اطلاع‌شناسی
INFORMOLOGY

سازماندهی اشیاء شبه مدرک در شبکهٔ جهانی وب

مهدی صفری

چکیده

شبکهٔ جهانی وب، با تمام امکاناتی که در ذخیره‌سازی، پردازش و خصوصاً انتقال اطلاعات به ارمغان آورده، متصدیان سازماندهی اطلاعات را با چالش‌های زیادی روبه‌رو ساخته است. سرعت سرسام آور منابع اطلاعاتی، ویژگی‌های ماهوی آنها و خصیصه های منحصر به فرد این رسانهٔ نوظهور موجب گردیده که روش‌های سنتی توصیف و سازماندهی منابع در این محیط کارآیی خود را از دست بدهند. بنابراین، آنچه توجه اکثر متخصصان را در امر سازماندهی منابع وب به خود جلب کرده است ابزاری است که امروزه "ابرداده" نامیده می شود. خلق قالب‌های ابرداده‌ای گوناگون با مراتب مختلف غنا در سطح توصیف از ویژگی‌های این عصر است که در این میان قالب ابرداده‌ای دابلین کور به‌عنوان نوعی قالب ابرداده‌ای ساده، و به‌دلیل ویژگی‌های درونی‌اش، اهمیت خاصی یافته است.

این مقاله، با ارائهٔ تصویری از ابرداده، به ضرورت‌ها و شرایط پدید آمدن قالب ابرداده‌ای دابلین کور می‌پردازد. ویژگی‌ها، رسالت، وظایف، و کارکردهای این قالب ابرداده‌ای در شبکهٔ جهانی وب بیان می‌شود و بستر نحوی به‌کارگیری عناصر ابرداده‌ای تعریف‌شده در این استاندارد مورد اشاره قرار می‌گیرد.

کلیدواژه‌ها: وب. سازماندهی. اشیاء شبه‌مدرک. ابرداده. دابلین کور

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

سازماندهی اشیاء شبه مدرک در شبکهٔ جهانی وب:

مقدمه‌ای بر قالب ابرداده‌ای دابلین کور

مهدی صفری[1]

مقدمه

شبکهٔ جهانی وب امکان تولید و اشاعهٔ منابع اطلاعاتی را توسط میلیون‌ها استفاده‌کننده در سرتاسر دنیا فراهم آورده است. رشد سریع اطلاعات در این شبکه و پویایی این رسانه، چه از لحاظ اندازه و چه محتوای منابع اطلاعاتی، سبب شده است که توصیف، سازماندهی و در نتیجه بازیابی منبع ــ یا به زبان وب، کشف منبع[2] ــ به بزرگ‌ترین مسئلهٔ این شبکه تبدیل شود. موتورهای کاوش به‌عنوان مهم‌ترین ابزارهای کاوش و کشف منابع وب، گرچه شمار زیادی منبع را در هر کاوش در اختیار قرار می‌دهند، ناکارآمدی آنها در بازیابی منابع وب از وجوه گوناگون در متون حوزه‌های مختلف مشخص شده است؛ از آن جمله می‌توان مطالعهٔ پترسون[3] (۱۹۹۷)، بار-ایلان[4] (۱۹۹۹)، روسو[5] (۱۹۹۹)، گردون و پاتاک[6] (۱۹۹۹) و مطالعهٔ متروپ و نیوون هوسن[7] (۲۰۰۱) را تنها به‌عنوان نمونه‌هایی ذکر کرد که هریک به نحوی ناکارآمدی موتورهای کاوش را در بازیابی منابع وب بررسی کرده‌اند. علاوه بر مسئلهٔ نقصان این ابزارها در بازیابی منابع، عدم توفیق آنها در تحت پوشش قراردادن همهٔ منابع وب از دیگر مسائل قابل توجه است. بر اساس مطالعهٔ لورنس و گیلز[8] (۱۹۹۹)، بزرگ‌ترین موتورهای

۱. کارشناس ارشد بنیاد دایرةالمعارف اسلامی
Email: mehsafari@yahoo.com

2. Resource discovery
3. Peterson
4. Bar-Ilan
5. Rousseau
6. Gordon & Pathak
7. Mettrop & Nieuwenhuysen
8. Lawrence & Giles

کاوش تنها ۱۶ درصد از منابع وب را نمایه کرده‌اند. این امر نشان می‌دهد که میزان معتنابهی از منابع وب ازطریق موتورهای کاوش نمایه و بالطبع بازیابی نمی‌شوند.

نمایه‌سازی تمام متن[1] صفحات وب به‌وسیلهٔ موتورهای کاوش سبب بازیابی مدارک نامرتبط زیادی می‌گردد. این روش در نمایه‌سازی صفحات وب، با توجه به حجم خارق‌العادهٔ منابع وب و سرعت رشد آن، طبعاً نمی‌تواند پاسخگوی نیازهای استفاده‌کنندگان باشد. به همین دلیل، برای مدیریت [کارآمد] اطلاعات باید توجه داشت که محتوای مدرک همیشه حاوی اطلاعات یا بافت مناسب نیست تا آن مدرک به نحو مؤثر نمایه‌سازی شود" (یانلا[2]، ۱۹۹۹). بنابراین، نیاز به اضافه‌کردن داده‌های توصیفی ورای آنچه در محتوای مدارک وب موجود است ضروری به‌نظر می‌رسد.

توصیف منابع فیزیکی اطلاعات در کتابخانه‌ها، به منظور تسهیل بازیابی آنها از دیرباز با استفاده از روش‌هایی چون نمایه سازی و خصوصاً فهرستنویسی انجام می‌شده است. اما، خصیصه های اصلی منابع الکترونیکی و شبکه‌ای و تفاوت آنها با منابع فیزیکی (هیری[3]، ۱۹۹۶) موجب شده است که به‌کارگیری روش‌های سنتی توصیف و کشف منبع، مانند قوانین فهرستنویسی، در شبکهٔ وب با دشواری روبه‌رو گردد. این دشواری‌ها و مشکلات در به‌کارگیری قوانین فهرستنویسی در محیط‌های الکترونیکی (بیکام[4]، ۲۰۰۰؛ لاگوتس[5]، ۲۰۰۰؛ ویس و کاستنز[6]، ۲۰۰۱) سبب گردیده که نقش آنها در محیط نوظهور وب کمرنگ گردد و به موازات آن ابزاری جدید که امروزه تحت عنوان "ابرداده[7]" شناخته می‌شود در توصیف و کشف منابع وب روز به روز از اهمیت بیشتری برخوردار گردد. افزایش شتابناک منابع اطلاعاتی در موضوع ابرداده و کاهش واژهٔ فهرستنویسی در متون مربوط (ارسگواک[8]، ۱۹۹۹) خود گواهی بر این مدعاست.

ابرداده چیست؟

این مفهوم نخستین‌بار با نام شرکتی امریکایی برای ایجاد و توسعهٔ محصولات مبتنی بر "ابرالگوها"[9] و توصیف مجموعهٔ داده‌ها در سال ۱۹۶۹ به‌کار رفت؛ اما ابتدا با چنین هدف گسترده‌ای مورد استفاده قرار نگرفت. کاربرد مکرر این اصطلاح بیشتر به دههٔ ۱۹۹۰ و شکل‌گیری شبکهٔ جهانی وب در ۱۹۹۳ باز می گردد (شیری، ۱۳۸۱). ساده‌ترین تعریفی که می‌توان از ابرداده ارائه داد، داده دربارهٔ داده است. اما به‌طور دقیق‌تر

1. Full text indexing
2. Iannella
3. Heery
4. Beacom
5. Lagoze
6. Weiss & Carstens
7. Metadata
8. Ercegovac
9. Metamodels

می‌توان ابرداده را دادهٔ ساخت‌یافته[1] دربارهٔ دیگر داده‌ها دانست (دی[2]، ۱۹۹۹؛ بورنت، ان جی و پارک[3]، ۱۹۹۹). به‌طور خاص‌تر، ابرداده مجموعهٔ ساخت‌یافته‌ای از عناصر است که منابع اطلاعاتی را به‌منظور شناسایی، کشف و استفاده از اطلاعات توصیف می‌کند (لی ـ اسملتسر[4]، ۲۰۰۰). گرینبرگ[5] (۲۰۰۲) تعریف دیگری از ابرداده ارائه می‌دهد که عبارت است از: "داده‌های ساخت‌یافته دربارهٔ شیء اطلاعاتی که عملکردهای مرتبط با شیء اطلاعاتی ویژه‌ای را پشتیبانی می‌کند".

آنچه از تعاریف بالا بر می‌آید این است که ابرداده داده‌ای است که منابع اطلاعاتی دیگر را توصیف کرده، بازیابی و استفاده از آنها را تسهیل می‌کند. طبق تعاریف ذکر شده، فهرست های کتابخانه‌ای و اصطلاحات نمایه‌ای ابرداده به شمار می آیند چون داده‌های ساخت‌یافته دربارهٔ دیگر داده‌ها (منابع کتابخانه‌ای) هستند و بازیابی و استفاده از آن منابع را تسهیل می‌کنند. اما، همان‌طور که دِی (۱۹۹۹) نیز خاطرنشان می‌کند، امروزه این اصطلاح به داده‌هایی اطلاق می‌شود که طوری ساخت یافته‌اند که بتوانند ماشین‌خوان و ماشین‌فهم باشند و غالباً در مورد مسائل کشف منبع در محیط وب به‌کار می‌روند.

طرح ابرداده‌ای دابلین کور: ضرورت‌ها و نیازها

طرح ابرداده‌ای دابلین کور[6] طرحی بین‌المللی و بین‌رشته‌ای[7] است که یک مجموعهٔ هسته از پانزده عنصر ابرداده‌ای را برای توصیف منابع ارائه می‌دهد (دی، ۱۹۹۹). خلق این قالب ابرداده‌ای مانند دیگر قالب‌های ابرداده‌ای برخاسته از نیازها و اهدافی بوده که پدیدآورندگان و توسعه‌دهندگان آن در نظر داشته‌اند.

در زمانی که عناصر ابرداده‌ای دابلین کور پدید آمد، قالب‌های ابرداده‌ای گوناگونی وجود داشت که هرکدام بنا بر نیازهای خاصی خلق شده بودند و از مراتب غنای مختلفی در سطح توصیف نیز برخوردار بودند. این طیف وسیع غنا در سطح توصیف، در دوران تحول و تطور دابلین کور باعث اختلافاتی میان گروه‌های مختلف شده است که ناشی از تجربیات آنها در دنیای بازیابی اطلاعات است. در عرصهٔ بازیابی اطلاعات، نگرش‌های متفاوتی در سطح و شیوهٔ توصیف منابع وجود داشت که می‌توان آنها را به‌صورت زیر خلاصه کرد (وایبل[8] یانلا و کاترو[9]، ۱۹۹۷):

۱. نمایه‌سازی تمام متن: نمایه‌سازی غیر ساخت یافته منابع تمام متن از طریق فایل مقلوب[10]: روش مورد استفاده در موتورهای کاوش.

1. Structured data
2. Day
3. Burnett, Ng & Park
4. Lee-Smeltzer
5. Greenberg
6. Dublin Core Metadata Initiative
7. Interdisciplinary
8. Weibel
9. Cathro
10. Inverted file

۲. جایگزین‌های حوزه‌بندی نشده[1]: مجموعهٔ اصطلاحاتی که توسط مؤلف یا نمایه‌ساز به‌عنوان ابرداده به یک منبع داده می‌شود بدون تمایز و تفاوت گذاشتن میان آنها (مثلاً امکان تفاوت گذاشتن میان منابعی که در مورد فرد "الف" است با منابعی که توسط فرد "الف" خلق شده است).

۳. جایگزین‌هایی با حداقل حوزه‌بندی[2]: وجود شمار محدودی از حوزه‌ها همراه با معنای شناخته‌شده (مانند پدیدآورنده، موضوع و عنوان). عناصر شناخته‌شده، کاوش حوزه‌بندی‌شده را پشتیبانی می‌کنند.

۴. جایگزین‌های تعدیل‌شده[3]: نام حوزه‌ها دارای نمادهای اضافی است که این امر پالایش و تعدیل این عناصر و محتوای آنها را پشتیبانی می‌کند.

۵. جایگزین‌های بسیار ساخت‌یافته[4]: داده‌هایی با سطح بالای ساخت‌یافتگی؛ به‌طور مثال، رکوردهای مارک و طرح کدگذاری متن[5].

هر کدام از موارد بالا در طیف توصیف منابع، نشان‌دهندهٔ نوعی توافق در مورد هزینه، سادگی خلق و نگهداری، و همچنین کاربرد و استفاده از رکوردهای توصیفی است. در یک سوی این طیف مزیت سادگی خلق رکوردهای توصیفی همراه با جامعیت نسبتاً بالا ولی مانعیت بسیار پایین وجود دارد؛ و در سوی دیگر، رکوردهایی وجود دارند که خلق و نگهداری آنها (چه از نظر مالی و چه از لحاظ نیروی انسانی) پرهزینه ولی دارای سطح مانعیت و انسجام بیشتری هستند.

زمانی که فکر توسعهٔ قالب ابرداده‌ای دابلین کور شکل گرفت، دو راهکار که هرکدام نشانگر دو انتهای طیف مذکور در توصیف منابع هستند از استفادهٔ زیادی برخوردار بودند. نمایه‌های خودکار مانند نمایه‌هایی که توسط خدمات مکان‌یابی تولید می‌شدند ازقبیل Lycos و Webcrawler و رکوردهای مارک که توسط متخصصان فهرستنویسی تولید می‌شدند. از طرفی رکوردهای خودکار غالباً اطلاعات توصیفی اندکی را در خود داشتند تا مفید واقع شوند و از طرف دیگر خلق و نگهداری رکوردهای دستی مانند مارک علاوه بر پرخرج بودن ، برای منابع موجود در اینترنت بسیار دشوار می نمود. راه حل این مسئله، خلق رکوردهای توصیفی بود که مفصل‌تر از مدخل‌های نمایه‌ای و خلاصه‌تر از یک رکورد فهرستنویسی باشد. اگر نیروی انسانی کمی برای توصیف منابع نیاز باشد و مهم‌تر از آن، اگر این کار توسط پدیدآورندگان منابع انجام گیرد منابع زیادی می‌توانند توصیف شوند؛ هم‌چنین اگر توصیف بر اساس استاندارد معیّنی صورت گیرد فقط خلق

1. Unfielded surrogates
2. Minimally fielded surrogates
3. Qualified surrogates
4. Richly-structured surrogates
5. Text Encoding Initiative (TEI)

رکوردها نیاز به دخالت انسان دارد و ابزارهای خودکار، این عناصر توصیفی را یافته و آنها را در پایگاه اطلاعاتی قابل جست‌وجو جمع‌آوری می‌کنند. پس در جواب به این پرسش که چنین قالب ابرداده‌ای ساده‌ای چگونه می‌تواند تعریف شود که بتواند گسترهٔ وسیعی از منابع الکترونیکی را به نحو مناسب توصیف کند و همچنین حد واسطی میان دو انتهای طیف یادشده در سطح توصیف منابع باشد، مرکز کتابخانهٔ رایانه‌ای پیوسته (اُ. سی. ال. سی)[1] و مرکز ملی کاربردهای ابررایانه‌ای[2] را بر آن داشت که کارگاهی آموزشی با موضوع ابرداده در مارس ۱۹۹۵ در شهر دابلین ایالت اوهایو امریکا برگزار کنند. ۵۲ کتابدار، آرشیویست، پژوهشگر علوم انسانی، جغرافی‌دان، تهیه‌کنندگان استانداردهای اینترنت، حوزه‌های مربوط به استاندارد Z39.50، و زبان استاندارد نشانه‌گذاری عمومی[3] در این کارگاه آموزشی شرکت کردند تا به بحث و تبادل‌نظر دربارهٔ مسئلهٔ ابرداده پرداخته و به نوعی اجماع کلی درباب عناصر ابرداده‌ای مورد نیاز در توصیف منابع الکترونیکی برسند (وایبل و دیگران، ۱۹۹۵).

کارگاه آموزشی ۱۹۹۵ و خلق قالب ابرداده‌ای دابلین کور

همان‌طور که اشاره شد، اساس طرح ابرداده‌ای دابلین کور در کارگاهی آموزشی که در مارس ۱۹۹۵ در شهر دابلین ایالت اوهایو امریکا برگزار گردید، ریخته شد. از مهم‌ترین اهداف برگزاری این کارگاه آموزشی "رسیدن به اجماع کلی در باب یک مجموعه هسته از عناصر ابرداده‌ای برای توصیف منابع شبکه‌ای بود" (وایبل و دیگران، ۱۹۹۵).

از آنجا که اکثر منابع اطلاعاتی موجود در شبکهٔ اینترنت نیاز شدید به روش‌های مؤثر برای بازیابی داشتند و نیز به‌دلیل وسعت و پیچیدگی مسئلهٔ توصیف منابع در شبکه‌های الکترونیکی، در نخستین کارگاه آموزشی، توصیف جنبه‌ها و ویژگی‌های اساسی و اصلی منبع که درکشف منبع مؤثر هستند مود توجه خاص قرارگرفت و دیگر عناصر ابرداده‌ای که می‌توانند نیازهای دیگر را مرتفع سازند مستثنی گردیدند. بنابراین، بحث و تبادل نظر به عناصری محدود شد که درکشف و بازیابی آنچه که تحت عنوان "اشیاء شبه مدرک"[4] نام گرفت به‌کار می‌آیند. فرض بر این بود که اشیاء شبه مدرک معمول‌ترین نوع منابع در اینترنت هستند و هر راه‌حلی که برای آنها ارائه شود می‌تواند برای دیگر انواع منابع نیز به‌کارگرفته شود. اشیاء شبه مدرک به‌صورت دقیق در این کارگاه

1. Online Computer Library Centre (OCLC)
2. National Center for Supercomputing Applications
3. Standard Generalized Markup Language (SGML)
4. Document Like Objects (DLO)

تعریف نشدند، بلکه مثال‌هایی مانند نسخه الکترونیکی روزنامه یا لغتنامه برای فهم ماهیت آنها ارائه گردید. شرکت‌کنندگان در این کارگاه آموزشی هیچ‌گونه تلاشی برای محدودکردن پیچیدگی اشیاء شبه مدرک نکردند، به‌جز اینکه محتوای اشیاء شبه مدرک در وهلهٔ اول متن دانسته شد و اینکه عناصر ابرداده‌ای مورد نیاز برای توصیف آنها در شکل شبیه ابرداده‌هایی هستند که برای توصیف منابع چاپی سنتی به‌کار می‌روند (وایبل و دیگران، ۱۹۹۵).

همان‌طور که گفته شد، در این کارگاه آموزشی عناصری مورد توجه بود که بر کشف منبع متمرکز هستند و عناصر مرتبط با توصیف کامل منابع، مانند حقوق معنوی اثر یا مسائل مالی و هزینه‌ای، از حیطهٔ توجه خارج گردید. نتیجهٔ این کارگاه آموزشی مجموعه‌ای از سیزده عنصر ابرداده‌ای بود که مجموعهٔ عناصر ابرداده‌ای دابلین کور[1] نام گرفت. این عناصر عبارت بودند از: موضوع، عنوان، مؤلف، ناشر، دیگر افراد[2]، تاریخ، نوع شیء[3]، شکل[4]، شناسه[5]، رابطه[6]، منبع، زبان و پوشش (وایبل و دیگران، ۱۹۹۵).

به‌طور کلی نتیجهٔ این بحث‌ها و تبادل‌نظرها را می‌توان به‌صورت زیر خلاصه کرد (دمپسی[7] و وایبل، ۱۹۹۶):

۱. فقط عناصر ابرداده‌ای که در کشف منبع مؤثر هستند مورد توجه قرار گرفتند. هدف توسعهٔ نوعی قالب ابرداده‌ای بود که برای توصیف کتابشناختی منابع الکترونیکی مفید باشد و قالبی ساده که توسط پدیدآوران منابع اطلاعاتی که هیچ‌گونه آموزش خاصی در این زمینه ندیده‌اند قابل استفاده باشد. عناصر ابرداده‌ای مربوط به مسائل مالی، اداری یا حقوقی منابع مستثنی شدند.

۲. بحث‌ها بر عناصری محدود بود که در کشف و بازیابی اشیاء شبه مدرک مورد نیاز هستند و این اشیاء عمدتاً با مثال توصیف شدند.

۳. عناصری ابرداده‌ای که معنای آنها کاملاً قابل فهم باشد مورد نظر بود. بستر نحوی برای عناصر ابرداده‌ای تعیین نشد.

۴. گسترش‌پذیری[8] به‌عنوان ویژگی اصلی برای عناصر ابرداده‌ای مورد توجه قرار گرفت. هدف دابلین کور، کنارزدن دیگر قالب‌های شناخته شدهٔ ابرداده‌ای نبود، بلکه نوعی قالب ابرداده‌ای با عناصر ساده است که می‌تواند در محتدالشکل کردن دیگر قالب‌های پیچیده توصیف مؤثر باشد. بنابراین، توسعهٔ روش‌ها و ابزارهایی برای گسترش عناصر دابلین کور ضروری به‌نظر رسید.

1. Dublin Metadata Core Element Set
2. OtherAgents
3. ObjectType
4. Form
5. Identifier
6. Relation
7. Dempsy
8. Extensibility

۶. عناصر ابرداده‌ای به صورت انتخابی[1]، تکرارپذیر[2] و اصلاح‌پذیر[3] تعریف گردید. عناصر می‌توانند به‌وسیلهٔ توضیحگر[4]ها تعدیل شوند.

دابلین کور و تصاویر شبکه‌ای به‌عنوان اشیاء شبه مدرک

با توجه به کاربرد عناصر دابلین کور برای اشیاء شبه مدرک و تعریف مبهم اشیاء شبه مدرک در اولین کارگاه آموزشی از یک سو، و رشد فزایندهٔ تصاویر مختلف در شبکهٔ جهانی وب از سوی دیگر، این مسئله که آیا تصاویر نیز اشیاء شبه مدرک قلمداد می‌شوند یا هستهٔ ابرداده‌ای دیگری برای توصیف تصاویر شبکه‌ای مورد نیاز است در سومین کارگاه آموزشی دابلین کور مورد توجه خاص قرار گرفت. در این کارگاه آموزشی که در سپتامبر ۱۹۹۶، توسط اُ.سی.ال.سی و ائتلاف اطلاعات شبکه‌ای[5] در شهر دابلین ایالت اوهایو برگزارشد بر سر این مسئله اتفاق‌نظر به‌وجود آمد که تصاویر تفاوت زیادی با اشیاء شبه مدرک ندارند؛ و نیز خصیصهٔ تعیین‌کنندهٔ اشیاء شبه مدرک گرافیکی بودن آنها نیست؛ بلکه ثابت بودن و مرزپذیری آنهاست، به‌نحوی که برای همهٔ استفاده‌کنندگان به یک شکل به‌نظر آیند. بنابراین، تصویر، فیلم، موسیقی، سخنرانی و دیگر اشیاء اطلاعاتی که ثابت هستند (برای همهٔ استفاده‌کنندگان محتوای یکسان و مشابهی دارند) می‌توانند اشیاء اطلاعاتی به شمار آیند. از سوی دیگر، اشیاء غیر شبه مدرک[6]، اشیایی را از قبیل پایگاه‌های اطلاعاتی مجازی[7] (پایگاه‌هایی که برونداد شبه‌مدرک خلق می‌کنند)، گرافیک‌های تجاری و برنامه‌های تعاملی[8]، که ممکن است برای استفاده‌کنندگان مختلف محتوای متفاوت داشته باشند، دربرمی‌گیرند (وایبل و میلر[9]، ۱۹۹۶).

با توجه به اینکه تصاویر، خصوصاً تصاویر الکترونیکی، دارای خصیصه‌هایی هستند که متفاوت از اشیاء اطلاعاتی متنی هستند (مانند نوع تصویر، قالب تصویر، جدول‌های رنگی، ابعاد، پویش[10]های متفاوت از تصویر، درجهٔ تحرک[11] درجهٔ شفافیت[12] و...) که همگی آنها در توصیف تصاویر از اهمیت زیادی برخوردارند، اما پرداختن به همگی این عناصر توصیفی با یکی از اهداف و ویژگی‌های زیربنایی قالب دابلین کور ــ یعنی سادگی ــ در تعارض است. اگر قرار بود که عناصر دابلین کور به روشی مفید برای توصیف تصاویر به‌کار گرفته شوند، جداسازی هستهٔ اصلی اطلاعات توصیفی مناسب برای یک رکورد توصیفی ساده و شناسایی یک روش گسترش به‌منظور ارائهٔ عناصر توصیفی غنی‌تر برای تصاویر لازم آمد. بحث‌ها و تبادل نظرهای

1. Optional
2. Repeatable
3. Modifiable
4. Qualifier
5. Coalition for Networked Information (CNI)
6. None-Document-Like Object
7. Virtual databases
8. Interactive applications
9. Miller
10. Scan
11. Dynamic range
12. Resolution

به‌عمل آمده منجر به تغییراتی در مجموعهٔ عناصر ابرداده‌ای دابلین کور گردید که طبق آن برخی عناصر مورد اصلاح و تغییر قرار گرفتند تا مقداری از حالت متن‌مداری[1] بیرون آیند، و نیز دو عنصر ابرداده‌ای جدید به مجموعهٔ عناصر ابرداده‌ای پیشین افزوده شد.

یکی از تغییراتی که در عناصر ابرداده‌ای به‌وجود آمد جداشدن موضوع و توصیف[2] بود که طی آن عنصر موضوع و عنصر توصیف جداگانه مورد توجه قرار گرفت، که بخشی از آن به دلیل توافق نظر در میان متخصصان تصویر بود که این عناصر را مفاهیمی کاملاً مجزّا و متفاوت برای تصاویر می‌دانستند. بنابراین، اصطلاحات واژگان کنترل‌شده[3]، کلیدواژه‌ها، و دیگر طرح‌های رده‌بندی رسمی در عنصر موضوع قرار می‌گیرند در حالی که عنصر توصیف برای توصیف یا توصیف متنی در محتوای تصاویر به‌کار برده می‌شود و همچنین، در مورد مدارک متنی، مکانی برای چکیده است. از دیگر تغییرات اصلی در عناصر دابلین کور اضافه‌شدن عنصر مدیریت حقوق[4] بود. گرچه این عنصر به توصیف خصیصه‌های بیرونی منبع می‌پردازد (که طبق اهداف اولیهٔ دابلین کور قابل بحث است)، اهمیت آن از طرف بسیاری از شرکت‌کنندگان در این کارگاه آموزشی مورد تأکید قرار گرفت. به‌طور کلی، تغییرات به‌وجود آمده در عناصر دابلین کور طی این کارگاه آموزشی را می‌توان به‌صورت زیر خلاصه کرد (وایبل و میلر، ۱۹۹۶):

۱. عنصر مؤلف به "مؤلف یا پدیدآورنده[5]" تغییر یافت.
۲. عنصر موضوع به "موضوع یا کلیدواژه‌ها" تغییر یافت.
۳. عنصر توصیف به مجموعه عناصر ابرداده‌ای افزوده شد.
۴. عنصر دیگر افراد به "دیگر همکاران[6]" تغییر یافت.
۵. عنصر "مدیریت حقوق" به تعداد عناصر افزوده شد.

طرح ابرداده‌ای دابلین کور: رسالت و فعالیت‌ها

بیانیهٔ رسالت دابلین کور در هشتمین کارگاه آموزشی که در سال ۲۰۰۰ در اوتاوا ـ کانادا برگزار شد با اتفاق نظر اعضای کمیتهٔ مشورتی تصویب شد. طبق این بیانیه، رسالت دابلین کور کمک به بهبود کشف منبع در شبکهٔ وب از طریق فعالیت‌های زیر است (وایبل و کخ[7]، ۲۰۰۰):

ـ توسعهٔ استانداردهای ابرداده‌ای برای کشف منبع به صورت بین حوزه‌ای؛

1. Text-centric
2. Description
3. Controlled vocabulary term
4. Rights management
5. Creator
6. Other contributors
7. Koch

ـ تعریف چارچوب‌هایی برای میان کنش‌پذیری[1] مجموعه‌های ابرداده‌ای؛

ـ تسهیل توسعهٔ مجموعه عناصر ابرداده‌ای خاص حوزه‌ها یا رشته‌ها، به‌طوری که بتوانند در چارچوب‌های کشف میان حوزه‌ای و میان کنش‌پذیری ابرداده‌ای کار کنند.

حیطهٔ فعالیت‌های دابلین کور بر اساس این رسالت عبارتند از:

۱. مستندسازی و ارتباطات؛ شامل:

ـ نگهداری وب سایت طرح ابرداده‌ای دابلین کور به‌عنوان منبع معتبر استانداردهای مربوط به طرح دابلین کور، پشتیبانی از مدارک، منابع مرتبط و اخبار مربوط به فعالیت‌های گروه‌های کاری[2]؛

ـ ارتقاء و پشتیبانی از توسعهٔ مواد و رخدادهای آموزشی که به اشاعهٔ اطلاعات مربوط به استانداردها و به‌کارگیری دابلین کور کمک می‌کند؛

ـ پشتیبانی از فعالیت‌های ارتباطی با دیگر حوزه‌های ابرداده‌ای برای ارتقاء همگرایی استانداردهای ابرداده‌ای در صورت عملی‌بودن، و پشتیبانی از گذرهای متقابل[3] در صورت مناسب‌بودن.

۲. توسعه و نگهداری استانداردها؛ شامل:

ـ سازماندهی کارگاه‌های آموزشی بین‌المللی و نشست‌های گروه‌های کاری در جهت نگهداری مدارک و پیشنهادهای موجود طرح ابرداده‌ای دابلین کور و توسعهٔ عناصر و بیانگرهای اضافی خاص حوزه‌های موضوعی؛

ـ شرکت در فعالیت‌های استانداردهای بین‌المللی به‌منظور بالابردن میزان پذیرش طرح به‌عنوان استانداردی جهانی.

۳. ابزارها و خدمات؛ شامل:

ـ توسعهٔ زیرساخت نرم‌افزاری (مانند بایگانی ابرداده‌ای طرح ابرداده‌ای دابلین کور[4]) برای پشتیبانی از فرایندهای مدیریت و نگهداری ابرداده‌های دابلین کور در زبان‌های مختلف؛

ـ فراهم‌آوری دسترسی قابل اطمینان به طرح‌های ابرداده‌ای برای مجریان، برنامه‌های کاربردی و استفاده‌کنندگان؛

ـ مشورت با طراحان نظام‌های ابرداده‌ای برای افزایش کاربرد دائمی و منسجم عناصر ابرداده‌ای دابلین کور.

1. Interoperability
2. Working group
3. Crosswalks
4. DCMI Metadata Registry

عناصر ابرداده‌ای دابلین کور

از آنجا که پیامد سومین کارگاه آموزشی تثبیت پانزده عنصر ابرداده‌ای به‌عنوان مجموعهٔ عناصر دابلین کور بود، پس از آن، مرحلهٔ توسعهٔ تعداد عناصر پایان یافت و مرحلهٔ تمرکز بر به‌کارگیری و استفاده از عناصر ابرداده‌ای آغاز گردید (وایبل، آینلا و کاترو، ۱۹۹۷).

معنای عناصر ابرداده‌ای در کارگاه آموزشی هلسینکی تثبیت شد و می‌توان آن را اساس نخستین استانداردسازی رسمی عناصر دابلین کور دانست. این کارگاه آموزشی، پنجمین کارگاه آموزشی دابلین کور بود که در اکتبر ۱۹۹۷ در شهر هلسینکی فنلاند با شرکت حدود هفتاد نفر از شصت کشور مختلف و با حمایت اُ.سی. ال. سی، کتابخانهٔ ملی فنلاند، و ائتلاف اطلاعات شبکه‌ای[1] برگزار شد. اکثر عناصر ابرداده‌ای، پذیرش عام و قطعی یافتند اما در مورد برخی از آنها بحث‌ها و تبادل‌نظرهای زیادی صورت گرفت که در نهایت منجر به تغییر و تعدیل تعاریف آنها گردید. حوزه‌های عمدهٔ مورد اختلاف و بحث در هلسینکی سه عنصر بود: تاریخ، پوشش، و رابطه.

عنصر تاریخ از همان ابتدای طرح ابرداده‌ای دابلین کور مورد اختلاف‌نظر بود و بحث‌های بسیاری در مورد آن انجام شده بود. تاریخ‌های زیادی در چرخهٔ حیات یک منبع وجود دارد که از اهمیت برخوردارند، اما اولویت‌دادن به یکی از آنها، بیشتر بر اساس کاربردهای خاص ابرداده‌هاست. بحث‌های فراوان در مورد این عنصر، به تغییر و تعدیل تعریف عنصر تاریخ منجر شد: تاریخی مرتبط با خلق یا دسترس‌پذیری منبع. همچنین انواع دیگری از تاریخ که در کاربردهای خاصی از اهمیت برخوردار است مشخص گردید.

عنصر پوشش نیز از عناصری بود که از همان آغاز دابلین کور مورد بحث و اختلاف‌نظر شدید قرار گرفته بود. آوردن این عنصر از نخستین کارگاه آموزشی دابلین کور مورد بحث بود و گرچه حامیان استفاده از این عنصر به دنبال روش‌هایی برای به‌کارگیری آن بودند، هیچ‌گونه راهنما یا روشی برای کاربرد آن وجود نداشت. تکامل وب نیاز به استفاده از این عنصر را برجسته‌تر ساخت. مطالعاتی که نشان می‌داد هدف بیشتر جست‌وجوهای اطلاعاتی در وب یافتن منابع محلی است، بر رشد سریع وب به‌عنوان نوعی منبع اطلاعات محلی برای مقاصد تجاری و فرهنگی صحه گذاشتند. توافق در استفاده از عنصر پوشش به‌منظور پشتیبانی از جست‌وجوی اطلاعات در یک حوزهٔ جغرافیایی خاص بود. استفاده از این عنصر برای شناسایی دورهٔ زمانی خاص (پوشش زمانی) نیز مورد توجه قرار گرفت؛ اما انتظار می‌رفت که این نوع کاربرد فقط بخش

1. Coalition for Networked Information (CNI)

کوچکی از کاربردهای عنصر پوشش را در عناصر ساده (بدون بیانگر) دابلین کور دربرگیرد.

مسئلهٔ روابط میان منابع در محیط‌های الکترونیکی که احتمال وجود شکل‌های مختلف و مرتبط منابع بسیار زیاد است از اهمیت فراوانی برخوردار است. ارائهٔ ابرداده‌های منسجم برای منابعی که از دیگر منابع مشتق شده‌اند و ممکن است ابرداده‌های مربوط به خود را داشته باشند نیز اهمیت خاصی دارد. این مسئله در یک نشست ابرداده‌ای که در ژوئیه ۱۹۹۷ در گروه کتابخانه‌های پژوهشی[1] برگزار گردید مطرح شد و در کارگاه آموزشی هلسینکی مورد توجه خاص قرار گرفت. بحث و تبادل‌نظر در این مورد به اصلی منجر شد که تحت عنوان "اصل ۱:۱" شناخته می‌شود. طبق این اصل هر منبع باید توصیف ابرداده‌ای مجزّایی داشته باشد و هر توصیف ابرداده‌ای باید دربرگیرندهٔ عناصری مرتبط با منبعی واحد باشد. برقراری پیوند میان این توصیف‌ها به روشی منسجم و پایدار ضروری است. عنصر پیوند ابزار منطقی برای فراهم آوردن چنین ارتباطی است. در برقراری پیوند میان منبع مبدأ و منبع مقصد، رابطهٔ نامگذاری‌شده[2] ــ تعیین نوع رابطه میان منابع ــ از اهمیت کلیدی برخوردار است. به همین سبب، نوع رابطه[3]های متعددی که در برقراری پیوند میان منابع مختلف می‌توانند به‌کار آیند مشخص گردید (وایبل و هاکالا[4]، ۱۹۹۸).

استانداردسازی عناصر دابلین کور در ششمین کارگاه آموزشی دابلین کور که از دوم تا چهارم نوامبر ۱۹۹۸ در شهر واشنگتن دی سی امریکا برگزار شد نیز از مسائل مهم مورد توجه بود. استانداردسازی دابلین کور از طرق مختلف مورد توجه قرار گرفته بود که اولین آنها، استفاده از استاندارد نیروی کاری مهندسی اینترنت[5] بود. نخستین گام‌ها در استانداردسازی، به‌کارگیری استانداردهای LETF RFC بود که اولین توصیف و ارائهٔ رسمی معنای ۱۵ عنصر ابرداده‌ای دابلین کور از طریق استاندارد RFC 2413 انجام گرفت و با عنوان DC 1.0 شناخته شد. مرحلهٔ بعدی استانداردسازی دابلین کور، ساختاردهی مجدد RFC 2413 و ارائهٔ آن به سازمان ملی استانداردهای اطلاعاتی[6] و مرکز معیارین‌سازی اروپا[7] بود که قبلاً نیز دابلین کور در موضوع کار آنها قرار داشت. با جرح و تعدیل در RFC قبل از استانداردسازی موافقت شد؛ و بنابراین، تغییرات برنامه‌ریزی شده در دو دسته قرار گرفت: اول بازنگری تعاریف عناصر ابرداده‌ای برای بالابردن وضوح و روشنی آنها بود که این کار برعهدهٔ گروه‌های کاری قرار داده شد، و دومین تغییر پیشنهادی تغییر در شکل ارائه و توصیف عناصر دابلین کور بر طبق نوعی استاندارد

1. Research Libraries Group (RLG)
2. Named relation
3. Relation type
4. Hakala
5. Internet Engineering Task Force (IETF)
6. National Information Standards Organization (NISO)
7. Center for European Normalization (CEN)

توصیف عناصر ابرداده‌ای بود که به نام استاندارد ISO 11179 شناخته می‌شد. این استاندارد روشی را برای توصیف یکدست عناصر داده‌ای ارائه می‌داد (وایبل، ۱۹۹۹).

آخرین ویرایش عناصر ابرداده‌ای دابلین کور ویرایش ۱/۱ است که دو نگارش معتبر از آن وجود دارد. در اوایل سال ۲۰۰۰، نگارش ۱/۱ عناصر ابرداده‌ای دابلین کور تحت عنوان توافقنامه کارگاه آموزشی سِن[1] مورد پذیرش قرار گرفت و در سپتامبر سال ۲۰۰۱ نگارش جدیدتری از عناصر ابرداده‌ای دابلین کور با کمی جرح و تعدیل تحت حمایت سازمان ملی استانداردهای اطلاعاتی تصویب شد (دابلین کور، ۲۰۰۳؛ دکرز[2] و وایبل، ۲۰۰۲). در جدول ۱ مجموعهٔ پانزده عنصر ابرداده‌ای دابلین کور و تعاریف آنها براساس استاندارد سازمان ملی استانداردهای اطلاعاتی نشان داده شده است (دابلین کور، ۲۰۰۳):

جدول ۱. عناصر ابرداده‌ای دابلین کور

	نام عنصر	تعریف
۱	عنوان	نامی که به منبع داده می‌شود.
۲	پدیدآورنده	موجودیتی که اساساً مسئولیت خلق محتوای اثر را دارد.
۳	موضوع	موضوع محتوای منبع
۴	توصیف	گزارشی از محتوای منبع
۵	ناشر	موجودیتی که مسئول دسترس‌پذیرسازی منبع است.
۶	پدیدآورنده همکار	موجودیتی که مسئولیت همکاری در خلق محتوای منبع را دارد.
۷	تاریخ	تاریخی مربوط به یک رویداد در چرخهٔ زندگی منبع
۸	نوع	ماهیت یا گونهٔ محتوایی منبع
۹	قالب	شکل رقومی یا فیزیکی منبع
۱۰	شناسه	ارجاعی واضح به منبع در بافتی مشخص
۱۱	منبع	ارجاع به منبعی که منبع حاضر از آن مشتق شده است.
۱۲	زبان	زبان محتوای فکری منبع
۱۳	رابطه	ارجاعی به یک منبع مرتبط
۱۴	پوشش	وسعت یا حوزهٔ محتوای منبع
۱۵	حقوق	اطلاعات مربوط به حقوق منبع

1. CEN Workshop Agreement
2. Dekkers

تعدیل عناصر ابرداده‌ای: ضرورت توضیحگرهای[1] ابرداده‌ای

در اولین کارگاه آموزشی دابلین کور قابلیت پالایش معنا و ارزش[2] عناصر ابرداده‌ای برای ارائه توصیف‌های غنی‌تر، با استفاده از توضیحگرها، به‌عنوان یکی از ویژگی‌های اساسی عناصر ابرداده‌ای دابلین کور مورد توجه قرار گرفت (دمپسی و وایبل، ۱۹۹۶). زمانی که دابلین کور در حوزه‌های مختلف مورد استفاده قرار گرفت، به تبع آن، نیاز به ساختارمندی بهتر و روابط معناشناختی خاص‌تر نیز احساس شد. کاربردهای خاصی از عناصر ابرداده‌ای دابلین کور نیازمند پالایش و توصیف عناصر ابرداده‌ای و ارزش‌های آنها بود. این نیاز باعث خلق و توسعهٔ توضیحگرهای عناصر ابرداده‌ای گردید. این توضیحگرها اخصیّت معناشناختی بهتری را در توصیف منابع حوزه‌های خاص فراهم می‌کنند و به‌عنوان مکمل غیراجباری و جداگانه‌ای برای عناصر ابرداده‌ای دابلین کور خلق شده‌اند. این توضیحگرها، معنا و مفهوم اساسی عناصر توصیفی را تغییر نمی‌دهند بلکه آنها را پالایش می کنند. در فرایند مبادلهٔ اطلاعات، توضیحگرها می‌توانند جدا شوند و فقط عنصر پایه‌ای و اصلی باقی بماند (اسلاویک و بایجت[3]، ۲۰۰۱).

مسئلهٔ توضیحگرهای عناصر ابرداده‌ای در چهارمین کارگاه آموزشی دابلین کور (کارگاه آموزشی کانبرا) در سال ۱۹۹۷ (وایبل ، آینلا و کاترو، ۱۹۹۷)، ششمین کارگاه آموزشی در واشنگتن دی سی در ۱۹۹۸ (وایبل، ۱۹۹۹) و سرانجام در هفتمین کارگاه آموزشی در فرانکفورت (وایبل و کخ، ۲۰۰۰)، مورد توجه خاص قرار گرفت. در نهایت دو نوع بیانگر تعریف شدند: توضیحگرهای پالایش عنصر ابرداده‌ای[4] و توضیحگرهایی طرح‌های کدگذاری[5]. توضیحگرهایی که مربوط به پالایش عنصر هستند معنای عنصر ابرداده‌ای را محدودتر یا خاص‌تر می‌سازند؛ مانند پیشنهاد استفاده از فهرست مطالب یا چکیده به‌عنوان عنصر توصیف. توضیحگرهای مربوط به طرح‌های کدگذاری عنصر، طرح‌های استانداردی را مشخص می‌سازند که در تفسیر و فهم ارزش و معنای عناصر ابرداده‌ای سودمندند. به‌طور مثال، پیشنهاد استفاده از سرعنوان‌های موضوعی کتابخانهٔ کنگره برای کدگذاری عنصر موضوع یا به‌کارگیری استاندارد ISO639-21 برای عنصر زبان (توضیحگرهای دابلین کور، ۲۰۰۰).

کارکردهای عناصر ابرداده‌ای دابلین کور

وظایف عناصر تعریف شده در دابلین کور را می‌توان بر اساس وظایف اساسی

1. Qualifier
2. Value
3. Slavic & Baiget
4. Element refinement
5. Encoding schemes

کاربران اطلاعات تقسیم‌بندی کرد. طبق الگویی که ایفلا ارائه داده است وظایف و نیازهای اساسی کاربران اطلاعات را می‌توان به‌صورت زیر تقسیم کرد (ایفلا، ۱۹۹۸):

۱. "یافتن" یا "کشف" موجودیتی[1] (منبعی) که با معیارهای ذکر شدهٔ کاوش فرد منطبق است (به‌طور مثال مکان‌یابی یک منبع یا موجودیت در یک پایگاه اطلاعاتی، با کاوش خصیصه یا خصیصه‌هایی از آن موجودیت).

۲. "شناسایی" موجودیت (تصدیق اینکه موجودیت توصیف شده با موجودیت جست‌وجوشده منطبق است یا تمایز بین دو یا چند موجودیت که دارای خصیصه‌ها و ویژگی‌های مشابهی هستند).

۳. "انتخاب" موجودیتی که با نیازهای اطلاعاتی فرد منطبق است (به‌طور مثال، انتخاب یک موجودیت بر حسب محتوا، شکل فیزیکی و ..که نیازهای اطلاعاتی فرد را رفع می‌کنند و یا رد آن موجودیت به دلیل نامناسب بودن بر حسب نیازهای اطلاعاتی فرد).

۴. "دسترسی" به موجودیت توصیف شده (به‌طور مثال، دست یافتن به موجودیت از طریق خرید، امانت، یا دسترسی به آن از طریق شبکه‌های پیوسته).

از آنجاکه استاندارد ابرداده ای دابلین کور نیز به‌منظور تسهیل فرایند کشف و بازیابی منابع توسعه یافته است، تقسیم‌بندی عناصر تعریف‌شده در این استاندارد ابرداده‌ای برحسب توانشان در رفع نیازهای استفاده‌کنندگان و کاربران اطلاعات می‌تواند تصویر روشن‌تری از آن بر اساس وظایف و کارکردهایش ارائه دهد. کاربردپذیری هر یک از عناصر ابرداده‌ای تعریف‌شده در دابلین کور بر اساس وظایف ذکرشده (دابلین کور، ۱۹۹۸) به قرار زیر است:

۱. عناصر مورد استفاده برای کشف منبع. عناصر دابلین کور اساساً به‌منظور پشتیبانی از کشف و یافتن منابع الکترونیکی توسعه یافته‌اند و هدف آنها ارائهٔ اطلاعات توصیفی اساسی است که کاربران بتوانند به کاوش اطلاعات مورد نیاز خود بپردازند. این عناصر عبارتند از: عنوان، پدیدآورنده، همکار و موضوع. علاوه‌بر این عناصر، که عناصر اصلی در جست‌وجو و کشف منبع محسوب می‌شوند، عناصر دیگری هستند که از اهمیت ثانویه در امر کشف منبع برخوردارند و به‌عنوان عناصر محدودکنندهٔ کاوش به‌کار می‌روند مانند عناصر پوشش، زبان و قالب.

۲. عناصر مورد استفاده برای شناسایی منبع. توانایی شناسایی منبعی خاص از میان

1. Entity

منابع بازیابی‌شده، و تشخیص منابع مشابه از اهداف روشن و واضح دابلین کور نبوده است. ولی، با این حال، دسته‌ای از عناصر دابلین کور به شناسایی و معرفی منابع می‌پردازند که عبارتند از: تاریخ، نوع منبع، قالب و شناسه.

۳. عناصر مورد استفاده برای انتخاب منبع. عناصر دابلین کور کلیهٔ اطلاعات لازم برای انتخاب منبعی خاص از میان دیگر منابع بازیابی شده را برای کاربر فراهم نمی‌کنند ولی میزان معیّنی اطلاعات به این منظور در اختیار وی قرار می‌گیرد. به‌ویژه عنصر توصیف از اهمیت خاصی برخوردار است. دیگر عنصری که در انتخاب منبع می‌تواند مورد استفاده قرار گیرد عنصر پوشش است.

۴. عناصر مورد استفاده برای فراهم‌آوری و دسترسی به منبع. در محیط شبکه‌ای، دستیابی به منبع می‌تواند از طریق آدرس دقیق منبع اطلاعاتی که از طریق عنصر شناسه ارائه می شود انجام گیرد.

دابلین کور و بستر نحوی انتقال ابرداده‌ها

در نخستین کارگاه آموزشی دابلین کور در سال ۱۹۹۵ هیچ‌گونه بستر نحوی برای انتقال ابرداده‌های دابلین کور مشخص نگردید (وایبل و دیگران، ۱۹۹۵). در دومین کارگاه آموزشی (کارگاه آموزشی وارویک در سال ۱۹۹۶) در حالی که تنها یک سال از توسعهٔ اولیهٔ قالب دابلین کور می گذشت، حوزه‌های گوناگونی از عناصر دابلین کور در توصیف منابع خود استفاده کرده بودند. از آن میان می‌توان از طرح ابرداده‌ای کشورهای اسکاندیناوی[1]، طرح میان کنش‌پذیری نام متحدالشکل منبع[2]، طرح کتابخانهٔ دیجیتالی اسکندریه[3] و بسیاری طرح‌های دیگر نام برد. در این طرح‌ها، نبود الگویی مشترک برای به‌کارگیری ابرداده‌های دابلین کور، روش‌ها و بسترهای نحوی مختلفی برای ورود این عناصر ابرداده‌ای به‌کار گرفته شدند. واضح است که استفادهٔ وسیع از هرگونه قالب ابرداده‌ای به طریقی نیازمند توافق و اجماع در باب بستر نحوی برای انتقال آن است. به این دلیل و نیز به سبب اینکه وب به عنوان اصلی‌ترین رسانه در اینترنت مورد توجه بود، ارائهٔ بستر نحوی برای به‌کارگیری این عناصر در وب مورد توجه قرار گرفت. در نتیجه، یک گروه کاری با هدف شرح بستر نحوی عناصر در این کارگاه آموزشی شکل گرفت. علاوه بر این، قراردادی برای واردکردن عناصر ابرداده‌ای در زبان نشانه‌گذاری فرامتن[4] در کارگاه آموزشی نمایه‌سازی و کاوش توزیعی وب[5]،

1. Nordic Metadata Project
2. Uniform Resource Name Interoperability Project (TURNIP)
3. Alexandria Digital Library (ADL)
4. Hyper Text Mark up Language (HTML)
5. W3C Distributed Indexing and Searching Workshop

که در همان سال با حضور نمایندگانی از دابلین کور، کارگاه آموزشی وارویک، و بسیاری از فراهم‌آورندگان خدمات وب مانند Lycos، WebCrawler، Microsoft، فروشندگان نرم‌افزار و ائتلاف شبکهٔ جهانی وب برگزار شد، پیشنهاد گردید (دمپسی و وایبل، ۱۹۹۶).

مسئلهٔ رسمی‌سازی بستر نحوی ابرداده‌های دابلین کور در ششمین کارگاه آموزشی نیز مورد بحث و بررسی قرار گرفت. گرچه اولین قرارداد در مورد شیوهٔ کدگذاری دابلین کور در زبان نشانه‌گذاری فرامتن از سال ۱۹۹۶ در دست بود، به دلیل تغییراتی که در این زبان به وجود آمد و نیاز به رسمی‌سازی بیشتر، در یک پیش‌نویس اینترنتی[1] که توسط جان کانز[2] (۱۹۹۹) منتشر گردید، به شیوهٔ واردکردن عناصر ابرداده‌ای دابلین کور در زبان نشانه‌گذاری فرامتن پرداخته شد. امروزه ساده‌ترین راه به‌کارگیری ابرداده‌ای برای توصیف منابع در وب واردکردن آنها در زبان نشانه‌گذاری فرامتن است. فایدهٔ وارد کردن عناصر ابرداده‌ای در مدارک وب این است که نظام دیگری برای استفاده از ابرداده‌ها مورد نیاز نیست. همچنین، ابرداده‌ها جز جدایی‌ناپذیری از مدارک می‌شوند و ممکن است توسط ابزارهای نمایه‌سازی وب نمایه گردند (وایبل، ۱۹۹۹). کاربردهای کنونی دابلین کور نیز غالباً براساس وارد کردن عناصر آن در زبان نشانه‌گذاری فرامتن است که از طریق ابرنشانه[3]های این زبان صورت می‌گیرد.

با پیدایش زبان نشانگذاری گسترش‌پذیر[4] و جهت‌گیری وب به سمت استفاده از این زبان در طراحی و ارائهٔ خدمات وب، شیوهٔ به‌کارگیری عناصر دابلین کور در بستر نحوی این زبان مورد توجه قرار گرفت. وارد کردن عناصر ابرداده‌ای در این زبان از طریق چارچوبی که به عنوان "چارچوب توصیف منبع[5]" شناخته می‌شود انجام می‌گیرد. چارچوب توصیف منبع که اختصاراً RDF خوانده می‌شود در اواسط سال ۱۹۹۸ توسط ائتلاف شبکهٔ جهانی وب به‌عنوان معماری خاصی برای شبکه وب خلق گردید. آنچه که این چارچوب انجام می‌دهد و آن را از دابلین کور متمایز می‌سازد این است که توصیف روابط معناشناختی و نحوی برای مجموعه‌های ابرداده‌ای متعدد و ناهمگون را میسّر می‌سازد (آرون هایم[6]، ۱۹۹۸). این چارچوب به‌عنوان زیرساختی عمل می‌کند که در آن، هرگونه استاندارد ابرداده‌ای قابل استفاده است و از این طریق میان کنش‌پذیری معناشناختی[7] در میان استانداردهای متعدد و مختلف ابرداده‌ای را ممکن می‌سازد (مدیروس[8]، ۲۰۰۰).

1. Internet draft
2. John Kunze
3. Meta tag
4. eXtensible Mark up Language (XML)
5. Resource Description Framework (RDF)
6. Ahronheim
7. Semantic interoperability
8. Medeiros

حاصل سخن

در میان قالب‌های مختلف ابرداده‌ای خلق‌شده، قالب ابرداده‌ای دابلین کور، احتمالاً به‌دلیل ویژگی‌ها، اهداف و کارکردهایش، نقش مهمی در توصیف منابع الکترونیکی خصوصاً منابع موجود در شبکهٔ جهانی وب خواهد داشت. سادگی قالب دابلین کور این امکان را به پدیدآورندگان منابع وب می‌دهد که بدون نیاز به دانش تخصصی در فهرستنویسی، از عناصر تعریف‌شده توسط این استاندارد در توصیف منابع استفاده کنند. شبکهٔ جهانی وب به عنوان بزرگ‌ترین شبکهٔ اطلاعاتی، نیازها و اضطرارهای خاصی را در زمینهٔ سازماندهی و بازیابی منابع به‌وجود آورده است. کارگاه‌های آموزشی بین‌المللی دابلین کور به‌طور مداوم با هدف حل مشکلات و توسعهٔ این استاندارد ابرداده‌ای شکل می‌گیرند و با تمرکز بر این نیازها و اضطرارها، در پی گسترش کاربردپذیری ابرداده‌های دابلین کور در شبکهٔ وب هستند. سیر تحول این قالب ابرداده‌ای، از جنبه‌های مختلفی چون عناصر و بیانگرهای ابرداده‌ای، بستر نحوی انتقال ابرداده‌ها، میانکنش‌پذیری ابرداده‌ها و دیگر مسائل، نشان‌دهندهٔ کوشش این قالب در پاسخگویی به پرسش‌های مطرح شده دربارهٔ سازماندهی منابع وب و رفع نیازهای این شبکه است. میزان کاربرد و رواج گستردهٔ این قالب ابرداده‌ای، مسلماً، به میزان موفقیت آن در رفع نیازها و الزام‌های محیط‌های الکترونیکی به‌طور اعم، و شبکهٔ جهانی وب به‌طور اخص، بستگی خواهد داشت.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

مآخذ

شیری، علی‌اصغر (۱۳۸۱). "ابرداده". *دایرةالمعارف کتابداری و اطلاع‌رسانی*، ج. ۱، ص. ۷۸-۸۰.

Ahronheim, Judith R. (1998), "Descriptive metadata: emerging standards". *Journal of Academic Librarianship*, 24(5): 395-399.

Bar-Ilan, Judit (1999), "Search engine results over time - A case study on search engine stability". *International Journal of Scientometrics, Informetrics and Bibliometrics*, 2/3(1). [online] Available: http:// www.cindoc.csic.es/ cybermetrics/ articles/ v2ilp1. html

Beacom, Matthew (2000). "Crossing a digital divide: AACR2 and unaddressed problems of networked resources". [online] Available: http:// www.loc.gov/ catdir/ bibcontrol/

patton_paper.html

Burnett, Kathleen; Ng, Kwong Bor; Park, Soyeon (1999)." A comparison of the two traditions of metadata developments". *Journal of the American Society for Information Science*, 50(13): 1209-1217.

Day, Michael. (1999). "Metadata and electronic information". [online] Available: http://www. ukoln.ac.uk/metadata/presentations/circe/birmingham.html

Dekkers, Makx; Weibel, Stuart, L. (2002), "Dublin core metadata initiative: progress report and workplan for 2002". *D-Lib Magazine*, 8(2). [online] Available: http:// www.dlib. org/dlib/february02/weibel/02weibel.html

Dempsey, Lorcan; Weibel, Stuart L.(1996), "The warwick metadata workshop: a framework for the development of resource description". *D-Lib Magazine*, July/ August. [online] Available: http://www.dlib.org/dlib/july96/07weibel.html

Dublin core metadata and the cataloging rules (1998). [online] Available: http:// www. libraries.psu.edu/iasweb/personal/jca/dublin/dcreport.html

Dublin core metadata element set, version 1, 1: Reference description. (2003). [online] Available: http://dublincore.org/documents/2003/02/04/dces/

Dublin core qualifiers (2000). [online] Available: "http:// dublincore. org/ documents/ 2000/ 07/ 11/dcmes-qualifiers

Ercegovac, Zorana. (1999), "Introduction". *Journal of American Society for Information Science*, 50(13): 1165-1168.

Gordon M.; Pathak P. (1999), "Finding information on the World Wide Web: the retrieval effectiveness of search engines". *Information Precessing and Management*, 35(2): 141-180.

Greenberg, Jane (2002)."Metadata and the World Wide Web". *Encyclopedia of Library and Information Science*, Vol. 72, Sup. 35, pp. 244-261.

Heery, Rachel (1996), "Review of metadata formats". *Program*, 30(4): 345-373.

Iannella, Renato (1999), "The future of internet metadata". [inline] Available: http:// archive. dstc.edu.au/RDU/reports/HKSYM99/

IFLA Study Group on the functional requirements for bibliographic records (1998). "Functional requirements for bibliographic records: final report". [online] Available: http:// www.ifla.org/VII/s13/frbr/frbr.htm

Kunze, J. (1999),"Encoding dublin core metadata in HTML". [online] Available: http:// www. fuqs.org/rfcs/rfcs2731/html

Lagoze, Carl (2000), "Business unusual: How Event-awareness may Breathe life into the catalog?". [online] Available: http://lcweb.loc.gov/catdir/bibcontrol/lagoze_paper.html

Lawrence, S. and Giles, C. L. (1999). "Accessibility and distribution of information on the Web". *Nature*, 400, pp. 107-110

Lee-Smeltzer, Kuang-Hwei (2000), "Finding the needle: controlled vocabularies, resource discovery and dublin core". *Library Collections, Acquisitions & Technical Services*, 24, pp. 205-215.

Medeiros, Norm. (2000). "XML and the resource description framework: The great Web Hope", *ONLINE*, September 2000, [online] Available: http://www.onlineinc.com/ onlinemag/DL2000/medeiros9.html

Mettrop, Wouter; Nieuwenhuysen, Paul (2001). "Internet search engines - fluctuations in document accessibility". *Journal of Documentation*, 57 (5): 623-651.

Peterson, R. E. (1997), "Eight internet search engines compared". *First Monday*, 2(2). [online] Available: http://www.firstmonday.dk/issues/issue2_2/peterson/index.html

Rousseau, R. (1999), "Time evolution of the number of hits in keyword searches on the internet". Post conference seminar - cybermetrics'99 at the seventh international conference on scientometrics and informetrics, Colima, Mexico. [online] Available: www.cindoc.csic.es/cybermetrics/cybermetrics99.html

Slavic, Aida; Baiget, Clara (2001). "Using dublin core in educational material: Some practical considerations based on the EASEL experience". *VINE*, 125, pp.74-85.

Weibel, Stuart; Godby, Jean; Miller, Eric; Daniel, Ron. (1995), "OCLC/NCSA Metadata Workshop Report". [online] Available: http://www.ifla.org/documents/libraries/ cataloging/ oclcmeta.htm

Weibel, Stuart; Iannella, Renato; Cathro, Warwick (1997). "The 4th dublin core metadata report: DC-4, March 3-5, 1997, National Library of Australia, Canberra". *D-Lib Magazine*, June 1997. [online] Available: http://www.dlib.org/dlib/june97/metadata/06weibel.html

Weibel, Stuart; Miller, Eric, (1996). "Image description on the internet: A summary of the CNI/OCLC image metadata workshop, september 24-25, 1996, Dublin, Ohio", *D-Lib Magazine*, January 1997. [online] Available: http://www.dlib.org/dlib/january97/oclc/01weibel. html

Weible, Stuart (1999). "The state of the dublin core metadata initiative. *D-Lib Magazine*. 5(4). [online] Available: http://www.dlib.org/dlib/april99/04weibel.html

Weible, Stuart L.; Koch, Traugott (2000). "The dublin core metadata initiative: Mission, current activities, and future directions". *D-Lib Magazine*, 6(12). [online] Available at: http://www.dlib.org/dlib/december00/weibel/12weibel.html

Weible, Stuart; Hakala, Juha (1998). "DC-5: The helsinki metadata workshop, a report on the workshop and subsequent developments". *D-Lib Magazine*. [online] Available: http://www.dlib.org/dlib/february98/02weibel.html

Weiss, Amy K.; Carstens, Timothy V. (2001),"The year's work in cataloging, 1999", *Library Resources and Technical Services*, 45(1):47-53.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی