

# ابرداده

## و اهمیت آن در سازماندهی و بازیابی اطلاعات

\*زهرا حسن‌خانی

کارشناس ارشد کتابداری و اطلاع‌رسانی سازمان سنجش آموزش کشور  
zhasankhani@gmail.com

رژوشگاه علوم انسانی و مطالعات فرهنگی  
پرتال جامع علوم انسانی

### چکیده

توسعه فناوری‌های اطلاعاتی و ارتباطی و پدید آمدن شبکه‌های اطلاعاتی مختلف، منجر به رشد تصاعدی حجم منابع اطلاعاتی الکترونیکی شده و ایجاد راهکارهای جدیدی برای ذخیره و بازیابی سریع و مؤثر اطلاعات را الزامی نموده است. ابر داده‌ها به عنوان یکی از راهکارهای اساسی، ابزار مناسبی برای مدیریت و کنترل منابع الکترونیکی هستند. مقاله حاضر به تشریح مفهوم ابر داده، تاریخچه، انواع و کاربرد آنها در سازماندهی اطلاعات می‌پردازد.  
**کلیدواژه‌ها:** ابر داده، طرح‌های ابر داده‌ای، سازماندهی اطلاعات

### مقدمه

در عصر حاضر که فناوری اطلاعات و ارتباطات، دگرگونی‌های گریزناپذیر و چشمگیری در تمام حرفه‌ها بوجود آورده است، علوم کتابداری و اطلاع‌رسانی نیز همانند سایر علوم، درگیر این تغییر و تحولات گردیده و از پیشرفت‌های ناشی از آن مصون نمانده است. رشد فناوری و پدید آمدن شبکه‌های اطلاعاتی مختلف، منجر به گسترش فزاینده انواع منابع اطلاعاتی در قالب‌های الکترونیکی

گردیده و از این روی، چالش‌ها و نیز عرصه‌های نوینی را فراروی دست‌اندرکاران و متخصصان امر اطلاع‌رسانی قرار داده است. وب جهان‌گستر به عنوان یکی از ابزارهای اینترنت که قابلیت‌ها، کارکردها و روش‌های جدید انتشار و اشاعه اطلاعات را داراست، بر فرایندهای مختلف تولید، سازماندهی و اشاعه اطلاعات نیز تأثیر گذارده است که این امر منجر به افزایش منابع اطلاعاتی متنی و غیر متنی در شبکه جهانی اینترنت گردیده (شیری، ۱۳۷۸) و تغییرات شگرفی را در حوزه ذخیره و بازیابی اطلاعات بوجود آورده است.

رشد و گسترش روزافزون انواع منابع اطلاعاتی ساختار نیافته و ضعف و ناکارآمدی موتورهای جستجوی اینترنتی در سازماندهی این منابع، مشکلاتی را در دسترسی به منابع مرتبط و ارزشمند فراروی جستجوگران اطلاعات قرار داده است، (محمدی، ۱۳۸۳) به گونه‌ای که حتی قوی‌ترین موتورهای جستجو نیز قادر به نمایه‌سازی تمام وب نبوده و در پاسخگویی به این مشکل ناتوان می‌باشند. برای سازماندهی و دسترس پذیرسازی بهینه منابع اینترنتی و تسهیل فرایند جستجو و بازیابی اطلاعات در اینترنت اقدام‌های

سازماندهی منابع اطلاعاتی الکترونیکی با روش‌های متداول و سنتی فهرست‌نویسی و سازماندهی اطلاعات که صرفاً در کتابخانه‌ها و مراکز اطلاع‌رسانی مورد استفاده قرار می‌گیرد، به کار می‌رود. اطلاعات کتابشناختی استاندارد، اصطلاحات نمایه‌ای و چکیده‌ها که توسط کتابخانه‌ها تهیه می‌شوند، نوعی ابر داده هستند. همانطور که فهرست‌ها در کتابخانه‌ها برای جستجو، سازماندهی و بازیابی منابع کتابی و غیره مورد استفاده قرار می‌گیرند، ابر داده همچون فهرستی برای جستجو، سازماندهی و بازیابی منابع الکترونیکی موجود در شبکه اینترنت مورد استفاده قرار می‌گیرد (شیری، ۱۳۷۸، ص ۱۲۶)، در واقع این واژه در محیط الکترونیکی کاربرد دارد. ابر داده‌ها، داده‌هایی ساختارمند درباره داده‌های دیگر می‌باشند که ماشین‌فهم و

متعددی انجام شده و طرح‌ها و پژوهش‌های بیشماری در این زمینه صورت گرفته است. چرا که استفاده کنندگان اطلاعات نیاز دارند به اطلاعات دقیق و مرتبط دسترسی یابند و وجود راهکارها و ابزارهایی برای ذخیره و بازیابی اطلاعات اینترنتی و الکترونیکی ضروری است.

ابرداده (فراداده) در واقع یکی از ابزارهای کمکی برای نمایه‌سازی و ارائه اطلاعات مرتبط با موضوع مورد جستجو در اینترنت هستند که دست اندرکاران امر، با ابداع آن توانستند تا حدودی بر اداره و سازماندهی منابع اینترنتی تسلط یابند. به بیان دیگر، ابر داده به عنوان روشی مؤثر برای ردیابی منابع گسترده اینترنتی شناخته شده است. (ثورنلی، ۱۳۸۱)



ماشین‌خوان نیز هستند.

بطور کلی می‌توان گفت ابر داده اطلاعاتی را برای توصیف خصوصیات محمول‌های اطلاعاتی، سامان‌دهی و رمزگذاری می‌کند تا به شناسایی، کشف، سنجش و مدیریت محمول‌های توصیف شده کمک کند. (ثورنلی، ۱۳۸۱)

#### تاریخچه

بررسی تاریخی مفهوم ابر داده نشان می‌دهد این اصطلاح که اکنون جزء تفکیک‌ناپذیری از دنیای وب گردیده است، پیش از ابداع شبکه جهانی اینترنت و فراگیر شدن وب بکار رفته است. واژه ابر داده اولین بار توسط یک شرکت تجاری آمریکایی برای ایجاد و توسعه محصولات مبتنی بر ابرالگوها<sup>۲</sup> و توصیف مجموعه داده‌ها توسط جک مایرز و کول جیان در سال ۱۹۶۹ بکار رفت. کاربرد این اصطلاح بطور عمده به دهه ۱۹۹۰ و شکل‌گیری شبکه وب در سال ۱۹۹۳ بازمی‌گردد.

دمپسی و هیبری<sup>۳</sup> در یک طرح تحقیقاتی گسترده ۲۲ قالب ابر داده‌ای رایج و مهم را بررسی کردند و در گزارش تحقیق خود به انواع قالب‌های ابر داده‌ای شکل گرفته در دهه‌های ۱۹۶۰، ۱۹۷۰

#### ابرداده چیست؟

واژه ابر داده یا metadata از ریشه یونانی meta به معنی ماهیت و data به معنی اطلاعات یا داده گرفته شده است. (محمدی، ۱۳۸۳)

ابرداده‌ها هر نوع اطلاعات یا داده‌های ساختار یافته هستند که به توصیف، توضیح و یا به تعیین محل منبع اطلاعاتی می‌پردازند. اطلاعات در یک محیط دیجیتالی «داده» و اطلاعات مربوط به آن داده «ابرداده» نامیده می‌شود. اصطلاح «ابر» نیز بیانگر چیزهایی است که بیش از ماهیت خود هستند.

ساده‌ترین تعریفی که در انواع متون از ابر داده ارائه شده عبارت است از «ابرداده اطلاعاتی درباره منابع اطلاعاتی» (شیری، ۱۳۷۸) یا «داده درباره داده». یعنی ابر داده‌ها، داده‌هایی هستند که به تشریح برنامه‌ها، پیام‌ها، فایل‌ها، منابع وب و سایر اشیای داده‌ای می‌پردازند. (اودل<sup>۱</sup>، ۲۰۰۵، ص ۳۳)

با توجه به تعاریف فوق می‌توان گفت که ابر داده روش و ابزاری برای فهرست‌نویسی، سازماندهی و بازیابی منابع الکترونیکی در شبکه اینترنت است و این اصطلاح برای ایجاد تمایز میان روش‌های جدید



اشتراک، یکپارچه‌سازی و استفاده مجدد از انواع اطلاعات را در محیط شبکه‌ای توزیع شده فراهم می‌سازد. (شمیری، ۱۳۷۸، ص ۱۲۹)

کشف منبع: فرایندی است که در آن از بر داده‌ها برای یافتن وجود، مکان و خصوصیات یک منبع خاص استفاده می‌شود. کشف منبع، کاوش، بازیابی و دسترسی به منابع را پشتیبانی می‌کند و روشی ارائه می‌دهد که شناسایی و مکان‌یابی منابعی که مناسب استفاده‌کننده است را میسر می‌سازد. (صفری، ۱۳۸۲، ص ۷۷)

واضح است که بر داده‌ها علاوه بر یافتن منبع کاربردهای دیگری نیز دارند. بر داده‌ها می‌توانند به مدیریت و بازیابی منابع بر اساس نیازهای فناورانه استفاده‌کنندگان یاری رسانده و نیز در پالایش اطلاعات از طریق خدمات نرخ‌گذاری، مدیریت دسترسی به منابع بر اساس قوانین حق مؤلف، حفاظت و نگهداری اطلاعات دیجیتالی و تأیید اصالت منابع کمک نمایند. (صفری، ۱۳۸۲، ص ۷۸)

### انواع طرح‌های بر داده‌ای

انواع طرح‌های بر داده‌ای برای مخاطبان متفاوت در حال ایجاد و گسترش هستند و دسته‌بندی‌های متفاوتی از این طرح‌ها ارائه شده است. از جمله می‌توان به بر داده‌های توصیفی<sup>۵</sup>، مدیریتی<sup>۶</sup> و ساختاری<sup>۷</sup> اشاره نمود.

بر داده توصیفی به توصیف منبع، به منظور کشف و شناسایی می‌پردازد و عنوان، پدیدآور، چکیده و کلیدواژه از عناصر مهم این نوع بر داده است.

بر داده مدیریتی با ارائه اطلاعاتی از زمان، چگونگی ایجاد فایل، اطلاعات مربوط به فراهم‌آوری، حقوق، مکان‌یابی و غیره امکان مدیریت منبع را فراهم می‌کند.

بر داده ساختاری چگونگی ترکیب اشیاء اطلاعاتی را مشخص می‌نماید.

با توجه به دلایل و انگیزه‌های مختلف شکل‌گیری انواع قالب‌های بر داده‌ای برای طیف وسیعی از منابع مختلف اطلاعاتی الکترونیکی موجود در اینترنت، استفاده از الگویی که بتواند از انسجام و جامعیت بیشتری برخوردار باشد و سهولت تولید رکوردهای بر داده‌ای را تضمین کند از اهمیت خاصی برخوردار است. (شمیری، ۱۳۷۸، ص ۱۳۰) به همین جهت تاکنون برای بر داده‌ها، قالب‌ها و استانداردهای گوناگونی بوجود آمده که از مهمترین آنها می‌توان قالب بر داده‌ای دوبلین کور<sup>۸</sup>، قالب مارک، قالب یو. اس. سمارک، قالب بر داده‌ای طرح کدگذاری متن را نام برد.

و ۱۹۸۰ به عنوان اولین قالب‌های بر داده‌ای اشاره نمودند. (شمیری، ۱۳۷۸، ص ۱۲۷)

علاوه بر استفاده از فهرست چاپی، کتابخانه‌ها از دهه ۱۹۶۰ استفاده از فهرست‌های ماشین‌خوان را آغاز کرده‌اند و آنها را بعنوان استاندارد برای اشتراک اطلاعات بین کتابخانه‌ها، شناسایی، توصیف، دسترسی موضوعی به منابع و همچنین ذخیره‌سازی و بازیابی به کار برده‌اند.

این فهرست‌ها را می‌توان نوعی قالب بر داده‌ای دانست که قبل از ظهور وب خلق شدند. از دیگر قالب‌های بر داده‌ای قبل از ظهور وب، می‌توان به طرح کدگذاری متن<sup>۹</sup> و توصیف کدگذاری شده منابع آرشیوی اشاره نمود. (صفری، ۱۳۸۲، ص ۷۶)

### اهمیت بر داده‌ها

بر داده روشی نظام‌مند است که منابع اطلاعاتی را برای کاربران دسترسی‌پذیر و قابل فهم می‌سازد. (محمدی، ۱۳۸۳). منابع موجود در شبکه جهانی وب، دارای ویژگی‌های متمایز و متفاوتی از منابع اطلاعاتی سنتی هستند.

سازماندهی مؤثر و کارآمد اطلاعات در محیط‌های دیجیتالی به‌ویژه محیط وب متکی به مدیریت و سازماندهی مؤثر و کارآمد بر داده‌ها شده است، بطوری که می‌توان گفت که در عصر اینترنت شاید هیچ داده‌ای ارزشمندتر از داده دربار داده نباشد.

طیف وسیعی از گروه‌های مختلف اعم از دانشمندان، محققان، ناشران، نهادهای دولتی، عرضه‌کنندگان تجاری، مریبان، محققان، آمارگران، متخصصان و ... مایلند برای جستجو و بازیابی اطلاعات سریع و دقیق منابع اینترنتی از الگوها و قالب‌های استاندارد استفاده کنند. (شمیری، ۱۳۷۸)

برای تولید و افزایش امکان ذخیره‌سازی دیجیتالی اطلاعات و نیز دسترسی نامحدود به اطلاعات در طول دوره زمانی آنها از بر داده‌ها استفاده می‌شود.

بر داده‌ها ایجاد، ثبت، رده‌بندی، دسترسی، نگهداری و نظم و ترتیب رکوردها را در زمان و مکان خاص خود، میسر می‌سازند.

از بر داده‌ها می‌توان برای تعیین خط‌مشی‌ها و اداره سیستم‌های مختلف بهره گرفت. سازمان‌های عمومی و اختصاصی می‌دانند که بر داده‌ها برای مدیریت اطلاعات تجاری حساس، بسیار کارایی داشته و بدین منظور در کانون سیستم مدیریت قرار دارند. (فرانکز و کوند، ۲۰۰۶، ص ۵۵)

از همه مهم‌تر اینکه فهرست‌نویسان اینترنتی، کتابداران، مدیران و طراحان شبکه، مدیران دانش و هر کس که منابع اطلاعاتی را در یک شبکه و یا روی اینترنت مدیریت می‌کند، باید بر داده را بشناسد.

چرا که همه دلایل مربوط به علت اهمیت نمایه‌سازی و فهرست‌نویسی برای مواد چاپی با شدت بیشتر درباره اهمیت و ضرورت بر داده برای منابع و اطلاعات موجود در وب نیز صادق هستند. (صفری، ۱۳۸۲، ص ۱۷۷)

### کارکرد بر داده‌ها

در نوشته‌های معتبر توصیف منبع و کشف منبع را دو کارکرد مهم بر داده‌ها بر شمرده‌اند.

توصیف منبع: بر داده‌ها با بیان ویژگی‌های منابع، به نحوی ساختارمند به توصیف آنها می‌پردازند. (صفری) از آنجا که منابع اطلاعاتی به اشکال مختلف و با ویژگی‌های متفاوت در پایگاه‌های اطلاعاتی گوناگون وجود دارد، توصیف بر داده‌ای امکان تطبیق،



برخی متخصصان معتقدند که کتابخانه‌ها می‌توانند از ابر داده برای ایجاد رکوردهای فهرست‌نویسی پیچیده‌تر و نیز توصیف پایگاه‌های اطلاعاتی و تصاویر دیجیتالی، فایل‌های صوتی و سایر منابع غیر متنی استفاده کنند. (شیری، ۱۳۷۸، ص ۱۳۰)

در آخر می‌توان گفت با ظهور محیط‌های الکترونیکی جدید و منابع الکترونیکی مختلف به الگوها، روشها، استانداردها و ابزارهای جدیدی برای ذخیره و بازیابی آنها نیاز است و در راستای رفع این نیاز، ابر داده‌ها به‌عنوان کلیدی برای فراهم‌آوری، دسترسی آسان و به اشتراک گذاشتن داده‌ها ابداع گردیده‌اند. به عبارت دیگر ابر داده‌ها، ابزار مناسبی برای مقابله با پدیده انفجار دانش و آلودگی اطلاعات می‌باشند.

### کاربرد ابر داده‌ها در سازماندهی اطلاعات

اصطلاح ابر داده اصطلاح جدیدی نیست، کتابداران سال‌هاست که به فهرست‌نویسی کتاب‌ها و مجلات می‌پردازند. فهرست کتابخانه خود نوعی ابر داده است که برای یافتن کتاب‌ها و مجلات و زمینه‌های موضوعی خاص و بازیابی آنها از قفسه کتابخانه مورد استفاده قرار می‌گیرد.

امروزه ابر داده در سازماندهی منابع الکترونیکی بویژه منابع وب، نقش و اهمیت برجسته‌ای یافته است (صفری، ۱۳۸۲، ص ۷۴) و با شکل‌گیری الگوهای ابر داده‌ای نیاز به استفاده از فهرست‌های پیوسته همگانی بعنوان گذرگاهی برای دسترسی به مجموعه‌های ابر داده‌ای در کتابخانه‌ها امری ضروری به نظر می‌رسد. نظام‌های مبدل ابر داده‌ای، کتابخانه‌ها را قادر به استخراج ابر داده‌های خارج از فهرست پیوسته در محیط اینترنت و تبدیل آن به قالب‌هایی چون مارک و نیز یکپارچه‌سازی آن با مجموعه پایگاه‌های اطلاعاتی کتابخانه می‌سازد.

الگوهای ابر داده می‌توانند دو کاربرد عمده در کتابخانه‌ها داشته باشند: اول آنکه کتابخانه‌ها پایگاه‌های اطلاعاتی خود را به عنوان بخشی از خدمات کاوش شبکه‌های گسترده‌تر در دسترس قرار دهند و این امکان را فراهم سازند تا استفاده‌کنندگان کتابخانه خود دیدگاهی دوجانبه نسبت به پایگاه‌های اطلاعاتی کتابخانه‌ای و غیر کتابخانه‌ای کسب کنند.

دوم آنکه کتابخانه‌ها منابع جدید را که از طریق روشهای فهرست‌نویسی سنتی هزینه‌بر و زمان‌بر است، از طریق ابر داده توصیف نمایند.

### پس‌گوشه

۱. Udell
۲. Metamodels
۳. Dempsey & Heery
۴. Text Encoding Initiative (TEI)
۵. Descriptive
۶. Administrative

۷. Structural
۸. Dublin Core

### منابع و مآخذ

- ۱- ثورنلی، جنی (۱۳۸۱). چگونگی ابر داده: ایجاد و استانداردها. ترجمه محسن حاجی زین‌العابدینی. فصلنامه کتاب. دوره سیزدهم، ۳ (پاییز): ۱۳۴.
- ۲- دیگان، ماریلین، تاتر، سیمون (۱۳۸۲). آینده دیجیتالی کتابخانه‌ها: راهبردهایی برای عصر اطلاعات، ترجمه عباس گیلوری. تهران: دبیزش، چاپار.
- ۳- شیری، علی اصغر (۱۳۷۹). ابر داده‌ها و تاثیر آن بر فهرست‌های ماشین‌خوان: الگوی فارسی ابر داده برای سازماندهی اطلاعات الکترونیکی فارسی در فهرست‌های رایانه‌ای: کاربرد و توسعه: مجموعه مقالات همایش کاربرد و توسعه فهرست‌های رایانه‌ای در کتابخانه‌های ایران، ۲۷ و ۲۸ آبان ماه ۱۳۷۸، دانشکده علوم تربیتی و روانشناسی دانشگاه فردوسی مشهد. به کوشش رحمت‌الله فتحی. تهران: مشهد: دانشگاه فردوسی مشهد؛ مرکز اطلاع‌رسانی و خدمات علمی وزارت جهاد سازندگی: ۱۲۴ - ۱۵۱.
- ۴- صفری، مهدی (۱۳۸۲). ابر داده و کاربرد آن در سازماندهی منابع وب با تاکید بر عناصر ابر داده‌ای دوبلین کور در بستر نحوی HTML. فصلنامه کتاب، دوره چهاردهم، ۳ (پاییز): ۷۳ - ۸۵.
- ۵- محمدی، حسین (۱۳۸۳). ابر داده: مفاهیم و کاربردها. مجله الکترونیکی مرکز اطلاعات و مدارک علمی ایران. سال دوم، شماره ۳.
- 6- Udell, Jon (2005). Managing metadata. InfoWorld. San Mateo: Oct 24, Vol 27, Iss. 43: 33 - 37. Available at <http://proquest.umi.com>