بخشی عناصر فراداده‌ای بر نمایه‌سازی و رتبه‌بندی اشیای محتوایی وب در موتورهای کاوش پرداخته‌اند، با اندکی تفاوت، از روش تجربی استفاده نموده‌اند (شریف، ۱۳۸۶).

1. Record.
2. Root element.
3. metadata tag.
4. Root element.
5. Embedding.
۶. به منظور توصیف، شناسایی و تسهیل کشف صفحات خانگی هر یک از حوزه‌های فرعی ایجاد شده، با استفاده از ابزارها (Tools) و فرمهای از پیش آماده (Templates) معرفی شده در وب سایت طرح فراداده‌ای هسته دابلین، دو پیشینه فراداده‌ای هسته دابلین در بستر نحوی HTML تولید، و توسط پژوهشگر ویرایش و تکمیل داده‌ها گردید. سپس این پیشینه‌ها در منبع (Source) صفحات خانگی حوزه‌های فرعی جاسازی شدند.
7. California Digital Library.
۸. گزینش پیشینه‌ها از یک رده موضوعی، برای فراهم نمودن شرایطی کاملاً یکسان در فرایند پژوهش انجام شده است.
9. Subject browsing.

در مجموعه منابع کتابخانه دیجیتالی کالیفرنیا، یکی از موضوعاتی که بیش از ۵۰ شیء محتوایی را دارا بود، انتخاب و از طریق امکان «پروتکل انتقال بایگانی (FTP)» که برای هر شیء محتوایی فراهم شده، منتقل<sup>۲</sup> گردید. در مجموعه کتابخانه دیجیتالی کالیفرنیا، همراه هر شیء محتوایی - در همان امکان FTP، پیشینه‌های فراداده‌ای هستهٔ دوبلین و مارک ۲۱ (مبتنی بر XML) - که برای بازیابی اشیای محتوایی مرتبط ایجاد شده‌اند - به صورت مجزا از آن شیء محتوایی وجود دارند. بنابراین، این پیشینه‌ها نیز انتقال یافته و اعتبار آنها مورد سنجش قرار گرفت.<sup>۳</sup> پس از انتقال اشیای مذکور، فرایند جاسازی پیشینه‌های فراداده‌ای در اشیای محتوایی مورد مطالعه بر اساس استانداردها و راهنماهای مرتبط انجام و پس از آماده‌سازی بر روی حوزه‌های فرعی یاد شده منتشر شد.



### نمونه‌ای از یک شیء فراداده‌ای حاوی عناصر فراداده‌ای هسته دوبلین در قالب XML

1. File Transfer Protocol.
2. Download.
۳. اعتبار پیشینه‌ها - از لحاظ ساختار و رعایت قواعد فنی - به وسیلهٔ نرم‌افزار ارائه شده از سوی دفتر توسعه شبکه و استانداردهای مارک صورت پذیرفت.

```

-- <DjVuXML>
-- <BODY>
-- <metadata>
  <leader>00826nam^a2200229la^4500</leader>
  <controlfield tag="001"> GLAD235110-B</controlfield>
  <controlfield tag="005">20060401233913.0</controlfield>
  <controlfield tag="008">850425s1906 enka 000 0 eng d</controlfield>
-- <datafield tag="035" ind1="1" ind2="1">
  <subfield code="a">ocm12210623</subfield>
</datafield>
-- <datafield tag="040" ind1="1" ind2="1">
  <subfield code="a">AAM</subfield>
  <subfield code="c">AAM</subfield>
  <subfield code="d">CUY</subfield>
</datafield>
-- <datafield tag="060" ind1="1" ind2="1">
  <subfield code="a">BD241</subfield>
  <subfield code="b">H7</subfield>
</datafield>
-- <datafield tag="100" ind1="1" ind2="1">
  <subfield code="a">Hooper, Charles E.</subfield>
</datafield>
-- <datafield tag="245" ind1="1" ind2="4">
  <subfield code="a">The anatomy of knowledge:</subfield>
  <subfield code="b">an essay in objective logic /</subfield>
  <subfield code="c">by Charles E. Hooper.</subfield>
</datafield>
-- <datafield tag="260" ind1="1" ind2="1">
  <subfield code="a">London:</subfield>
  <subfield code="b">Watts.</subfield>

```

نمونه‌ای از یک شیء محتوایی حاوی عناصر فراداده‌ای مارک ۲۱ در قالب XML

از آنجا که هدف پژوهش، بررسی کیفیت نمایه‌سازی و رتبه‌بندی اشیای محتوایی حاوی پیشینه‌های فراداده‌ای به وسیلهٔ موتورهای کاوش عمومی بود، بدیهی است باید بهترین موتورهای کاوش عمومی که بتوانند هدفهای این پژوهش را محقق سازند، انتخاب می‌گردید. بسیاری از منابع معتبر اذعان دارند و نیز بر اساس مشاهدات، تجارب مدیران وب‌سایتها، تعداد کاربران، بررسی ساختار و ویژگیها، و سایر شواهد و دلایل، موتورهای کاوش گوگل<sup>۱</sup> و یاهو<sup>۲</sup> در بین موتورهای کاوش عمومی دارای بالاترین کیفیت و بالاترین بسامد استفاده هستند (شریف، ۱۳۸۶؛ ویکی پدیا، ۲۰۰۸k؛ الکسا<sup>۳</sup>، ۲۰۰۸؛ لوییس<sup>۴</sup>، ۲۰۰۸؛

- .....
1. Google search engine: <http://www.google.com>
  2. Yahoo search engine and subject directory: <http://www.yahoo.com>
  3. Alexa
  4. Lewis



کمپکس<sup>۱</sup>، (۲۰۰۸).

## ۲. انتخاب و معرفی حوزه‌های فرعی به موتورهای کاوش

گام بعدی بر اساس مطالعات و بررسی‌های انجام شده، معرفی حوزه‌های فرعی جامعه پژوهش به موتورهای گوگل و یاهو، با استفاده از پیشرفته‌ترین روشهای توصیه شده از سوی آنها بود. بدین منظور، حوزه‌های فرعی مورد مطالعه به وسیله امکان Webmaster Tools<sup>۲</sup> و XML Sitemap<sup>۳</sup> به موتور کاوش گوگل؛ و با بهره‌مندی از امکان Suggest a site<sup>۴</sup>، Yahoo! Search URL Status Review Form<sup>۵</sup>، و ROR & Text Sitemap<sup>۶</sup> به موتور کاوش یاهو<sup>۷</sup> در شرایط کاملاً یکسان معرفی شدند. شایان ذکر است، اعتبار همه حوزه‌های فرعی جامعه پژوهش - با تعبیه کدهای مرتبط و ارائه شده از سوی گوگل و یاهو- مورد تأیید موتورهای مورد مطالعه قرار گرفت.<sup>۸</sup>

.....  
1. Campex.

2. <https://www.google.com/accounts/ServiceLogin?service=sitemaps&passive=true&nui=1&continue=http%3A%2F%2Fwww.google.com%2Fwebmasters%2Ftools%2Fsiteoverview&followup=http%3A%2F%2Fwww.google.com%2Fwebmasters%2Ftools%2Fsiteoverview&hl=en>

۳. این بایگانی (فایل) مبتنی بر XML از طریق تولید کننده پیوسته گوگل ( Google Online Sitemap Generator) با آدرس: <http://www.xml-sitemap.com> تولید شد.

4. <http://search.yahoo.com/info/submit.html>

5. <http://help.yahoo.com/l/us/yahoo/search/urlstatus.html>

۶. این بایگانی‌ها (فایل‌ها) نیز به وسیله تولید کننده پیوسته گوگل تهیه گردید.

۷. برای معرفی وبسایت‌های جامعه پژوهش به موتور کاوش یاهو، به دلیل کندی و ضعف روباتها و خزنده‌های (crawler) این موتور نسبت به گوگل، از روشهای متنوع و مورد تأیید این موتور استفاده شد. این مشکل موجب صرف وقت و زحمت فراوان در حین فرایند پژوهش گردید.

8. Google:

<https://www.google.com/accounts/ServiceLogin?service=sitemaps&passive=true&nui=1&continue=http%3A%2F%2Fwww.google.com%2Fwebmasters%2Ftools%2Fsiteoverview&followup=http%3A%2F%2Fwww.google.com%2Fwebmasters%2Ftools%2Fsiteoverview&hl=en>

Yahoo:

[https://login.yahoo.com/config/login\\_verify2?.src=siteexplorer&.done=http%3A%2F%2Fsiteexplorer.search.yahoo.com%2Fmysites](https://login.yahoo.com/config/login_verify2?.src=siteexplorer&.done=http%3A%2F%2Fsiteexplorer.search.yahoo.com%2Fmysites)

### ۳. گردآوری داده‌ها

روش گردآوری داده‌ها در این پژوهش از طریق مشاهده طراحی شده<sup>۱</sup> انجام گرفته است - تماشای دقیق به شیوه‌ای علمی و نظام‌مند با استفاده از راهبردهای از پیش تعیین شده<sup>۲</sup> که اغلب در مطالعه‌های کمی برای توصیف‌های نظام‌مند یا آزمون فرضیه‌های علمی، به کار می‌رود (پاول، ۱۳۷۹، ص. ۱۷۸). ابزار گردآوری داده‌ها، سیاهه<sup>۳</sup> و ارسسی (چک لیست) بود که با استفاده از نرم افزار آماری SPSS - به عنوان ابزار ثبت مکانیکی، بر اساس سؤالهای پژوهش توسط پژوهشگر طراحی شده است. سپس داده‌های گردآوری شده در قالب جدولها و نمودارها که آنها نیز با نرم افزار SPSS تهیه گردید، ارائه شد. ابزار دیگری که برای مشاهده و ثبت داده‌ها استفاده شده، تهیه<sup>۴</sup> تصاویری نمونه در فاصله‌های زمانی نامنظم با استفاده از امکان Screen shot و ذخیره آنها از طریق نرم افزار Paint بوده است. بدیهی است، این نوع ابزارها توان ثبت نظام‌مند داده‌ها و مقوله‌بندی آنها را ندارند و این کار باید توسط پژوهشگر انجام گیرد (پاول، ۱۳۷۹، ص. ۱۷۸).

به منظور گردآوری داده‌های مربوط به سؤالهای اول، دوم، سوم، و چهارم، و فرضیه‌های اول و دوم پژوهش از راهبردهای جستجوی:

"keyphrase" site:dcmixml.islamicdoc.com

"keyphrase" site:marcxml.islamicdoc.com

که در کادر جستجوی پایه<sup>۳</sup> موتورهای گوگل و یاهو وارد شدند، استفاده شد. و

برای فرضیه‌های سوم و چهارم این راهبردها طراحی گردید:

"keyphrase" site:dcmixml.islamicdoc.com OR  
site:marcxml.islamicdoc.com

"keyphrase" site:marcxml.islamicdoc.com OR  
site:dcmixml.islamicdoc.com

.....  
1. Structured observation.

۲. در این پژوهش، منظور راهبردهای جستجو (Search strategies) تعیین شده بر اساس قوانین و توصیه‌های ارائه شده از سوی موتورهای کاوش گوگل و یاهو است.

3. Basic search.

چنان‌که پیشتر اشاره شد، هدف از طرح فرضیه‌های سوم و چهارم، بررسی رتبه‌بندی هر یک از اشیای محتوایی گروه اول و دوم نسبت به یکدیگر بود. بنابراین، بررسی رتبه‌بندی آنها نسبت به سایر پیشینه‌های نتایج جستجو در موتورهای کاوش، مد نظر نبود. عملیات کاوش با استفاده از راهبردهای جستجوی یاد شده، تنها دو پیشینه را به عنوان نتیجه ارائه می‌داد؛ یک پیشینه مربوط به گروه اول و یک پیشینه مربوط به گروه دوم که رتبه آنها نسبت به یکدیگر سنجیده می‌شد.

به طور کلی، در مرحله گردآوری داده‌ها، بیش از دو هزار عملیات جستجو با استفاده از راهبردهای جستجوی مرتبط در بازه زمانی اول تا سی و یکم اردیبهشت ماه ۱۳۸۷ صورت گرفت.

### تجزیه و تحلیل یافته‌ها

۱. نتایج آزمون یومان - ویتنی<sup>۱</sup> مربوط به فرضیه اول.

	F
Mann-Whitney U	1250.000
Wilcoxon W	2575.000
Z	.000
Asymp. Sig. (2-tailed)	1.000

a Grouping Variable: Metadata

چنان‌که نتیجه آزمون نشان می‌دهد، میان کیفیت نمایه‌سازی سه عنصر اصلی (پدیدآورنده، عنوان، و موضوع) اشیای محتوایی حاوی پیشینه‌های فراداده‌ای هسته‌دوبلین و مارک ۲۱ توسط موتور کاوش گوگل، تفاوت معناداری وجود ندارد.

۲. نتایج آزمون یومان - ویتنی مربوط به فرضیه دوم.

	F
Mann-Whitney U	11250.000
Wilcoxon W	22575.000
Z	.000
Asymp. Sig. (2-tailed)	1.000

a Grouping Variable: Metadata

۱. در آزمونهای مربوط به چهار فرضیه پژوهش،  $\alpha=5\%$  در نظر گرفته شد.

میان کیفیت نمایه سازی سه عنصر اصلی اشیای محتوایی حاوی پیشینه‌های فراداده‌ای هسته دوبلین و مارک ۲۱ توسط موتور کاوش یا هو نیز تفاوت معناداری مشاهده نشد.

۳. نتایج آزمون یومان- ویتنی مربوط به فرضیه سوم.

	F
Mann-Whitney U	1250.000
Wilcoxon W	2525.000
Z	.000
Asymp. Sig. (2-tailed)	1.000

a. Grouping Variable: Metadata

b. SE = Google

در موتور کاوش گوگل، ۲۵ پیشینه فراداده‌ای هسته دوبلین و ۲۵ پیشینه فراداده‌ای مارک ۲۱ درج شده در اشیای محتوایی، نسبت به یکدیگر بالاتر قرار گرفتند. بنابراین، تفاوت معناداری میان رتبه‌بندی اشیای محتوایی عناصر فراداده‌ای مربوط به هسته دوبلین و مارک ۲۱ در گوگل وجود ندارد.

۴. نتایج آزمون یومان- ویتنی مربوط به فرضیه چهارم.

	F
Mann-Whitney U	1250.000
Wilcoxon W	2525.000
Z	.000
Asymp. Sig. (2-tailed)	1.000

a. Grouping Variable: Metadata

b. SE = Google

اما رتبه اشیای محتوایی در موتور کاوش یا هو متفاوت بود. ۴۰ شیء محتوایی حاوی عناصر فراداده‌ای هسته دوبلین و ۱۰ شیء محتوایی حاوی عناصر فراداده‌ای مارک ۲۱ در این موتور کاوش نسبت به یکدیگر رتبه بالاتری کسب نمودند. پاسخ مربوط به پرسشهای اول و دوم. کیفیت نمایه‌سازی اشیای محتوای حاوی

عناصر فراداده‌ای هسته دوبلین و مارک ۲۱ توسط موتورهای کاوش گوگل و یاهو

جدول ۱. وضعیت نمایه‌سازی عناصر فراداده‌ای مارک ۲۱

و هسته دوبلین توسط موتور کاوش گوگل

درصد نمایه‌سازی	تعداد پیشینه‌ها	تعداد عناصر مورد مطالعه	وب سایت	طرح فراداده‌ای (در گوگل)
٪۱۰۰	۵۰	۱۰	marxml.islamicdoc.com	مارک
٪۱۰۰	۵۰	۹	dcmixml.islamicdoc.com	هسته دوبلین

موتور کاوش گوگل محتوای تمامی ۱۰ عنصر قالب مارک و ۹ عنصر طرح هسته

دوبلین را به طور کامل (صد در صد) نمایه‌سازی کرده است.

جدول ۲. وضعیت نمایه‌سازی عناصر فراداده‌ای مارک ۲۱

و هسته دوبلین توسط موتور کاوش گوگل

درصد نمایه‌سازی	تعداد پیشینه‌ها	تعداد عناصر مورد مطالعه	وب سایت	طرح فراداده‌ای (ياهو)
٪۱۰۰	۵۰	۱۰	marxml.islamicdoc.com	مارک
٪۱۰۰	۵۰	۹	dcmixml.islamicdoc.com	هسته دوبلین

کیفیت نمایه‌سازی عناصر مربوط به طرحهای فراداده‌ای هسته دوبلین و مارک

۲۱ توسط موتور کاوش یاهو ٪۱۰۰ است.

بر اساس یافته‌های مربوط به پرسشهای اول و دوم، می‌توان به پرسشهای سوم و چهارم نیز پاسخ داد. برای موتورهای کاوش گوگل و یاهو، بین اشیای محتوایی با ساختار مسطح و سلسله مراتبی تفاوتی وجود ندارد. موتورهای یاد شده، الگوریتمی برای نمایه‌سازی پیشینه‌های فراداده‌ای با برچسب‌های عناصر زبان-پایه و بدون آن نیز ندارد.

پاسخ پرسش پنجم در نتایج چهار فرضیه و چهار پرسش پیشین نهفته است. یافته‌ها حاکی از آن است که هیچ یک از دو طرح فراداده‌ای هسته دوبلین و مارک از لحاظ

دسترس پذیری توسط موتورهای کاوش عمومی، بر دیگری برتری ندارند، و هر دو طرح برای سازماندهی اشیای محتوایی وب مناسب می‌باشند.

به طور اجمالی، مطالعه بر روی میان‌کنش‌پذیری اشیای محتوایی حاوی عناصر هسته‌دوبلین و مارک ۲۱ مبتنی بر XML با موتورهای کاوش گوگل و یاهو بیانگر آن است که:

- تمامی عناصر فراداده‌ای دو طرح هسته‌دوبلین و مارک ۲۱ توسط موتورهای کاوش نمایه‌سازی شده‌اند. بنابراین، اشیای محتوایی حاوی این عناصر، از طریق جستجوی کلیدواژه‌ای و عبارتهای کلیدی، به آسانی قابل‌بازیابی هستند.

- موتورهای کاوش الگوریتمی برای انتخاب عناصری خاص نداشته‌اند. به عبارت دیگر، تمامی عناصر فراداده‌ای برای نرم‌افزارهای نمایه‌سازی موتورهای کاوش گوگل و یاهو، ارزش نمایه‌شدن داشته‌اند.

به طور ضمنی، این یافته می‌تواند نویدی برای فراهم‌کنندگان اطلاعات دسترس‌پذیر از طریق شبکه جهانی وب باشد، زیرا پیاده‌سازی پیشینه‌های فراداده‌ای در بستر «زبان نشانه‌گذاری فرامتن (HTML)» چنین قابلیت را دارا نبود.

- بستر XML برای طرح‌های فراداده‌ای، رهایی از محدودیتهای بستر HTML در میان‌کنش‌پذیری موتورهای کاوش - به عنوان مهمترین ابزار جستجو در اینترنت - با پیشینه‌های فراداده‌ای - به عنوان ابزاری برای توصیف، شناسایی، و کشف اشیای محتوایی - است.

- پژوهشهای پیشین تنها بخشی از عناصر (۴ یا ۵ عنصر اصلی) را که تقریباً با فرا برچسبهای HTML مطابقت داشته، انتخاب نموده‌اند. اما پژوهش حاضر تمامی عناصر موجود در پیشینه‌های فراداده‌ای مربوط به اشیای محتوایی جامعه پژوهش را مورد بررسی قرار داد، و بازیافت پذیری همه عناصر را تأیید نمود.

- قالب مارک پیش از ایجاد قالب مبتنی بر XML خود (MARXML)، قالب پایگاه اطلاعاتی - پایه<sup>۱</sup> بود، و امکان پیاده‌سازی آن در صفحات وب وجود نداشت. مارک در قالب «زبان نشانه‌گذاری گسترش‌پذیر» (HTML)، به دلیل محدودیتهای این زبان (محدودیت برجسبهای آن) پیاده‌سازی نشد. بنابراین، نمی‌توانست در صفحات وب جاسازی شود و در نتیجه مانع حضور فعال آن در سازماندهی اشیای محتوایی نظامهای ایستای وب گردید. پیاده‌سازی مارک در بستر XML باعث حذف این محدودیت گردید، و از همه مهمتر بازیافت پذیری آنها از طریق ابزارهای عمومی کاوش را میسر ساخت.

- تمامی اشیای محتوایی حاوی عناصر فراداده‌ای هسته دابلین و مارک، از طریق سه عنصر اصلی مذکور (عنوان، موضوع، پدیدآورنده) دسترس‌پذیر می‌باشند. بنابراین، هیچ تفاوتی بین نمایه‌سازی پیشینه‌های فراداده‌ای هسته دابلین و مارک درج شده در اشیای محتوایی مبتنی بر XML توسط هر دو موتور کاوش گوگل و یاهو مشاهده نمی‌شود. در نتیجه، نمی‌توان گفت کدام طرح فراداده‌ای توسط موتورهای کاوش بهتر نمایه‌سازی می‌شود.

- بررسی فرضیه‌های سوم و چهارم با استفاده از آزمون یومان-ویتنی، وجود تفاوت بین رتبه‌بندی پیشینه‌های دو طرح فراداده‌ای مورد مطالعه در موتور کاوش یاهو را نشان داد.

- نرم افزارهای پیمایشگر - نمایه‌ساز موتورهای کاوش گوگل و یاهو حساسیت خاصی (الگوریتم) نسبت به مدارک مبتنی بر XML با ساختار مسطح (مانند هسته دابلین) و سلسله‌مراتبی ندارند (مانند مارک ۲۱)، و کنش مشابهی را نشان می‌دهند. نیز بین پیشینه‌های با برجسبهای زبان - پایه و دیگر پیشینه‌ها، تفاوتی قایل نمی‌شوند؛ یعنی خط مشی‌هایی که نسبت به نمایه‌سازی برجسبهای HTML داشتند، حداقل در مورد برجسبهای فراداده‌های درج شده در مدارک مبتنی بر XML ندارند. به عبارت دیگر، کنش

.....  
1. Database-based format.

موتورهای کاوش گوگل و یاهو نسبت به طرح‌های فراداده‌ای با برجسبهای عناصر زبان - پایه (هسته دابلین) و بدون برجسبهای عناصر زبان - پایه (مارک) یکسان است.

- به منظور تعیین برتری دو طرح فراداده‌ای هسته دابلین و مارک ۲۱ از لحاظ دسترس پذیری (کیفیت نمایه سازی و رتبه بندی) چهار عامل<sup>۱</sup> در نظر گرفته شد. از میان این چهار عامل، یافته‌های سه عامل حاکی از برتری نداشتن هر یک از دو طرح نسبت به یکدیگر بود. اما عامل چهارم (رتبه بندی اشیای محتوایی مورد مطالعه در موتور کاوش یاهو) برتری اشیای محتوای حاوی عناصر هسته دابلین را نسبت به گروه دوم نشان داد. با توجه به اینکه سه عامل اول مناسب بودن هر دو طرح را از لحاظ دسترس پذیری نشان می‌داد، و از سوی دیگر تفاوت در رتبه بندی در مهمترین و قویترین موتور کاوش حال حاضر وب (گوگل) - که الگوی سایر موتورهای کاوش مهم در طراحی الگوریتمهای نمایه سازی و رتبه بندی نیز هست - دیده نشد، و از همه مهمتر آنکه رتبه اشیای محتوایی مربوط به هر دو گروه جامعه پژوهش در یاهو با تغییر عناصر مورد جستجو تغییر می‌نمود. هیچ یک از طرح‌های فراداده‌ای هسته دابلین و قالب فراداده‌ای مارک ۲۱ از لحاظ دسترس پذیری (کیفیت نمایه سازی و رتبه بندی) به وسیله موتورهای کاوش عمومی بر دیگری برتری ندارند.

### نتیجه گیری

همان طور که در مقدمه اشاره شد، برای تسهیل و تسریع دسترسی به حجم گسترده اشیای محتوایی انتشار یافته بر روی شبکه وب، دو جنبش نرم افزاری با دو رویکرد متفاوت رخ داد. جنبش نخست، جنبش کاملاً فناوری-پایه بود. ضعفهای دستاوردهای آن را می‌توان ناشی از نوعی فن زدگی دانست؛ اما جنبش دیگر بر پایه مدیریت اطلاعات شکل گرفت و استفاده از فناوری را به منظور سازگاری با رسانه‌ها و محیط جدید مد نظر قرار داد (فتاحی، ۱۳۸۶). محبوبیت ابزارهای مبتنی بر جنبش نخست و قابلیت‌های محصول جنبش



دوم، بیانگر اهمیت میان‌کنش‌پذیری آنهاست. بررسی تلاش‌های صورت گرفته از سوی دو جنبش در جهت میان‌کنش‌پذیری، نشان می‌دهد جنبش دوم گام‌های جدی‌تر و برجسته‌تری برداشته است.

گرایش دو طرح فراداده‌ای یاد شده به سوی استفاده از فناوری پیشرفته «زبان نشانه‌گذاری گسترش‌پذیر» (XML) به عنوان بستر نحوی (یا محیط پیاده‌سازی)، با هدف استفاده از قابلیت‌های این فناوری، در راستای حرکت آگاهانه به سوی سازگاری و پاسخگویی به تحولات اخیر است. این امر ارزشهای افزوده فراوانی برای این دو طرح داشته است:

- امکان جاسازی پیشینه‌های فراداده‌ای هسته دوبلین و مارک ۲۱ در اشیای دیجیتال در قالب‌های مختلف، از جمله قالب PDF با استفاده از RDF
- امکان نمایه‌سازی کامل آنها توسط موتورهای کاوش عمومی (ابزاری که بیش از ۸۰٪ جستجوهای محیط وب را به خود اختصاص داده‌اند).
- تسهیل تبدیل طرح‌های فراداده‌ای به یکدیگر، و در نتیجه افزایش میان‌کنش‌پذیری نظام‌های فراداده‌ای
- امکان ایجاد پروفایل‌های کاربردی<sup>۱</sup> به منظور تأمین نیازهای خاص سازمانها و محیط‌های اطلاعاتی مختلف (مانند قالب METS)
- بستری بسیار منعطف و با قابلیت‌های منحصر به فرد برای انجام پردازش‌های سطح بالا بر روی پیشینه‌های فراداده‌ای.

.....  
1. Application profiles.

مجموعه عناصر فراداده‌ای (از یک یا چند استاندارد فراداده‌ای)، خط مشی‌ها، و رهنمودهایی که به منظور کاربردهای خاص (محلی) تعریف شده است (فتاحی و دیگران، ۱۳۸۷).

### منابع

- پاول، رونالد آر. (۱۳۸۵). روش‌های اساسی پژوهش برای کتابداران. ترجمه نجلا حریری. تهران: دانشگاه آزاد اسلامی
- شریف، عاطفه (۱۳۸۶). «بررسی میزان اثر بخشی عناصر ابر داده‌ای بر رتبه بندی صفحات وب توسط موتورهای کاوش». فصلنامه کتابداری و اطلاع‌رسانی. شماره ۳۸ (تابستان ۱۳۸۶) [پیوسته]، دسترس پذیر:  
[www.aftab.ir/articles/publications/information/c17c1205315159\\_ketabdari\\_va\\_etelarasani38\\_p1.php](http://www.aftab.ir/articles/publications/information/c17c1205315159_ketabdari_va_etelarasani38_p1.php). [۳۰ آبان ۱۳۸۶]
- رجبعلی بگلو، رضا (۱۳۸۵). «بررسی رابطه وب سنجی و میزان استفاده از عناصر فراداده‌ای دابلین کور در مجلات الکترونیکی دسترسی آزاد در رشته علوم کتابداری و اطلاع‌رسانی». پایان نامه کارشناسی ارشد کتابداری و اطلاع‌رسانی، دانشکده علوم تربیتی و روانشناسی، دانشگاه شیراز. [پیوسته]، دسترس پذیر:  
<http://alpha-5.irandoc.ac.ir/scripts/wxis.exe?a=5:49:42>. [۳۰ آبان ۱۳۸۶]
- طاهری، مهدی (۱۳۸۷). «مقایسه کارایی طرح فراداده ای هسته دوبلین و قالب فراداده مارک ۲۱ در سازماندهی منابع اطلاعاتی شبکه جهانی وب». فصلنامه کتابداری و اطلاع‌رسانی، شماره ۴۳ (پاییز ۱۳۸۷) [پیوسته]، دسترس پذیر:  
<http://aqlibrary.ir/Issue/ShowOneArticle.aspx?ArticleCode=622>  
[۱۴ تیر ۱۳۸۷]
- فتاحی، رحمت‌الله (۱۳۸۶). «از آرمانها تا واقعیت: تحلیلی از مهمترین چالشها و رویکردهای سازماندهی اطلاعات در عصر حاضر». فصلنامه کتابداری و اطلاع‌رسانی، شماره ۴۰ (زمستان ۱۳۸۶) [پیوسته]، دسترس پذیر:  
<http://aqlibrary.ir/Issue/ShowOneArticle.aspx?ArticleCode=469>  
[۱۴ تیر ۱۳۸۷]

- فتاحی، رحمت‌الله، مهدی طاهری و فرشته ناقد احمدی (۱۳۸۷). فهرست‌نویسی رایانه‌ای: مفاهیم، شیوه‌ها، و ابزارهای فهرست‌نویسی در محیط رایانه‌ای. تهران: کتابدار.

- Alexa: The Web Information Company (2008). "Global Top 500". [online], available at: [http://www.alexa.com/site/ds/top\\_sites?ts\\_mode=global&lang=none](http://www.alexa.com/site/ds/top_sites?ts_mode=global&lang=none) . [14 Dec. 2007].

- Bifet, Albert; Castillo, Carlos (2005). "An Analysis of Factors Used in Search Engine Ranking". [online], available at: <http://airweb.cse.lehigh.edu/2005/bifet.pdf>. [14 Dec. 2007].

- Campex (2005). "Top Search Engines". [online], available at: <http://capmex.biz/resources/top-search-engines> .[14 Dec. 2007].

- Google (2006). "Add your URL to Google". [online], available at: <http://www.google.com/addurl/>. [14 Dec. 2007].

- Google (2008a). "Create your Google Sitemap Online". [online], available at: <http://www.xml-sitemaps.com/>. [25 May 2008].

- Google (2008a). "Dashboard". [online], available at: <https://www.google.com/accounts/ServiceLogin?service=sitemaps&passive=true&nui=1&continue=http%3A%2F%2Fwww.google.com%2Fwebmasters%2Ftools%2Fsiteoverview&followup=http%3A%2F%2Fwww.google.com%2Fwebmasters%2Ftools%2Fsiteoverview&hl=en> . [25 May 2008].

- Google (2008b). "Google Webmaster Tools" [online], available at: <https://google.com/webmasters/tools/docs/en/about.html>. [25 May 2008].

- Henshaw, Robin; Valauskas, Edward J (2001). " Metadata as a Catalyst: Experiments with Metadata and Search Engines in theInternet Journal, First Monday ". [online], available at: [www.librijournal.org/pdf/1999-3pp125-131.pdf](http://www.librijournal.org/pdf/1999-3pp125-131.pdf) . [5 Dec. 2007].

- Kelly (2006). "MARC and MARC XML". [online], available at: <http://threegee.files.wordpress.com/2006/05/marcxml.pdf>. [14 Dec. 2007].

- Lewis, Edward (2008). "Top Ten Search Engines". [online], available at: <http://www.seoconsultants.com/search-engines/>. [25 May 2008].

- Luk, Robert, et al (2000). "A Survey of Search Engines for XML Documents". [online], available at: <http://www.haifa.il.ibm.com/sigir00-xml/final-papers/Luk/XMLSUR.htm>. [14 Dec. 2007].

- The National Information Standards Organization (2005). "Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies". [online], available at: <http://www.niso.org/standards/resources/Z39-19-2005.pdf>. [5 Dec. 2006].

- Quevedo-Torrero, 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. [online], available at: <http://www.lib.umi.com/dissertations/fullcit/3156028>. [5 Dec. 2007].

- Safari, Mehdi (2005). "Search Engine and Resource Discovery on the Web: Is Dublin Core an Impact Factor". [online], available at: [www.webology.ir/2005/v2n2/a13.html](http://www.webology.ir/2005/v2n2/a13.html). [5 Dec. 2007].

- Sokvine, Lloyd (2000). "An Evaluation of the Effectiveness of Current Dublin Core Metadata for Retrieval". [online], available at: [www.vala.org.au/vala2000/2000pdf/Sokvitne.PDF](http://www.vala.org.au/vala2000/2000pdf/Sokvitne.PDF). [5 Dec. 2007].

- Turner, Thomas P.; Brackbill, Lise (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". [online], available at: <http://cat.inist.fr/?aModele=afficheN&>

cpsidt=1748620 . [5 Dec. 2007].

- Wikipedia (2008k). "Web search engine". [online], available at: [http://en.wikipedia.org/wiki/Search\\_engines](http://en.wikipedia.org/wiki/Search_engines). [25 May 2008].

- Yahoo! (2008a). "Yahoo! Site Explorer". [online], available at: [http://en.wikipedia.org/wiki/Yahoo!\\_Site\\_Explorer](http://en.wikipedia.org/wiki/Yahoo!_Site_Explorer). [25 May 2008].

- Yahoo! (2008b). "Yahoo Search Content Quality Guidelines". [online], available at: <http://help.yahoo.com/l/us/yahoo/search/basics/basics-18.html> . [25 May 2008].

- Yahoo! (2008c). "Yahoo! Submit Your Site". [online], available at: <http://search.yahoo.com/info/submit.html> . [25 May 2008].

- Zhang, Jin; Dimitroff, Alexandra (2004). "Internet search engine's response to metadata Dublin Core implementation ". [online], available at: <http://portal.acm.org/citation.cfm?id=1142111> . [5 Dec. 2007].

- Zhang, Jin; Dimitroff, Alexandra (2005a). "The impact of metadata implementation on Webpage visibility in search engine result (Part II)". [online], available at: [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](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) . [5 Dec. 2007].

- Zhang, Jin; Dimitroff, Alexandra (2005b). "The impact of Webpage content characteristics on webpage visibility in search engine result (Part I)". [online], available at: [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=33927751b92200b392f8c79b950dcd1](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=33927751b92200b392f8c79b950dcd1) . [5 Dec. 2007].

<http://www.archive.org>  
<http://www.dcmixml.islamicdoc.com>  
<http://www.google.com>  
<http://www.marxml.islamicdoc.com>  
<http://www.yahoo.com>