

بررسی پایگاه‌های کتاب الکترونیکی با تأکید بر ابر داده

الهه امتی^۱ مینا توسلی فرحی^{۲*}

۱. کارشناس ارشد، کتابداری و اطلاع‌رسانی پزشکی، مرکز توسعه تحقیقات بالینی مرکز آموزشی درمانی و پژوهشی هفده شهریور، دانشگاه علوم پزشکی گیلان، رشت، ایران.

۲. استادیار، کتابداری و اطلاع‌رسانی پزشکی، مرکز تحقیقات مدیریت اطلاعات سلامت، دانشگاه علوم پزشکی هرمزگان، بندرعباس، ایران.

مجله اطلاع‌رسانی پزشکی نوین؛ دوره اول؛ شماره اول؛ بهار و تابستان ۱۳۹۴؛ صفحات ۴۱-۳۳

چکیده

مقدمه: با رشد تصاعدی منابع الکترونیکی در محیط وب، کاربرد ابر داده‌ها، باعث افزایش دقت بازیابی و تسهیل جستجو در منابع الکترونیکی می‌شوند. لذا هدف از این مطالعه، تعیین ابر داده در پایگاه‌های کتاب الکترونیکی بوده است.

روش‌ها: این مطالعه از نوع کاربردی و به روش پیمایشی در سال ۱۳۹۳ انجام شده است. جامعه آماری شامل ۹ پایگاه اطلاعاتی کتاب الکترونیکی؛ *Connect Palgrave*، *MyiLibrary*، *Credo Reference*، *EBL-Ebook Library*، *Gale Virtual Reference Library*، *Taylor & Francis eBookstore*، *Dawsonera*، *Ebrary*، *NetLibrary* بود؛ که به روش هدف‌دار انتخاب شده‌اند. اطلاعات مربوط به ابر داده‌های مورد استفاده در وبسایت «بازار ارزیابی پایگاه‌های علمی» (<http://adat.crl.edu>) که به مقایسه پایگاه‌ها از جنبه‌های مختلف می‌پردازد، قابل نمایش بود. در این وبسایت با انتخاب پایگاه‌ها می‌توان نتایج را مشاهده کرد و کاربرد انواع ابر داده‌ها در آن‌ها را باهم مقایسه کرد. داده‌های به دست آمده بعد از ورود به نرم‌افزار آماری *Excel* مورد تحلیل قرار گرفته‌اند.

یافته‌ها: یافته‌ها نشان داد در پایگاه‌های مورد نظر، ابر داده‌های استفاده شده، به ترتیب *MARC* به صورت (۱۰۰ درصد)، *Onix* (۵۵/۵۵ درصد)، *Dublin Core* (۱۱/۱۱ درصد) به کار گرفته شده بود و *UKLom* در هیچ کدام از پایگاه‌ها استفاده نشده بود.

بحث و نتیجه‌گیری: بر اساس نتایج اخذ شده، پیشنهاد می‌گردد متخصصان اطلاع‌رسانی و طراحان پایگاه‌ها در ایجاد پایگاه‌های کتاب الکترونیکی، ابر داده‌های مناسب را در نظر بگیرند.

کلیدواژه‌ها: ابر داده، *MARC*، *UKLom*، *Dublin Core*، *Onix*

نوع مقاله: پژوهشی

پذیرش مقاله: ۹۳/۱۰/۲۱

اصلاح نهایی: ۹۳/۰۵/۰۵

دریافت مقاله: ۹۳/۰۷/۰۹

ارجاع: امتی الهه، توسلی فرحی مینا. بررسی پایگاه‌های کتاب الکترونیکی با تأکید بر ابر داده. اطلاع‌رسانی پزشکی نوین. ۱۳۹۴؛ (۱۱): ۴۱-۳۳.

مقدمه:

سازی این منابع قرار داده است. اقدامات متعددی در این زمینه صورت گرفته است که گرایش به استفاده از ابزار نوینی تحت عنوان ابر داده یکی از این موارد است. ابر داده در تعریفی ساده به داده‌ها و اطلاعاتی درباره منابع اطلاعاتی اطلاق می‌شود. در واقع ابر داده، ابزاری است که برای شناسایی، توصیف و سازمان‌دهی منابع اطلاعاتی الکترونیکی موجود در اینترنت به کار می‌رود [۱].

این اصطلاح اولین بار توسط شرکت تجاری آمریکایی به وسیله Jack Myers Voss در سال ۱۹۶۹ برای توصیف متا الگوها به کار

تکامل عصر الکترونیک منجر به توسعه پایگاه‌های الکترونیکی کتاب در موضوعات مختلف شده است و این پایگاه‌ها نقش بسزایی در دسترسی آسان و سریع به کتاب‌های الکترونیک مختلف را بدون مراجعه به کتابخانه ایفا می‌کنند، در واقع شبکه جهانی وب یکی از روش‌های رایج برای به دست آوردن اطلاعات است، اما گسترش فزاینده منابع اطلاعاتی در قالب‌های الکترونیکی بالأخص در محیط اینترنت چالش‌های جدیدی را فرا روی کتابداران و فهرست‌نویسان برای سازمان‌دهی و دسترس‌پذیر

نویسنده مسئول:

مینا توسلی فرحی

استادیار، کتابداری و اطلاع‌رسانی پزشکی

دانشگاه علوم پزشکی هرمزگان، بندرعباس، ایران

پست الکترونیکی: mfarahi20@gmail.com

تلفن: +۹۸۷۶۲۳۶۶۳۶۷

Dublin Core طوری طراحی شده است که غیرمتخصصان فهرست‌نویسی هم بتوانند آن را به‌کارگیرند، به‌گونه‌ای که نویسندگان و مسئولین وب سایت‌ها بدون داشتن تخصص فهرست‌نویسی قادر به استفاده از **Dublin Core** برای توصیف منابعشان هستند و بدین ترتیب منابعشان را برای موتورهای کاوش و سیستم‌های بازیابی قابل‌رؤیت‌تر می‌سازند [۴].

ONIX نیز به استاندارد اطلاق می‌شود که توسط ناشرین برای توزیع اطلاعات الکترونیکی کتاب‌هایشان به عمده و خرده‌فروشان و ناشران و کلیه کسانی که در فروش کتاب نقش دارند، به کار گرفته می‌شود.

ONIX به‌عنوان راه‌حلی برای دو مشکل ایجاد شد:

الف) نیاز به اطلاعات کامل‌تر کتاب‌ها به‌صورت آنلاین؛

ب) وجود تنوع فراوان فرمت‌های موردنیاز فروشندگان مطرح کتاب.

در سال ۱۹۹۹ انجمن ناشرین آمریکایی با همکاری عمده‌فروشان مهم کتاب و خرده‌فروشان آنلاین و سرویس‌دهندگان اطلاعات کتاب اقدام به ایجاد استاندارد بین‌المللی کردند که توسط آن ناشرین بزرگ و کوچک قادر به تبادل اطلاعات کتاب‌هایشان باشند. بدین ترتیب نسخه نخست آن **ONIX 1.0** - در سال ۲۰۰۰ انتشار یافت [۶].

ایجاد **UK LOM** از مقاله‌ای که در آوریل ۲۰۰۲ به گروه ابرده‌های آموزشی بریتانیا ارائه شد، سرچشمه می‌گیرد. این مقاله فراخوانی برای تشکیل انجمنی از دست‌اندرکاران برای شناسایی روش‌های شایع در بریتانیا جهت استفاده از ابرده‌ها در محتوای بسته‌های آموزش الکترونیکی بود. هدف اصلی استاندارد **LOM** افزایش قابلیت همکاری نمونه‌های ابرده‌ها و پروفایل‌های نرم‌افزارهای ایجادشده در بخش آموزش انگلستان است. بر اساس تعریف مؤسسه **IEEE** در سال ۱۹۹۰ می‌توان قابلیت همکاری متقابل را در چنین بستری به‌صورت زیر تعریف کرد: «توانایی دو یا چند سیستم یا مؤلفه‌های آن‌ها برای تبادل اطلاعات و استفاده از اطلاعاتی که ردوبدل شده‌اند» [۷].

در پیشینه‌های موردبررسی، طاهری ضمن معرفی و بررسی نقاط قوت و ضعف طرح فراداده‌ای **Dublin Core** و قالب نسخه ۲۱ **MARC** در سازمان‌دهی منابع اطلاعاتی شبکه جهانی وب، با استفاده از روش تحلیلی، کارایی دو طرح مذکور را در نمایه‌سازی اطلاعات الکترونیکی باهم مقایسه کرده است. نتیجه مقایسه نشان داد که درنهایت، قالب نسخه ۲۱ **MARC** برای ذخیره، پردازش و مبادله اطلاعات محیط وب، مناسب‌تر است.

گرفته‌شده بود، اما بیشترین کاربرد این اصطلاح با شکل‌گیری شبکه وب در سال ۱۹۹۳ صورت گرفته است.

اولین کاربردهای ابرده‌ها در قالب انواع مختلف **MARC** در سال‌های ۱۹۸۰، ۱۹۷۰، ۱۹۶۰ ظاهر شده بود. افزایش گسترده ابرده‌ها که از طریق صفحه گسترده وب قابل‌دسترس است، ضرورت‌هایی را برای ایجاد و استفاده از استانداردهای مربوط به ابرده‌ها به وجود آورده است. ابرده‌ها، کاربردهای مهمی دارد که ارزیابی و افزایش دقت بازیابی منابع شبکه‌ای را شامل می‌شود [۲]. هدف غایی تمامی فعالیت‌های کتابخانه دسترسی محققین به اطلاعات است، برای اینکه این دسترسی محقق شود، لازم است که اطلاعات گردآوری‌شده سازمان‌دهی شود. به‌منظور تحقق استانداردهای مناسب، طرح‌های ابرده‌های متعددی به وجود آمده‌اند که هدفشان دسترسی مؤثر به اطلاعات است [۳].

در این مطالعه، چهار ابرده‌ها موردبررسی در پایگاه‌های کتاب الکترونیکی شامل: **Dublin Core**, **MARC (Machine-Readable Catalogue)**, **ONIX (Online Information Exchange)**, **UK LOM (UK Learning Object Metadata)** بودند و مفاهیم نظری به شرح زیر، توضیح داده‌شده‌اند.

اصطلاح ابرده‌ها در بین جامعه اینترنت به مفهوم آنچه که کتابداران به‌عنوان فهرست‌نویسی اطلاعات یا توصیف منابع است، به‌کاربرده می‌شود [۴]. ابرده **MARC (Machine-Readable Cataloging)** تا اواسط دهه ۹۰ شناخته‌شده‌ترین ابرده توصیفی استاندارد بود که به‌عنوان قالبی قدرتمند ده‌ها سال در کتابخانه‌های سراسر جهان استفاده می‌شد. **MARC** به معنای رکورد فهرست‌نویسی ماشین‌خوان است؛ یعنی یک کامپیوتر قادر به خواندن و تفسیر داده‌ها در رکورد فهرست‌نویسی است. رکورد فهرست‌نویسی به معنای رکورد کتاب‌شناختی و یا اطلاعاتی است که به‌طور سستی بر روی یک کارت فهرست‌نویسی نشان داده‌شده است. رکورد کتاب‌شناختی، شامل توصیف مدرک، شناسه‌های اصلی، شناسه‌های افزوده، سرعت‌خوان‌های موضوعی، رده‌بندی و شماره راهنمای مدرک است (رکوردهای **MARC** حاوی اطلاعات بیشتری است). کتابخانه‌ها با استفاده از استاندارد **MARC** از دوباره‌کاری جلوگیری به عمل می‌آورند و امکان به اشتراک‌گذاری منابع کتاب‌شناختی را می‌یابند [۵].

Dublin Core یک ابرده پانزده عنصری است که به‌منظور سهولت در بازیابی منابع الکترونیکی و در اصل برای توصیف منابع وب توسط پدیدآورندگان ایجادشده است، اما به‌تدریج موردتوجه کتابخانه‌ها و موزه‌ها برای توصیف منابعشان قرار گرفته است.

است؛ به عنوان ابزاری، برای فهرست‌نویسی منابع دیجیتالی و بهترین ابزار فهرست‌نویسی که مورد مقبولیت بین‌المللی قرار گرفته است، محسوب می‌شود [۳].

در مطالعه‌های دیگر، ابادر و بابایی در پژوهشی که از نوع پیمایشی توصیفی بود، در مجموع ۱۶ صفحه اصلی وبسایت کتابخانه‌های مرکزی دانشگاه‌های علوم پزشکی ایران را با استفاده از دو مرورگر Internet Expelorer و Net Scape مورد بررسی قرار دادند و با مشاهده Source از طریق منوی View عناصر فراداده‌ای موجود در صفحه HTML هر یک از وبسایت‌ها را استخراج کردند و در چک‌لیست قرار دادند. در وبسایت‌های مورد بررسی آن‌ها، هیچ‌یک از وبسایت‌ها از عناصر فراداده‌ای Dublin Core استفاده نکرده‌اند و در طراحی تمام این وبسایت‌ها از فراداده‌های عام زبان نشانه‌گذاری فرامتن استفاده شده بود. نتایج حاکی از آن بود که طراحان صفحات وب به ابر برچسب‌ها (Meta Tags) بیشتر اهمیت می‌دهند تا طرح‌های فراداده‌ای خاصی همچون Dublin Core [۱۰].

شفیعی و همکاران از روش تحلیل محتوا، عناصر ابر داده‌ای Dublin Core و مجموعه ابر نشانه‌های زبان نشانه‌گذاری فرامتن موجود در فیلدهای موجود در صفحات جست‌وجوی پایگاه‌های اطلاعاتی پزشکی MD Consult, Medline Ovid, Pub Med, Medline Plus, Iranmedex, Pars Medline و ایران و جهان را بررسی کردند. یافته‌ها نشان داد عناصر ابر داده‌ای Dublin Core و ابر نشانه‌های زبان نشانه‌گذاری فرامتن در صفحات باز یابی شده بسیار کم هستند. این امر در همه پایگاه‌های اطلاعاتی به جز پایگاه اطلاعاتی Medline Plus مشاهده شد و تنها در پایگاه اطلاعاتی Medline Plus عناصر ابر داده‌ای Dublin Core وجود داشت. بیشترین تعداد فیلد ابر داده‌ای مربوط به پایگاه اطلاعاتی PubMed و کمترین تعداد مربوط به پایگاه اطلاعاتی Medline Plus بود، این مطالعه نشان داد که فیلدها و عناصر ابر داده‌ای در پایگاه‌های اطلاعاتی پزشکی ایران و جهان بسیار کم و نامناسب به کار گرفته شدند [۱۱] و اما پیشینه‌های مورد بررسی در خارج کشور، Fast و Campbell پیش‌بینی کردند که قابلیت‌های ابر داده‌ها در آینده می‌تواند کتابخانه‌های دانشگاهی را قادر سازد که اطلاعات موجود در وب معنایی را در فهرست کتابخانه‌هایشان به کار گیرند، در حالی که معماری‌های ابر داده‌ها در حال حاضر استفاده از وب را تنها جهت تبادل داده‌های کتاب‌شناسی که توسط کتابخانه‌ها ایجاد شده‌اند را میسر می‌سازد [۱۲].

طرح فراداده‌ای Dublin Core، نه تنها از لحاظ تعداد عناصر توصیف، نسبت به نسخه ۲۱ MARC قابلیت‌های کمتری دارد، بلکه از لحاظ انواع قالب برای انواع اطلاعات و نیز انواع قالب برای انواع مواد اطلاعاتی، با قالب نسخه ۲۱ MARC قابل مقایسه نیست [۸].

در مطالعه‌های دیگر نیز شفییعی و همکاران در پژوهش خود، ۹۰ صفحه وب از طریق جستجو در موتورهای جستجوی عمومی (google, yahoo و MSN) را به عنوان نمونه انتخاب کردند و میزان حضور عناصر ابر داده‌ای (Dublin Core) و ابر نشانه‌های زبان نشانه‌گذاری فرامتن) در این صفحات و وجود رابطه معنادار بین حضور عناصر ابر داده‌ای و نوع موتور جستجو را بررسی کردند. نتایج، نشان‌دهنده حضور بسیار کم عناصر ابر داده‌ای Dublin Core و ابر نشانه‌های زبان نشانه‌گذاری فرامتن در صفحات باز یابی شده بود که بیانگر کاربرد بسیار کم عناصر ابر داده‌ای در صفحات وب بود [۹].

همچنین کریمی و فیاض در مطالعه‌ای ضمن معرفی مهم‌ترین استانداردهای سازمان‌دهی در کتابخانه‌های دیجیتال، به این نتیجه رسیدند، قالب فراداده‌ای Dublin Core با ساختاری ساده و قابل استفاده حتی برای پدیدآور اثر و برای طیف وسیعی از منابع شبکه‌ای به وجود آمده است. استاندارد MARC دارای ساختاری پیچیده‌تر، جزئی‌تر و کاربردی‌تر، نسبت به طرح فراداده Dublin Core است.

طرح فراداده‌ای توصیف شی (MODS: Metadata Object Description Schema) با ساختار ذخیره‌سازی مبتنی بر XML (Extensible Markup Language) و منطبق با MARC، دارای مجموعه عناصری بیشتر از Dublin Core و ساده‌تر از MARC است و با هدف پاسخگویی به نیازهای اطلاعاتی در محیط دیجیتال تهیه شده است. قالب فراداده‌ای طرح کدگذاری متن (Text Encoding Initiative) و توصیف آرشیوی رمزگذاری شده (Encoded Archival Description)، بیشتر در آرشیوها مورد استفاده قرار می‌گیرد و قالب فراداده‌ای خدمات مکان‌یاب اطلاعات دولتی (Government Information Locator Services) برای مکان‌یابی اطلاعات مناسب نهادها و ادارات دولتی کاربرد دارد. فراداده‌ای توصیف مستند (Metadata Authority Description: MADS Schema) در کنار استاندارد فراداده‌ای توصیف شی، به منظور توصیف افراد حقیقی و حقوقی، رویدادهای مهم و شناسه‌های موضوعی و جغرافیایی به وجود آمده است. قواعد استاندارد توصیف و دستیابی به منابع (RDA: Resource Description and Access) که در واقع ویرایش سوم قواعد فهرست‌نویسی Englo-American

طرح فراداده‌های Dublin Core و نسخه ۲۱ MARC مبتنی بر XML منتشرشده در محیط‌های اطلاعاتی پیوسته پویا از لحاظ دسترس‌پذیری توسط موتورهای کاوش نسبت به دیگری برتری ندارند [۱۵].

Debus و همکاران استفاده از ابر داده ONIX را در کتابخانه کنگره مورد بررسی قرار دادند. آنان معتقدند که استاندارد ONIX برای کتابداران نویدبخش است، چراکه بسیاری از عناصر داده در داخل آن راه می‌توان به صورت استاندارد MARC نگاشت. کتابخانه کنگره برای ایجاد توصیف‌های کتاب‌شناسی MARC از ابر داده‌های ONIX ناشرین برای انتشارات جدید دریافتی از طریق فهرست‌نویسی الکترونیکی در برنامه انتشارات کتاب، تبدیل‌کننده ONIX به MARC را ایجاد کرد. نتایج نشان داد که استفاده از مبدل ONIX به MARC، زمان لازم برای ایجاد توصیف‌های کتاب‌شناسی را کاهش می‌دهد، باعث تسهیل گنجاندن ابر داده‌ها در پیشینه‌های کتاب‌شناسی می‌شود و برای کارکنان فهرست‌نویسی پیشینه‌هایی را فراهم می‌سازد که قابل مقایسه با پیشینه‌هایی با کیفیت فهرست‌نویسی غیر بنیادی است [۱۶].

Baker در مطالعه خود مطرح ساخت که چگونه مسائل پیرامون ابر داده Dublin Core، نمونه بارز مشکلاتی است که نیازمند حل و فصل است تا داده‌های کتابخانه‌ها با استانداردهای سستی کیفیت و سازگاری مطابقت داشته باشند و در عین حال در محیط داده‌های مرتبط با منابع آن‌ها سازگار باشند. او در مقاله‌اش نشان داد که چگونه استاندارد Dublin Core که عموماً به عنوان یک قالب ساده پیشینه در نظر گرفته می‌شود، به طور روزافزونی به عنوان واژگان RDF برای استفاده در ابر داده‌ها و بر اساس یک مدل «شرح» دیده می‌شود و چگونه برای پر کردن شکاف بین این مدل‌ها روش‌های جدید ابر داده‌ها ظهور کرده‌اند. نتایج نشان داد در کتابخانه‌ها باید امکان پیوند و ادغام داده‌ها با منابعی که مطابق استانداردهای دیگر هستند، وجود داشته باشد [۱۷].

بنابراین نقش ویژه طرح‌های ابر داده‌ای در سازمان‌دهی، نظم‌بخشی، بازیابی دقیق منابع اطلاعاتی الکترونیکی [۱] و هزینه‌های خرید آن‌ها از جمله مسائل مهمی به شمار می‌آیند که اهمیت بررسی ابر داده‌ها را نشان می‌دهد، حال این مسئله پیش می‌آید که پایگاه‌های کتاب الکترونیکی از نظر به‌کارگیری ابر داده‌ها در چه وضعیتی قرار دارند؟ لذا هدف از این مطالعه تعیین ابر داده‌های پایگاه‌های کتاب الکترونیکی است.

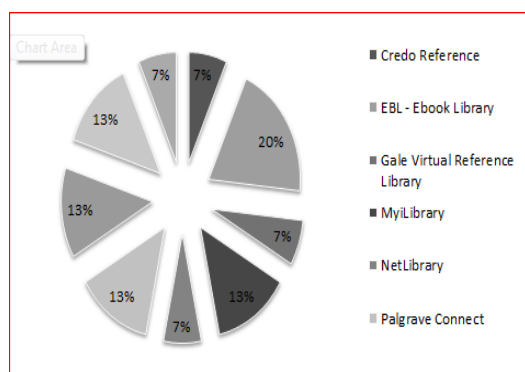
روش‌ها:

در مطالعه دیگر، علیجانی و جوکار باهدف بررسی تمایل کتابداران کتابخانه‌های ملی جهان و میزان توجه آنان به استاندارد Dublin Core، با استفاده از چک‌لیست و مراجعه به منبع صفحه اصلی وبسایت هر کتابخانه به بررسی عناصر فراداده Dublin Core در وبسایت‌های ۷۰ کتابخانه ملی پرداختند. عناصر Dublin Core به سه دسته عناصر توصیف محتوا، عناصر مربوط به حق مالکیت معنوی و عناصر مربوط به نمایش ساختاری تقسیم شدند. نتایج نشان داد که فقط ۱۴ (۲۰ درصد) کتابخانه‌های ملی از عناصر Dublin Core استفاده کرده‌اند. کتابخانه ملی عربستان با استفاده از هر ۱۵ عنصر Dublin Core، حداکثر عناصر را و کتابخانه ملی سوئیس با استفاده از ۲ عنصر، حداقل عناصر Dublin Core را استفاده کرده‌اند [۱۳].

Kurtz در مطالعه‌ای ضمن بررسی اجمالی Dublin Core و DSpace به مطالعه و مقایسه مخازن Repositories سازمانی سه دانشگاه New Mexico, Washington, OHIO State University پرداخت. دانشگاه‌های مورد بررسی برای ایجاد و مدیریت مخازن خود از Dublin Core و DSpace استفاده می‌کردند. سپس از هر مخزن بیست رکورد به طور تصادفی نمونه‌گیری شد و با استفاده از ملاک‌های کیفیت ابر داده‌ها یعنی معیارهای جامعیت، دقت و سازگاری مورد بررسی قرار گرفتند. نتایج نشان داد که دانشگاه New Mexico برای نظارت بر روند آرشیو مخزن خود، از کتابداران استفاده می‌کند، در حالی که دو دانشگاه دیگر از دو استراتژی مختلف به عنوان بخشی از فرایند خود-آرشیو (Self-archiving) استفاده می‌کردند. این مطالعه نشان داد که آرشیوهای تحت نظارت کتابداران در دانشگاه New Mexico کامل‌ترین و دقیق‌ترین رکوردها را برای مدخل‌های DSpace را دارا بودند [۱۴].

فرج‌پهلوی و طباطبائی پژوهشی باهدف تبیین نمایه‌سازی و رتبه‌بندی اشیای محتوایی حاوی عناصر فراداده‌های Dublin Core و نسخه ۲۱ MARC مبتنی بر زبان نشانه‌گذاری گسترش‌پذیر (XML) در محیط‌های اطلاعاتی پیوسته پویا توسط موتورهای کاوش google و yahoo و مقایسه آن‌ها با یکدیگر با رویکرد تحلیلی - مقایسه‌ای انجام دادند. جامعه پژوهش را ۱۰۰ شیء محتوایی (کتاب الکترونیکی) مبتنی بر زبان نشانه‌گذاری گسترش‌پذیر (XML) در قالب دو گروه تشکیل داد؛ گروه اول، حاوی عناصر فراداده‌های Dublin Core و گروه دوم، حاوی عناصر فراداده‌های نسخه ۲۱ MARC بود که بر روی وبسایت <http://www.marcdcmi.ir> انتشار یافته بودند. ابزار گردآوری داده‌ها سیاهه واریسی بود. در مجموع، یافته‌ها حاکی از آن بود که هیچ‌یک از دو

هیچ کدام از پایگاهها استفاده نشده است. به طور کلی، بیشترین میزان استفاده از ابر دادهها، MARK و سپس ابر داده ONIX بود. در مجموع ۴۱/۶۶ درصد از ابر دادهها در پایگاهها به کار برده شده بود. ابر دادههای مورد استفاده در پایگاههای مورد نظر به ترتیب MARK به صورت (۱۰۰ درصد)، ONIX (۵۵/۵۵ درصد)، Dublin Core (۱۱/۱۱ درصد) بود و UK LOM در هیچ کدام از پایگاهها مورد استفاده قرار نگرفت. به طور کلی بیشترین میزان توجه به ابر دادهها در پایگاههای مورد نظر، MARC و کمترین میزان توجه به ابر دادهها UK LOM بود. همچنین نتایج یافتهها نشان داد که یک پایگاه (۷۵ درصد) ابر دادهها، چهار پایگاه (۵۰ درصد) ابر دادهها و چهار پایگاه (۲۵ درصد) ابر دادهها را به کار برده اند (جدول شماره ۲).



نمودار شماره ۱- نمودار دایره ای توزیع فرآوانی کاربرد ابر دادهها در پایگاههای کتاب الکترونیکی

این مطالعه از نوع کاربردی و به روش پیمایشی در سال ۱۳۹۳ انجام شده است. جامعه آماری شامل ۹ پایگاه اطلاعاتی کتاب الکترونیکی غیر رایگان بود؛ که به روش هدفدار انتخاب شده اند. اطلاعات مربوط به ابر دادههای مورد استفاده در وبسایت «ابزار ارزیابی پایگاههای علمی» (<http://adat.crl.edu>) که به مقایسه پایگاهها از لحاظ مختلف می پردازد، قابل نمایش بود. در این وبسایت با انتخاب پایگاهها می توان نتایج مقایسه را مشاهده کرد و کاربرد انواع ابر دادهها، در آنها را با هم مقایسه کرد. ۹ پایگاه اطلاعاتی کتاب الکترونیک به همراه ابر دادههای به کار برده شده در آن، به صورت جدول نمایش داده شدند، چنانچه ابر دادهای در پایگاه وجود داشته باشد، با علامت (√) و در غیر این صورت با علامت (-) نمایش داده شده است. دادههای به دست آمده، بعد از ورود به نرم افزار آماری Excel مورد تحلیل قرار گرفته اند. آدرس پایگاههای اطلاعاتی کتاب الکترونیک در جدول شماره یک نشان داده شده است (جدول شماره ۱).

یافتهها:

پس از مرحله گردآوری دادهها، توزیع فرآوانی اطلاعات جمع آوری شده پایگاهها، به صورت نمودار شماره یک و جدول شماره دو نشان داده شده است.

جدول شماره دو، نشانگر توزیع فرآوانی پایگاهها بر اساس ابر دادههای مورد استفاده است؛ بنابراین جدول شماره دو، حاکی از آن است که ابر داده MARC در تمامی پایگاهها، ولی ابر داده Dublin Core تنها در یک پایگاه به نام EBL - Ebook Library و ابر داده ONIX در ۵ پایگاه به کار گرفته شده است و ابر داده UK LOM در

جدول شماره ۱- آدرس و نام پایگاههای اطلاعاتی کتاب الکترونیک

آدرس پایگاه	نام پایگاه	ردیف
www.credoreference.com	Credo Reference	۱
http:// www.ebrary.com	Ebrary	۲
http://www.dawsonera.com	Dawsonera	۳
http://www.ebookstore.tandf.co.uk/html/splash.asp	Taylor & Francis eBookstore	۴
http://www.netlibrary.com	NetLibrary	۵
http://www.palgraveconnect.com/index.html	Palgrave Connect	۶
http://www.gale.cengage.com	Gale Virtual Reference Library	۷
http://www.ebilib.com	EBL - Ebook Library	۸
http://lib.mylibrary.com	MyLibrary	۹

جدول شماره ۲- توزیع فراوانی ابردادهای موردبررسی در پایگاههای کتاب الکترونیکی

نام پایگاه	Dublin	UK LOM	ONIX	MARC	جمع فراوانی	درصد فراوانی
Credo Reference	-	-	-	√	۱	۲۵
EBL - Ebook Library	√	-	√	√	۳	۷۵
Gale Virtual Reference Library	-	-	-	√	۱	۲۵
MyiLibrary	-	-	√	√	۲	۵۰
NetLibrary	-	-	-	√	۱	۲۵
Palgrave Connect	-	-	√	√	۲	۵۰
Taylor & Francis eBookstore	-	-	√	√	۲	۵۰
Dawsonera	-	-	√	√	۲	۵۰
Ebrary	-	-	-	√	۱	۲۵
جمع فراوانی	۱	۰	۵	۹	۱۵	۴۱/۶۶
درصد فراوانی	۱۱/۱۱	۰	۵۵/۵۵	۱۰۰	-	-

بحث و نتیجه گیری:

در مطالعه Kurtz، دانشگاه‌های موردبررسی برای ایجاد و مدیریت مخازن سازمانی خود از Dublin Core و DSpace استفاده می‌کردند [۱۴]، علیجانی و جوکار، در مطالعه خود نشان دادند که فقط ۱۴ (۲۰ درصد) کتابخانه ملی از عناصر Dublin Core استفاده کرده‌اند و بسیاری از کتابخانه‌های ملی که کتابخانه‌های اصلی جهان هستند، در طراحی صفحات وب خود، توجه زیادی به عناصر Dublin Core معطوف نداشته‌اند، همچنین در مطالعه حاضر، ۱۱/۱۱ از پایگاه‌ها، از ابرداده Dublin Core استفاده کردند [۱۳].

فرچ‌پهلوی و طباطبائی نیز اشاره کردند که هیچ‌یک از دو طرح ابردادهای Dublin Core و نسخه ۲۱ MARC مبتنی بر XML منتشر شده در محیط‌های اطلاعاتی پیوسته پویا از لحاظ دسترس‌پذیری توسط موتورهای کاوش نسبت به یکدیگر برتری ندارند [۱۵].

همچنین در مطالعه Campbell و Fast، پیش‌بینی کردند که قابلیت‌های ابردادها در آینده می‌تواند کتابخانه‌های دانشگاهی را قادر سازد که اطلاعات موجود در وب معنایی را در فهرست کتابخانه‌هایشان به‌کارگیرند، درحالی‌که معماری‌های ابردادها در حال حاضر استفاده از وب را تنها جهت تبادل داده‌های کتاب‌شناسی که توسط کتابخانه‌ها ایجاد شده‌اند را میسر می‌سازد [۱۲]. درنهایت همان‌طور که کریمی و فیاض اشاره کردند با توجه به اینکه استفاده از تمامی استانداردهای سازمان‌دهی نه ضروری و نه امکان‌پذیر است، پیشنهاد می‌شود از میان استانداردهای عمومی، استاندارد مناسب انتخاب شود و با توجه به نیاز می‌توان از یک یا چند استاندارد مستند استفاده کرد [۳].

طاهری در مطالعه خود قالب نسخه ۲۱ MARC را برای ذخیره و پردازش اطلاعات محیط وب مناسب‌تر دانست [۸].

در یک نگاه کلی، ابردادهای مورد استفاده در پایگاه‌های موردنظر به ترتیب MARK، ONIX، Dublin Core و UK LOM بود و به‌طور کلی بیشترین میزان استفاده به ابردادها در پایگاه‌های موردنظر، MARC و کمترین میزان استفاده به ابردادها UK LOM بود. همچنین نتایج یافته‌ها نشان داد که یک پایگاه (۷۵ درصد) ابردادها، چهار پایگاه (۵۰ درصد) ابردادها و چهار پایگاه (۲۵ درصد) ابردادها را به‌کار برده‌اند. به‌کارگیری استاندارد MARC امکان تبادل اطلاعات را تقویت می‌کند [۵]. انجمن ناشران آمریکایی نیز به‌منظور تبادل اطلاعات کتاب‌ها، اقدام به ایجاد ابرداده ONIX کردند، در این مطالعه نیز بیشترین استفاده از ابرداده ابتدا MARC و سپس ابرداده ONIX (۵۵/۵۵ درصد) بود. Debus و همکاران اشاره کردند، استفاده از مبدل ONIX به MARC، زمان لازم برای ایجاد توصیف‌های کتاب‌شناسی را کاهش می‌دهد، همچنین باعث تسهیل گنجاندن ابردادها در پیشینه‌های کتاب‌شناسی می‌شود [۱۶].

در مطالعه شفیع و همکاران همچون مطالعه حاضر تنها یک پایگاه از ابرداده Dublin Core استفاده کرده بود [۱۱]، در مطالعه ابازر و بابایی هیچ‌کدام از وب سایت‌ها از ابرداده Dublin Core استفاده نکرده‌اند [۱۰]. در مطالعه دیگر شفیع و همکاران، نشان از حضور بسیار کم عناصر ابرداده Dublin Core بود [۹]، در مطالعه Baker، اشاره شد که چگونه در استاندارد Dublin Core که عموماً به‌عنوان یک قالب ساده پیشینه در نظر گرفته می‌شود، در کتابخانه‌ها باید امکان پیوند و ادغام داده‌ها با منابعی که مطابق استانداردهای دیگری هستند، وجود داشته باشد [۱۷].

از جمله محدودیت های این پژوهش، عدم بررسی مستقیم ابر داده های پایگاه ها؛ به دلیل غیر رایگان بودن آنها بود، به همین دلیل از ابزار ارزیابی پایگاه های علمی (ADAT) که به مدیریت منابع الکترونیکی می پردازد، استفاده شده است.

با رشد تصاعدی منابع الکترونیکی در محیط وب، لزوم بازیابی سریع اطلاعات، صرفه جویی در وقت و هزینه خرید این دسته از منابع الکترونیکی، جهت بازیابی دقیق اطلاعات، بهتر است منابعی که شامل ابر داده های مناسب هستند توسط متخصصان اطلاع رسانی خریداری شوند. طبق یافته ها و نتایج حاصل از این مطالعه، لازم است در طراحی پایگاه ها از ابر داده مناسب استفاده شود تا کاربران در بازیابی اطلاعات به نتایج مطلوبی دست یابند. به طور کلی ابر داده MARC در وب سایت ها بیشترین میزان کاربرد را در پایگاه ها داشته است، بعد از MARC بیشترین میزان کاربرد ابر داده، به ONIX و سپس Dublin Core تعلق می گیرد و ابر داده UK LOM در هیچ کدام از پایگاه ها مورد توجه قرار نگرفته بود. بر اساس نتایج اخذ شده پیشنهاد می گردد متخصصان اطلاع رسانی و طراحان پایگاه ها در ایجاد پایگاه های الکترونیکی، ابر داده های مناسب را در نظر بگیرند. همچنین پیشنهاد می شود، مطالعه مشابهی در مورد شناسایی انواع عناصر ابر داده ای مورد استفاده در ابر داده های مورد نظر انجام گیرد و با توجه به اهمیتی که عناصر ابر داده ای در ابر داده های مورد نظر، برای بازیابی دقیق تر اطلاعات دارند، توصیه هایی به متخصصان اطلاع رسانی و طراحان پایگاه ها جهت به کارگیری آنها می شود.

در مطالعه فیاض و کریمی ابر داده MARC نسبت به ابر داده Dublin Core پیچیده تر، کاربردی تر و جزئی تر شناخته شده، در مطالعه حاضر نیز ابر داده MARC که کاربردی تر بود، بیشتر استفاده شد [۳]. در حال حاضر، قالب نسخه ۲۱ MARC، قالب های مختلفی برای انواع اطلاعات ارائه می دهد که عبارتند از: اطلاعات کتاب شناختی، قالب مستندات، اطلاعات رده بندی، اطلاعات جامعه و اطلاعات مربوط به موجودی. همچنین، از آنجا که نسخه ۲۱ MARC، بر اساس قواعد فهرست نویسی Englo-American طراحی شده، قالب های خاص هر نوع رسانه اطلاعاتی را از چاپی تا مواد سمعی و بصری و منابع الکترونیکی، شامل می شود؛ در صورتی که Dublin Core فقط منحصر به منابع الکترونیکی است و سایر رسانه های اطلاعاتی را پشتیبانی نمی کند و برای اطلاعات خاص فاقد قالب خاص نیز است؛ بنابراین قالب فراداده ای نسخه ۲۱ MARC برای ذخیره، پردازش، تبادل و بازیابی منابع شبکه جهانی وب، مناسب تر و کارآمدتر است. تعدد عناصر کتاب شناختی، وجود قالب های مختلف اطلاعات، سابقه دیرینه تولید و کاربرد قالب MARC، همراه با تعدد ویرایش ها، تهیه کنندگان پیشینه ها، گستره کاربردها، اجماع بین المللی و امکان کنترل نامها و موضوعها، از مهم ترین دلایل برتری نسخه ۲۱ MARC نسبت به Dublin Core است [۱۱]. در این مطالعه نیز از ابر داده MARC به صورت ۱۰۰ درصد استفاده شده بود. در واقع کاربرد ابر داده ها باعث افزایش دقت بازیابی و تسهیل جستجو در منابع الکترونیکی می شوند.

References

- Mohamadi SH. Metadata: Concepts and Applications. Scientific Communication, Monthly Journal of Irandoc 2004; 2(3): 48-53. [In Persian]
- Biranavand A, Sharif Moghadam H. The Comparison of Metadata with the Traditional Cataloging Standards: A Prospective Approach to the Function of Metadata in Resource Organization. Scientific Communication, Monthly Journal of Irandoc 2010; 16(1): 1-11. [In Persian]
- Karimi M, Fayaz S. Standards of information organization in digital libraries. Scientific Communication, Monthly Journal of Irandoc 2010; 16(2): 1-10. [In Persian]
- Weibel S. The Dublin Core: A Simple Content Description Model for Electronic Resources. Bulletin of the American Society for Information Science and Technology 1997; 24(1): 9-11.
- Furrie B. Understanding marc: machine-readable cataloging. 8th ed. Network Development and Marc Standards Office Loc, Editor. Washington dc: cataloging distribution service. Library of congress; 2009. Available from: <https://catalog.loc.gov/vwebv/search?searchCode=LCCN&searchArg=2009044593&searchType=1&permalink=y>
- Daly F. Onix: The Metadata Standard for the Information and Entertainment Industries. Publishing Research Quarterly 2002; 18(2): 28-40.
- Campbell L. Learning Object Metadata (LOM). In: Ross S, Day M, Editors. DCC Digital Curation Manual; 2007; Available from: <https://www.era.lib.ed.ac.uk/bitstream/handle/1842/3334/Campbell%20learning-object-metadata.pdf?sequence=1> (Accessed 12 July 2013)

8. Taheri SM. The Comparison of the Effectiveness of Dublin Core Initiative with Marc 21 Format in Organizing the World Wide Web Information Resources. *Library and Information Science*. 2008; 43: 66-79. [In Persian]
9. Shafie Alavijeh S, Ghaebi A, Rezaie Sharifabadi S. Review Of Metadata Elements Within The Web Pages Resulting From Searching In General Search Engines. *Iranian Research Institute for Information Science and Technology (IRANDOC)*. 2009; 25(1): 71-89. [In Persian]
10. Abazari Z, Babaei K. Metadata Elements of Dublin Core and Those of Central Libraries Websites of Medical Universities in Iran: A Comparative Study. *Journal of Health Administration* 2011; 14(43): 7-18. [In Persian]
11. Shafie Alavijeh S, Rahimi A, Shafie Alavijeh P. A Comparison Of Metadata Elements And The Hypertext Markup Language In Records Retrieved From Specialized Medical Databases In Iran And Around The World. *National Studies on Librarianship and Information Organization* 2012; 23(2): 160-76. [In Persian]
12. Campbell DG, Fast KV. Academic Libraries And The Semantic Web: What The Future May Hold For Research-Supporting Library Catalogues. *The Journal of Academic Librarianship* 2004; 30(5): 382-390.
13. Alijani AS, Jowkar A. Dublin Core Metadata element set usage in national libraries' web sites. *Electronic Library* 2009; 27(3): 441-7. [In Persian]
14. Kurtz M. Dublin Core, Dspace and a Brief Analysis of Three University Repositories. *Information Technology and Libraries* 2010; 29(1): 40-6.
15. Farajpahlou AH, Tabatabai F. How Are Xml-Based Marc 21 And Dublin Core Records Indexed And Ranked By General Search Engines In Dynamic Online Environments? *Aslib Proceedings* 2011; 63(6): 586-592.
16. Debus-Lopez K, Williamson D, Saccucci C, Williams C. Bringing Publisher Metadata Directly To The Library Use Of Onix At The Library Of Congress. *Libr Resour Tech Serv* 2012; 56(4): 266-79.
17. Baker T. Libraries, Languages of Description, And Linked Data: A Dublin Core Perspective. *Library Hi Tech* 2012; 30(1): 116-33.

Survey of e-book databases with emphasis on metadata

Elaheh Ommati¹ Mina Tavassoli Farahi^{2*}

1. Master of Science in Medical Library and Information Sciences, Clinical Research Development Center, 17 Shahrivar Educational, Research and Remedial Center, Guilan University of Medical Sciences, Rasht, Iran.

2. Assistant Professor, Medical Library and Information Science, Health Information Management Research Center, Hormozgan University of Medical Sciences, Bandar Abbas, Iran.

(Received 01 Oct, 2014 Accepted 11 Jan, 2015)

Original Article

ABSTRACT

Introduction: With the exponential growth of electronic resources on the Web, the application of metadata has enhanced the precision of retrieval and facilitated the search of electronic resources. Hence, the aim of this study was to determine the application of metadata in e-book databases.

Methods: This study is an applied work, which was carried out through survey methods in 2013. The population consisted of 9 e-books databases, including Gale Virtual Reference Library, EBL – E-book Library, Credo Reference, MyiLibrary, Palgrave Connect, NetLibrary, Ebrary, Dawsonera, Taylor & Francis eBookstore which were selected by an intentional method. Data was collected from Academic Database Assessment Tool Website (<http://adat.crl.edu>), which compares databases from different perspectives, through direct observation. The obtained data was analyzed using 'Excel' software.

Results: Results showed that metadata used in e-books databases, were as follows, MARC metadata (100 per cent), Onix (55.55 per cent), Dublin Core (11.11 per cent) and UK Lom was not used in any of the databases.

Conclusion: Results suggested that information specialists and database designers should consider applying appropriate metadata in developing e-book databases.

Key words: Metadata, Onix, Dublin Core, UK LOM, MARC.

Citation: Ommati E, Tavassoli Farahi M. Survey of e-book databases with emphasis on metadata. *Journal of Modern Medical Information Sciences*. 2015; 1(1): 33-41.

Correspondence:

Mina Tavassoli Farahi

Assistant Professor, Medical Library and Information Science

Hormozgan University of Medical Sciences, Bandar Abbas, Iran.

Tel: +98 7633666367

Email: mfarahi20@gmail.com