

بررسی پایگاههای کتاب الکترونیکی با تأکید بر ابر داده

الهه امتی^۱ دکتر مینا توسلی فرحی^۲

۱. کارشناس ارشد، کتابداری و اطلاع‌رسانی پزشکی، دانشکده پیرا پزشکی، دانشگاه علوم پزشکی تهران، تهران، ایران.

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

مجله اطلاع‌رسانی پزشکی نوین؛ سال اول؛ شماره اول؛ بهار ۱۳۹۴؛ صفحات ۳۳-۴۱

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.

روش کار: 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* بوده که به روش هدفدار انتخاب شده‌اند. اطلاعات مربوط به ابر داده‌های مورد استفاده در وب سایت "انبار ارزیابی پایگاههای علمی" (<http://adat.crl.edu>) که به مقایسه پایگاهها از جنبه های مختلف می‌پردازد، قابل نمایش بود. در این وب سایت با انتخاب پایگاهها می‌توان نتایج را مشاهده کرد و کاربرد انواع ابر داده‌ها در آنها را با هم مقایسه کرد. داده‌های به دست آمده بعد از ورود به نرم‌افزار آماری *Excel* مورد تجزیه و تحلیل قرار گرفته‌اند.

یافته‌ها: یافته ها نشان داد در پایگاههای موردنظر، 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 هیچکدام از پایگاهها استفاده نشده بود.

نتیجه‌گیری: براساس 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، اونیکس، دابلین کور، یوکی لام، مارک.

دریافت مقاله: ۹۳/۰۷/۰۹ اصلاح نهایی: ۹۳/۱۰/۱۰ پذیرش مقاله: ۹۳/۱۰/۲۱

مقدمه:

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 گرفته شده بود، اما بیشترین کاربرد این اصطلاح با شکل‌گیری شبکه وب در سال ۱۹۹۳ صورت گرفته است. اولین کاربردهای ابر داده در

تکامل عصر الکترونیک منجر به توسعه پایگاههای الکترونیکی کتاب در موضوعات مختلف شده است و این پایگاهها نقش به سزایی 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 بالاخص در محیط اینترنت چالش‌های جدیدی را فرا روی کتابداران و فهرست‌نویسان برای سازماندهی و دسترس پذیرسازی این منابع قرار داده است. اقدامات

نویسنده مسئول:

دکتر مینا توسلی فرحی

مرکز تحقیقات مدیریت اطلاعات سلامت، دانشگاه علوم پزشکی هرمزگان

بندرعباس - ایران

پست الکترونیکی: mfarahi20@gmail.com

تلفن ۹۸۷۶۱-۶۶۶۶۳۶۷

قالب مارک‌های مختلف در سال‌های ۱۹۸۰، ۱۹۷۰، ۱۹۶۰ ظاهر شده بود. افزایش گسترده ابرداده که از طریق صفحه گسترده وب قابل دسترس است، ضرورت‌هایی را برای ایجاد و استفاده از استانداردهای مربوط به ابرداده به وجود آورده است. ابرداده، 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 گردآوری شده سازماندهی شود. به منظور تحقق استانداردهای مناسب، طرح‌های ابرداده‌ای متعددی به وجود آمده‌اند که هدفشان دسترسی مؤثر به اطلاعات است [۳].

در این مطالعه، چهار ابرداده مورد بررسی 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 داده شده‌اند.

اصطلاح ابرداده در بین جامعه اینترنت به مفهوم آنچه که کتابداران به عنوان فهرست‌نویسی اطلاعات یا توصیف منابع است، به کار برده می‌شود [۴]. ابرداده مارک تا اواسط دهه ۹۰ شناخته شده‌ترین ابرداده توصیفی استاندارد بود که به عنوان قالبی قدرتمند دهها سال در کتابخانه‌های سراسر جهان استفاده می‌شد. مارک 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 به طور سستی بر روی یک کارت فهرست‌نویسی نشان داده شده است. رکورد کتاب‌شناختی، شامل توصیف مدرک، شناسه‌های اصلی، شناسه‌های افزوده، سرعنوان‌های موضوعی، رده‌بندی و شماره راهنمای مدرک است (رکوردهای مارک حاوی اطلاعات بیشتری است). کتابخانه‌ها با استفاده از استاندارد مارک از دوباره کاری جلوگیری به عمل می‌آورند و امکان به اشتراک‌گذاری منابع کتاب‌شناختی را می‌یابند [۵].

دابلین کور یک ابرداده پانزده عنصری است که به منظور سهولت در بازیابی منابع الکترونیکی و در اصل برای توصیف منابع وب توسط پدیدآورندگان ایجاد شده است، اما به تدریج مورد توجه کتابخانه‌ها و موزه‌ها برای توصیف منابع‌شان قرار گرفته است. دابلین کور طوری طراحی شده است که غیرمتخصصان فهرست‌نویسی هم بتوانند آن را به کار گیرند، به گونه‌ای که نویسندگان و مسئولین وب سایت‌ها بدون داشتن تخصص فهرست‌نویسی قادر به استفاده از دابلین کور برای

توصیف منابع‌شان هستند و بدین ترتیب منابع‌شان را برای موتورهای کاوش و سیستم‌های بازیابی قابل رؤیت‌تر می‌سازند [۴].

اونیکس نیز به استاندارد اطلاعاتی 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. اونیکس به عنوان راه‌حلی برای دو مشکل ایجاد شد: الف: نیاز به اطلاعات کامل‌تر کتاب‌ها به صورت آنلاین؛ و ب: وجود تنوع فراوان فرمت‌های مورد نیاز 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 اونیکس (ONIX 1.0) در سال ۲۰۰۰ انتشار یافت [۶].

ایجاد UK LOM از مقاله‌ای که در آوریل ۲۰۰۲ به گروه ابرداده‌های آموزشی بریتانیا ارائه شد، سرچشمه می‌گیرد. این مقاله فراخوانی برای تشکیل انجمنی از دست‌اندرکاران برای شناسایی روش‌های شایع در بریتانیا جهت استفاده از ابرداده‌ها در محتوای بسته‌های آموزش الکترونیکی بود. هدف اصلی استاندارد LOM افزایش قابلیت همکاری نمونه‌های ابرداده‌ها و پروفایل‌های نرم‌افزارهای ایجاد شده در بخش آموزش انگلستان است. براساس تعریف مؤسسه IEEE (۱۹۹۰) می‌توان قابلیت همکاری متقابل را در چنین بستری به صورت زیر تعریف کرد: "توانایی دو یا چند سیستم یا مؤلفه‌های آنها برای تبادل اطلاعات و استفاده از اطلاعاتی که رد و بدل شده‌اند" [۷].

در پیشینه‌های مورد بررسی، طاهری (۱۳۸۷) ضمن معرفی و بررسی نقاط قوت و ضعف طرح فراداده‌ای هسته دابلین و قالب مارک ۲۱، در سازماندهی منابع اطلاعاتی شبکه جهانی وب، با استفاده از روش تحلیلی، کارایی دو طرح مذکور را در نمایه‌سازی اطلاعات الکترونیکی با هم مقایسه کرده است. نتیجه مقایسه نشان داد که در نهایت، قالب مارک ۲۱ برای ذخیره، پردازش و مبادله اطلاعات محیط وب، مناسب‌تر است. طرح فراداده‌ای هسته دابلین، نه تنها از لحاظ تعداد عناصر توصیف، نسبت به مارک ۲۱ قابلیت‌های کمتری دارد، بلکه از لحاظ انواع قالب برای انواع اطلاعات و نیز انواع قالب برای انواع مواد اطلاعاتی، با قالب مارک ۲۱، قابل مقایسه نیست [۸].

در مطالعه‌ای دیگر نیز شفیعی و همکاران (۱۳۸۸) در پژوهش خود، ۹۰ صفحه وب از طریق جستجو در موتورهای جستجوی عمومی

(گوگل، یاهو و ام اس ان) را به عنوان نمونه انتخاب کردند و میزان حضور عناصر ابر داده‌ای (هسته دابلین 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 هسته دابلین و ابر نشانه‌های زبان نشانه‌گذاری فرامتن در صفحات بازیابی شده بود، که بیانگر کاربرد بسیار کم عناصر ابر داده‌ای در صفحات وب بود [۹].

همچنین کریمی و فیاض (۱۳۸۹) در مطالعه‌ای ضمن معرفی مهمترین استانداردهای سازماندهی در کتابخانه‌های دیجیتال، به این نتیجه رسیدند، قالب فراداده‌ای هسته دویلین (Dublin core) با ساختاری ساده و قابل استفاده حتی برای پدیدآور اثر و برای طیف وسیعی از منابع شبکه‌ای به وجود آمده است. استاندارد مارک دارای ساختاری پیچیده‌تر، جزئی‌تر و کاربردی‌تر، نسبت به طرح فراداده دویلین است. طرح فراداده‌ای توصیف شی (MODS: Metadata Object Description Schema) با ساختار ذخیره‌سازی مبتنی بر XML (Extensible Markup Language) و منطبق با مارک، دارای مجموعه عناصری بیشتر از دویلین کور و ساده‌تر از مارک 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 مکان یاب اطلاعات دولتی (Government Information Locator Services) برای مکان‌یابی اطلاعات مناسب نهادها و ادارات دولتی کاربرد دارد. فراداده‌ای توصیف مستند (MADS: Metadata Authority Description 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 فهرست نویسی انگلو آمریکن است؛ به عنوان ابزاری، برای فهرست‌نویسی منابع دیجیتالی و بهترین ابزار فهرست‌نویسی که مورد مقبولیت بین‌المللی قرار گرفته است، محسوب می‌شود [۳].

در مطالعه‌ای دیگر، ابازر و بابایی (۱۳۹۰) 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 اینترنت اکسپلورر و نت اسکپ مورد بررسی قرار دادند و با مشاهده Source از طریق منوی View عناصر فراداده‌ای موجود در صفحه HTML هر یک از وب سایت‌ها را استخراج کردند و در چک‌لیست قرار دادند. در وب‌سایت‌های مورد بررسی آنها، هیچ یک از وب‌سایت‌ها از عناصر فراداده‌ای هسته دویلین 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 دویلین کور [۱۰].

شفیعی و همکاران (۱۳۹۱) از روش تحلیل محتوا، عناصر ابر داده‌ای دابلین کور و مجموعه ابر نشانه‌های زبان نشانه‌گذاری فرامتن موجود در فیلدهای موجود در صفحات جست‌وجوی پایگاه‌های اطلاعاتی پزشکی Medline, MD Consult, Medline Ovid, Pub Med, Iranmedex, Pars Medline و ایران و جهان، را بررسی کردند. یافته‌ها نشان داد، عناصر ابر داده‌ای دابلین کور و ابر نشانه‌های زبان نشانه‌گذاری فرامتن در صفحات بازیابی 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 دابلین کور وجود داشت. بیشترین تعداد فیلد ابر داده‌ای مربوط به پایگاه اطلاعاتی Pub Med 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 پیشنهادها مورد بررسی در خارج کشور؛ FastCampbell (۲۰۰۴) پیشنهادی 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].

در مطالعه دیگر، علیجانی و جوکار (۲۰۰۹) با هدف بررسی تمایل کتابداران کتابخانه‌های ملی جهان و میزان توجه آنان به استاندارد دویلین کور، با استفاده از چک‌لیست و مراجعه به منبع صفحه اصلی وب سایت هر کتابخانه به بررسی عناصر فراداده دویلین کور در وب سایت‌های ۷۰ کتابخانه ملی پرداختند. عناصر دویلین کور 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 دوبلین کور استفاده کرده‌اند. کتابخانه ملی عربستان با استفاده از هر ۱۵ عنصر دوبلین کور، حداکثر عناصر را و کتابخانه ملی سوئیس با استفاده از ۲ عنصر، حداقل عناصر دوبلین کور را استفاده کرده‌اند [۱۳].

Kurtz (۲۰۱۰) در مطالعه‌ای ضمن بررسی اجمالی دوبلین کور (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 دوبلین کور و دی‌اسپیس استفاده می‌کردند. سپس از هر مخزن بیست رکورد به طور تصادفی نمونه‌گیری شد و با استفاده از ملاک‌های کیفیت ابرداها 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]. همچنین فرج‌پهلوی و طباطبائی (۲۰۱۱) پژوهشی با هدف تبیین نمایه‌سازی و رتبه‌بندی اشیای محتوایی حاوی عناصر فراداده‌ای هسته دوبلین و مارک ۲۱ مبتنی بر زبان نشانه‌گذاری گسترش‌پذیر (XML) در محیط‌های اطلاعاتی پیوسته پویا توسط موتورهای کاوش گوگل، یاهو و مقایسه آنها با یکدیگر با رویکرد تحلیلی - مقایسه‌ای انجام دادند. جامعه پژوهش را ۱۰۰ شی محتوایی (کتاب الکترونیکی) مبتنی بر زبان نشانه‌گذاری گسترش‌پذیر (XML) در قالب دو گروه تشکیل داد؛ گروه اول، حاوی عناصر فراداده‌ای هسته دوبلین و گروه دوم، حاوی عناصر فراداده‌ای مارک ۲۱ بود که بر روی وب‌سایت <http://www.marcdcmi.ir> انتشار یافته بودند. ابزار گردآوری داده‌ها سبانه واری بود. در مجموع، یافته‌ها حاکی از آن بود، که هیچ یک از دو طرح فراداده‌ای هسته دوبلین و مارک ۲۱ مبتنی بر XML منتشرشده در محیط‌های اطلاعاتی پیوسته پویا از لحاظ دسترس‌پذیری توسط موتورهای کاوش نسبت به دیگری برتری ندارند [۱۵].

Debus و همکاران (۲۰۱۲) استفاده از ابرداها ONIX را در کتابخانه کنگره مورد بررسی قرار دادند. آنان معتقدند که استاندارد

اونیکس برای کتابداران نویدبخش است، چرا که بسیاری از عناصر داده‌ها در داخل آن را، می‌توان به صورت استاندارد MARC نگاشت. کتابخانه کنگره برای ایجاد توصیف‌های کتاب‌شناسی مارک از ابرداها‌های اونیکس ناشرین برای انتشارات جدید دریافتی از طریق فهرست‌نویسی الکترونیکی در برنامه انتشارات کتاب، تبدیل‌کننده اونیکس به مارک را ایجاد کرد. نتایج نشان داد که استفاده از مبدل اونیکس به مارک، زمان لازم برای ایجاد توصیف‌های کتاب‌شناسی را کاهش می‌دهد، باعث تسهیل گنجاندن ابرداها در پیشینه‌های کتاب‌شناسی می‌شود و برای کارکنان فهرست‌نویسی پیشینه‌هایی را فراهم می‌سازد که قابل مقایسه با پیشینه‌هایی با کیفیت فهرست‌نویسی غیربنیادی است [۱۶].

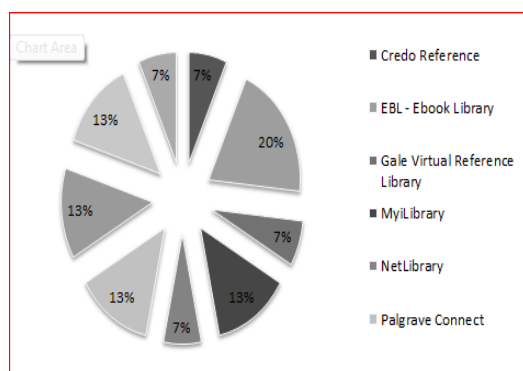
Baker (۲۰۱۲) در مطالعه خود مطرح ساخت که چگونه مسائل پیرامون ابرداها دوبلین کور، نمونه بارز مشکلاتی است که نیازمند حل و فصل است تا داده‌های کتابخانه‌ها با استانداردهای سستی کیفیت و سازگاری مطابقت داشته باشند و در عین حال در محیط داده‌های مرتبط با منابع آنها سازگار باشند. او در مقاله‌اش نشان داد که چگونه استاندارد دوبلین کور که عموماً به عنوان یک قالب ساده پیشینه در نظر گرفته می‌شود، به طور روزافزونی به عنوان واژگان RDF برای استفاده در ابرداها و بر اساس یک مدل "شرح" دیده می‌شود و چگونه برای پرکردن شکاف بین این مدل‌ها روش‌های جدید ابرداها 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 ابرداها در سازماندهی، نظم‌بخشی، بازیابی دقیق منابع اطلاعاتی الکترونیکی [۱] و هزینه‌های خرید آنها از جمله مسائل مهمی به شمار می‌آیند که اهمیت بررسی ابرداها را نشان می‌دهد، حال این مسئله پیش می‌آید که پایگاه‌های کتاب الکترونیکی از نظر به کارگیری ابرداها در چه وضعیتی قرار دارند؟ لذا هدف از این مطالعه تعیین ابرداها‌های پایگاه‌های کتاب الکترونیکی است.

روش کار:

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 شده‌اند. اطلاعات مربوط به ابرداها‌های مورد استفاده در وب سایت "ابزار ارزیابی پایگاه‌های علمی" (<http://adat.cri.edu>) که به مقایسه پایگاه‌ها از لحاظ مختلف می‌پردازد، قابل نمایش بود. در این وب سایت با انتخاب پایگاه‌ها می‌توان

پایگاهها به کار برده شده بود. ابر دادههای مورد استفاده در پایگاههای موردنظر به ترتیب مارک به صورت (۱۰۰ درصد)، اونیکس (۵۵/۵۵ درصد)، دویلین کور (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 موردنظر، 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 بردهاند (جدول شماره ۲).



نمودار شماره ۱- نمودار دایره‌ای توزیع فراوانی کاربرد ابر دادهها در پایگاههای کتاب الکترونیکی

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 تجزیه و تحلیل قرار گرفته‌اند. آدرس پایگاههای اطلاعاتی کتاب الکترونیک در جدول شماره یک نشان داده شده است (جدول ۱).

یافته‌ها:

پس از مرحله گردآوری دادهها، توزیع فراوانی اطلاعات جمع‌آوری شده پایگاهها، به صورت نمودار شماره ۱ و جدول شماره ۲ نشان داده شده است.

جدول شماره ۲، نشانگر توزیع فراوانی پایگاهها بر اساس ابر دادههای مورد استفاده است. بنابراین جدول شماره ۲، حاکی از آن است که ابر داده مارک در تمامی پایگاهها، ولی ابر داده دویلین کور تنها در یک پایگاه به نام EBL - Ebook Library و ابر داده اونیکس در ۵ پایگاه به کار گرفته شده است و ابر داده UK LOM در هیچ کدام از پایگاهها استفاده نشده است. به طور کلی، بیشترین میزان استفاده از ابر دادهها، مارک و سپس ابر داده اونیکس بود. در مجموع ۶۱/۶۶ درصد از ابر دادهها در

جدول شماره ۱- آدرس و نام پایگاههای اطلاعاتی کتاب الکترونیک

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

جدول شماره ۲- توزیع فراوانی ابر داده‌های مورد بررسی در پایگاه‌های کتاب الکترونیکی

Row	Site Name	MARC	ONIX	UK LOM	Dublin	جمع فراوانی	درصد فراوانی
1	Credo Reference	✓	—	—	—	۱	۲۵
2	EBL - Ebook Library	✓	✓	—	✓	۳	۷۵
3	Gale Virtual Reference Library	✓	—	—	—	۱	۲۵
4	MylLibrary	✓	✓	—	—	۲	۵۰
5	NetLibrary	✓	—	—	—	۱	۲۵
6	Palgrave Connect	✓	✓	—	—	۲	۵۰
7	Taylor & Francis eBookstore	✓	✓	—	—	۲	۵۰
8	Dawsonera	✓	✓	—	—	۲	۵۰
9	Ebrary	✓	—	—	—	۱	۲۵
	جمع فراوانی	۹	۵	۰	۱	۱۵	۴۱/۶۶
	درصد فراوانی	۱۰۰	۵۵/۵۵	۰	۱۱/۱۱	—	—

بحث و نتیجه‌گیری:

در یک نگاه کلی، ابر داده‌های مورد استفاده در پایگاه‌های مورد نظر به ترتیب مارک، اونیکس، دوبلین کور بود و UK LOM در هیچ کدام از پایگاه‌ها مورد استفاده قرار نگرفت. به طور کلی بیشترین میزان استفاده به ابر داده‌ها در پایگاه‌های مورد نظر، MARC و کمترین میزان استفاده 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 برده‌اند. به کارگیری استاندارد MARC امکان تبادل اطلاعات را تقویت می‌کند [۵].

انجمن ناشران آمریکایی نیز به منظور تبادل اطلاعات کتاب‌ها، اقدام به ایجاد ابر داده اونیکس کردند، در این مطالعه نیز بیشترین استفاده از ابر داده ابتدا MARC و سپس ابر داده اونیکس (۵۵/۵۵ درصد) بود. Debus و همکاران (۲۰۱۲) اشاره کردند، استفاده از مبدل اونیکس به مارک، زمان لازم برای ایجاد توصیف‌های کتاب‌شناسی را کاهش می‌دهد، همچنین باعث تسهیل گنجاندن ابر داده‌ها در پیشینه‌های کتاب‌شناسی می‌شود [۱۶].

در مطالعه شفیع و همکاران (۱۳۸۸) همچون مطالعه حاضر تنها یک پایگاه از ابر داده دوبلین کور استفاده کرده بود [۱۱]، در مطالعه اباد و بابایی (۱۳۹۰) هیچ کدام از وب سایت‌ها از ابر داده دوبلین کور استفاده نکرده‌اند [۱۰]. در مطالعه دیگر شفیع و همکاران (۱۳۹۱)، نشان از حضور بسیار کم عناصر ابر داده دوبلین کور بود [۹]، در مطالعه Baker (۲۰۱۲)، اشاره شد که چگونه در استاندارد دوبلین کور که عموماً به

عنوان 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].

در مطالعه Kurtz (۲۰۱۰)، دانشگاه‌های مورد بررسی برای ایجاد و مدیریت مخازن سازمانی خود از دوبلین کور و دی‌اسپیس استفاده می‌کردند [۱۴]. علیجانی و جوکار (۲۰۰۹)، در مطالعه خود نشان دادند که فقط ۱۴ (۲۰ درصد) کتابخانه ملی از عناصر دوبلین کور استفاده کرده‌اند و بسیاری از کتابخانه‌های ملی که کتابخانه‌های اصلی جهان هستند، در طراحی صفحات وب خود، توجه زیادی به عناصر دوبلین کور معطوف نداشته‌اند، همچنین در مطالعه حاضر، ۱۱/۱۱ از پایگاه‌ها، از ابر داده دوبلین کور استفاده کردند [۱۳].

فرچ‌پهلو و طباطبائی (۲۰۱۱) نیز اشاره کردند که هیچ یک از دو طرح فراداده‌ای دوبلین کور و مارک ۲۱ 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 (۲۰۰۴)، 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]. در نهایت همان طور که کریمی و فیاض (۱۳۸۹) اشاره کردند با توجه به اینکه استفاده از تمامی استانداردهای

درصد استفاده شده بود. در واقع کاربرد ابر داده‌ها باعث افزایش دقت بازیابی و تسهیل جستجو در منابع الکترونیکی می‌شوند.

از جمله محدودیت‌های این پژوهش، عدم بررسی مستقیم ابر داده‌های پایگاه‌ها؛ به دلیل غیر رایگان بودن آنها بود، به همین دلیل از ابزار ارزیابی پایگاه‌های علمی (ADAT) که به مدیریت منابع الکترونیکی می‌پردازد، استفاده شده است.

با رشد تصاعدی منابع الکترونیکی در محیط وب، لزوم بازیابی سریع اطلاعات، صرفه‌جویی در وقت و هزینه خرید این دسته از منابع الکترونیکی، جهت بازیابی دقیق اطلاعات، بهتر است منابعی که شامل ابر داده‌های مناسب هستند توسط متخصصان اطلاع‌رسانی خریداری شوند. طبق یافته‌ها و نتایج حاصل از این مطالعه، لازم است در طراحی پایگاه‌ها از ابر داده مناسب استفاده شود، تا کاربران در بازیابی اطلاعات به نتایج مطلوبی دست یابند. به طور کلی ابر داده مارک در وب سایت‌ها بیشترین میزان کاربرد را در پایگاه‌ها داشته است، بعد از مارک بیشترین میزان کاربرد ابر داده، به ONIX و سپس Dublin Core تعلق می‌گیرد و ابر داده UK LOM در هیچ کدام از پایگاه‌ها مورد توجه قرار نگرفته بود. براساس نتایج اخذ شده پیشنهاد می‌گردد متخصصان اطلاع‌رسانی و طراحان پایگاه‌ها در ایجاد پایگاه‌های الکترونیکی، ابر داده‌های مناسب را در نظر بگیرند. همچنین پیشنهاد می‌شود، مطالعه مشابهی در مورد شناسایی انواع عناصر ابر داده‌ای مورد استفاده در ابر داده‌های مورد نظر انجام گیرد و با توجه به اهمیتی که عناصر ابر داده‌ای در ابر داده‌های مورد نظر، برای بازیابی دقیق‌تر اطلاعات دارند، توصیه‌هایی به متخصصان اطلاع‌رسانی و طراحان پایگاه‌ها جهت به کارگیری آنها می‌شود.

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 (۱۳۸۷) در مطالعه خود قالب مارک ۲۱ را برای ذخیره و پردازش اطلاعات محیط وب مناسب‌تر دانست [۸]. در مطالعه فیاض و کریمی (۱۳۸۹) ابر داده مارک نسبت به ابر داده دوبلین کور، پیچیده‌تر، کاربردی‌تر و جزئی‌تر شناخته شده، در مطالعه حاضر نیز ابر داده مارک که کاربردی‌تر بود، بیشتر استفاده شد [۳].

در حال حاضر، قالب مارک ۲۱، قالب‌های مختلفی برای انواع اطلاعات ارائه می‌دهد که عبارتند از: اطلاعات کتاب‌شناختی، قالب مستندات، اطلاعات رده‌بندی، اطلاعات جامعه و اطلاعات مربوط به موجودی. همچنین، از آنجا که مارک ۲۱، بر اساس قواعد فهرست‌نویسی انگلو امریکن طراحی شده، قالب‌های خاص هر نوع رسانه‌ای را از چاپی تا مواد سمعی و بصری و منابع الکترونیکی، شامل می‌شود؛ در صورتی که هسته دوبلین، فقط منحصر به منابع الکترونیکی است و سایر رسانه‌های اطلاعاتی را پشتیبانی نمی‌کند و برای اطلاعات خاص فاقد قالب خاص نیز است. بنابراین قالب فراداده‌ای مارک ۲۱ برای ذخیره، پردازش، تبادل و بازیابی منابع شبکه جهانی وب، مناسب‌تر و کارآمدتر است. تعدد عناصر کتاب‌شناختی، وجود قالب‌های مختلف اطلاعات، سابقه دیرینه تولید و کاربرد قالب مارک، همراه با تعدد ویرایش‌ها، تهیه‌کنندگان پیشینه‌ها، گستره کاربردها، اجماع بین‌المللی و امکان کنترل نام‌ها و موضوع‌ها، از مهمترین دلایل برتری مارک ۲۱ نسبت به «هسته دوبلین» است [۱۱]. در این مطالعه نیز از ابر داده مارک به صورت ۱۰۰

References

منابع

1. Mohamadi Sh. Metadata: Concepts and applications. *Scientific Communication, Monthly Journal of Irandoc*. 2004; 2(3): 48-53. [Persian]
2. 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. [Persian]
3. Karimi M, Fayaz S. Standards of information organization in digital libraries. *Scientific Communication, Monthly Journal of Irandoc*. 2010; 16(2):1-10. [Persian]
4. 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.
5. 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 : <http://www.Loc.Gov/Marc/Umb/>.(accessed 29 April 2013)
6. Daly F. ONIX: The metadata standard for the information and entertainment industries. *Publishing Research Quarterly*. 2002; 18(2): 28-40.
7. Campbell L. Learning Object Metadata (LOM). In: Ross S, Day M, Editors. DCC Digital Curation Manual; 2007; Available from: <http://www.Era.Lib.Ed.Ac.Uk/Handle/1842/3334>.(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. [Persian]
9. ShafieAlavijeh S, Ghaebi A, RezaieSharifabadi S. Review of metadata elements within the web pages resulting from searching in general search engines. *Journal of Information Proccessing and Management*. 2009; 25(1): 71-89. [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. *J Health Adm*. 2011; 14(43): 7-18. [Persian]
11. ShafieAlavijeh S, Rahimi A, ShafieAlavijeh P. The Comparison of metadata elements and hypertext markup language in records retrieved from Iranian specialized medical databases with those of other countries. *National Studies on Librarianship and Information Organization*. 2012; 23(2): 160-76. [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.[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, et al. 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

E. Ommati, MSc¹ M. Tavassoli Farahi, PhD²

1. MSc, Medical Library and Information Science, Tehran University of Medical Sciences, Tehran, Iran.

2. Assistant Professor, 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)

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

Correspondence:

M. Tavassoli Farahi, PhD. Research Center for Health Information Management, Hormozgan University of Medical Sciences, Bandar Abbas, Iran

Tel: +98 7616666367

Email:

mfarahi20@gmail.com