



وزارت علوم، تحقیقات و فناوری
پروژه‌های علوم و فناوری اطلاعات ایران

طرح پژوهشی

**بررسی یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های
دیجیتال و ارائه الگوی پیشنهادی برای کتابخانه‌های دیجیتال
ایران**

مجری

مهدی علیپور حافظی

فروردین ۱۳۹۳

حقوق: پژوهشگاه علوم و فناوری اطلاعات ایران

طرح پژوهشی «بررسی یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال و ارائه الگوی پیشنهادی برای کتابخانه‌های دیجیتال ایران» پیرو ابلاغیه شماره ۳۳/۶۰۴۵ مورخ ۱۳۹۰/۳/۱ میان پژوهشگاه علوم و فناوری اطلاعات ایران (کارفرما) و (مجری) اجرا شده است. گزارش حاضر گزارش نهایی این طرح است.

این گزارش و تمامی حقوق مادی آن بر اساس «قانون حمایت حقوق مؤلفان و مصنفان و هنرمندان، مصوب سال ۱۳۴۸ و اصلاحیه‌های بعدی آن و همچنین آیین‌نامه‌های اجرایی این قانون متعلق به پژوهشگاه علوم و فناوری اطلاعات ایران است و هر گونه استفاده از تمامی یا پاره‌ای از آن، شامل: نقل قول، تکثیر، انتشار، کاربرد نتایج، تکمیل و مانند آنها به صورت چاپی، الکترونیکی یا وسایل دیگر فقط با اجازه کتبی پژوهشگاه امکان‌پذیر است. نقل قول در حد هزار واژه در انتشارات علمی مانند کتاب و مقاله با درج اطلاعات کامل کتاب‌شناختی، نیازی به مجوز پژوهشگاه ندارد. صحت مندرجات گزارش بر عهده مجری طرح پژوهشی است.

به نام آنکه جان را فکرت آموخت

چکیده

هدف اصلی پژوهش حاضر ارائه الگوی مناسب پیشنهادی برای یکپارچه‌سازی معنایی کتابخانه‌های دیجیتال ایران است. در این راستا و برای وصول به هدف اصلی ذکر شده شش پرسش اساسی طرح شد. در گام‌های مطالعاتی پژوهش حاضر از روش پژوهش کتابخانه‌ای استفاده شده است. در مراحل بعدی برای شناسایی کیفیت و نوع محتوای کتابشناختی کتابخانه‌های دیجیتال ایران از روش پیمایش تحلیلی استفاده شده است. در این روش از ابزارهای معمول آن برای گردآوری و تحلیل یافته‌ها استفاده شده است. در مرحله سوم برای ارائه الگوی پیشنهادی پژوهش نیز از روش تحلیل سیستم استفاده شده است. جامعه پژوهش حاضر را در مرحله شناسایی وضعیت موجود محتواهای فراداده کتابخانه‌های دیجیتال مورد بررسی مسئولان کتابخانه‌های دیجیتال (۲۶ کتابخانه دیجیتالی) تشکیل دادند. جهت گردآوری اطلاعات در این زمینه از ابزار پرسشنامه محقق ساخته استفاده شد. برای گردآوری اطلاعات در زمینه وضعیت خروجی فراداده در سیستم‌های اطلاعاتی کتابخانه دیجیتال نیز با توجه به شناسایی ۱۱ نرم‌افزار، جامعه پژوهش شامل این نرم‌افزارها شد. جهت گردآوری اطلاعات از این نرم‌افزارها از سیاهه واری و مشاهده مستقیم توسط پژوهشگر استفاده شد.

نتایج پژوهش نشان داد که کتابخانه‌های دیجیتالی مورد بررسی از استانداردهای فراداده متنوعی استفاده می‌نمایند. همچنین خروجی‌های فراداده استاندارد مورد استفاده نیز دارای تنوع دسترسی به آنها هستند. علاوه بر این مدل ارائه داده‌ها نیز در برخی موارد مدل برداشت اطلاعات و در برخی موارد دیگر مدل جستجوی هم زمان مورد استفاده می‌تواند قرار گیرد. به لحاظ تکمیل فیلدهای محتوایی نیز نتایج این پژوهش نشان داد که وضعیت این فیلدها در حال حاضر مطلوب نیست و پاسخگوی نیازهای آتی جهت برقراری روابط معنایی نیست و باید بیشتر به این موارد فکر شود. در نهایت با توجه به قابلیت‌ها، امکانات و وضعیت موجود مدل ترکیبی یکپارچه‌سازی معنایی در این پژوهش برای یکپارچه‌سازی معنایی کتابخانه‌های دیجیتالی مورد بررسی پیشنهاد شد.

کلیدواژه‌ها: کتابخانه‌های دیجیتالی، یکپارچه سازی، مبادله معنایی، تعامل معنایی، یکپارچه‌سازی معنایی، ایران، مدل‌های یکپارچه‌سازی، مدل یکپارچه سازی معنایی

فهرست مندرجات

چکیده	۴
فهرست مندرجات	۵
فهرست جدول‌ها	۷
فهرست شکل‌ها	۸
فصل ۱	۹
۱-۱. مقدمه	۱۰
۱-۲. بیان مسئله	۱۱
۱-۳. تعریف عملیاتی اجزاء مسئله	۱۲
۱-۴. پرسش‌های اساسی	۱۳
۱-۵. هدف و فایده	۱۳
۱-۶. متغیرهای اساسی پژوهش	۱۴
۱-۷. جنبه‌های نو بودن پژوهش	۱۴
۱-۸. محدودیت‌های پژوهش	۱۵
فصل ۲. مبانی نظری و ادبیات پژوهش	۱۶
۲-۱. مقدمه	۱۷
۲-۲. یکپارچه‌سازی محتوا و خدمات در کتابخانه‌های دیجیتال	۱۸
۲-۳. معنا در بازایی اطلاعات	۲۱
۲-۴. یکپارچه‌سازی معنایی و روش‌های انجام آن	۲۴
۲-۵. وب معنایی	۲۶
۲-۶. ابزارهای یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال	۲۷
۲-۶-۱. هستی‌نگاری‌ها	۲۸
۲-۶-۲. سیستم‌های سازماندهی دانش (کی.او.اس.)	۳۰
۲-۷. زبان‌های مورد استفاده در یکپارچه‌سازی معنایی	۳۳
۲-۷-۱. زبان توصیفی ایکس.ام.ال.	۳۴
۲-۷-۲. آر.دی.اف.	۳۷
۲-۷-۳. دَمَل +اوایل	۳۹
۲-۷-۴. او.دبلیو.ال.	۴۱
۲-۸. نمونه‌های یکپارچه‌سازی معنایی	۴۲
۲-۸-۱. پروژه دلاس	۴۲
۲-۸-۲. پروژه تی.اس.آی.ام.آی.اس.	۴۴
۲-۸-۳. پروژه یورپانا	۴۵
۲-۸-۴. ان.دی.ال.تی.دی.	۴۶
۲-۸-۵. پروژه ان.اس.دی.ال.	۴۶

۲-۹. نتیجه گیری.....	۴۷
۲-۱۰. پیشینه پژوهش.....	۴۸
۲-۱۰-۱. پیشینه پژوهش در داخل.....	۵۰
۲-۱۰-۲. پیشینه‌های خارجی.....	۵۳
۲-۱۰-۳. تحلیل پیشینه‌های پژوهش.....	۵۸
فصل ۳. روش‌شناسی پژوهش.....	۶۰
۳-۱. نوع و روش پژوهش.....	۶۱
۳-۲. شیوه و ابزار گردآوری اطلاعات.....	۶۱
۳-۳. روایی و پایایی ابزار پژوهش.....	۶۲
۳-۴. جامعه پژوهش.....	۶۲
۳-۵. روش تجزیه و تحلیل یافته‌ها.....	۶۴
فصل ۴. توصیف یافته‌های پژوهش.....	۶۵
۴-۱. مقدمه.....	۶۶
۴-۲. توصیف یافته‌های پژوهش.....	۶۶
۴-۲-۱. توصیف یافته‌های مربوط به وضعیت موجود محتواهای پیشینه‌های کتابشناختی.....	۶۶
۴-۲-۲. توصیف یافته‌های مربوط به وضعیت موجود خروجی‌های فراداده نرم‌افزارهای کتابخانه‌های دیجیتال.....	۷۵
فصل ۵. تحلیل یافته‌های پژوهش.....	۸۳
۵-۱. مقدمه.....	۸۴
۵-۲. تحلیل یافته‌ها.....	۸۴
۳-۵. اعتبارسنجی مدل پیشنهادی.....	۹۵
۵-۴. پیشنهادهای پژوهش.....	۹۳
فهرست منابع.....	۹۴
پیوست‌ها.....	۱۰۱
پیوست ۱. پرسشنامه بررسی وضعیت موجود محتواهای پیشینه‌های کتابشناختی کتابخانه‌های دیجیتال ایران.....	۱۰۲
پیوست ۲. سیاهه واریاسی ارزیابی قابلیت‌های خروجی نرم‌افزارهای کتابخانه دیجیتال.....	۱۰۷
پیوست ۳. نمونه فایل ایکس.ام.ال. دریافتی از کتابخانه دیجیتال مجلس.....	۱۰۹

فهرست جدول‌ها

جدول ۱-۲. عمده‌ترین برچسب‌های اسکاس و توضیحات آنها.....	۳۳
جدول ۲-۲. اجزاء تشکیل دهنده دَمَل +اوایل.....	۴۰
جدول ۳-۲. قواعد دَمَل +اوایل.....	۴۱
جدول ۱-۴. انواع منابع اطلاعاتی موجود در کتابخانه‌های دیجیتال ایران.....	۶۶
جدول ۲-۴. تکمیل فیلدهای کتابشناختی توسط کتابخانه‌های دیجیتال مورد بررسی.....	۶۷
جدول ۳-۴. مستند سازی اسامی پدیدآورندگان و ناشران.....	۶۸
جدول ۴-۴. روش‌های تحلیل موضوعی منابع اطلاعاتی توسط کتابخانه‌های دیجیتال مورد بررسی.....	۶۸
جدول ۵-۴. انتخاب موضوع برای تحلیل منابع اطلاعاتی توسط کتابخانه‌های دیجیتال مورد بررسی.....	۶۹
جدول ۶-۴. انتخاب کلیدواژه برای تحلیل منابع اطلاعاتی توسط کتابخانه‌های دیجیتال مورد بررسی.....	۶۹
جدول ۷-۴. منابع مورد استفاده در فهرست‌نویسی بنیادی توسط کتابخانه‌های دیجیتال مورد بررسی.....	۷۳
جدول ۸-۴. اصطلاحنامه‌های مورد استفاده جهت نمایه‌سازی کنترل شده.....	۷۴
جدول ۹-۴. اصطلاحنامه‌های لاتین مورد استفاده جهت نمایه‌سازی کنترل شده.....	۷۴
جدول ۱۰-۴. خروجی‌های فراداده‌ای نرم‌افزارهای مورد بررسی.....	۷۶
جدول ۱۱-۴. نحوه ارتباط فراداده با فایل منبع.....	۷۷
جدول ۱۲-۴. نمایه‌سازی و دسترسی به فراداده توسط موتورهای جستجوی عمومی.....	۷۸
جدول ۱۳-۴. وجود امکان مبادله اطلاعات.....	۷۸
جدول ۱۴-۴. استفاده از تفاهم‌نامه جهت مبادله اطلاعات.....	۷۹
جدول ۱۵-۴. فرمت‌های مورد استفاده برای مبادله اطلاعات.....	۸۰

فهرست شکل‌ها

- شکل ۱-۲. یکپارچه‌سازی کتابخانه‌های دیجیتال..... ۱۸
- شکل ۲-۲. الگوی نشانه‌شناسی پیرس..... ۲۱
- شکل ۲-۳. مثلث معنایی آگدن و ریچاردز..... ۲۱
- شکل ۲-۴. شبکه تولید و استفاده از محتوا..... ۲۲
- شکل ۲-۵. عوامل دخیل در یکپارچه‌سازی نحوی و روابط آنها با هم..... ۲۵
- شکل ۲-۶. یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتالی..... ۲۶
- شکل ۲-۷. ساختار سلسله مراتبی ایکس.ام.ال..... ۳۵
- شکل ۲-۸. ساختار درختی ایکس.ام.ال..... ۳۶
- شکل ۲-۹. دی.تی.دی. یک فایل ایکس.ام.ال..... ۳۶
- شکل ۲-۱۰. مجموعه عناصر مورد استفاده در استفاده از ایکس.ام.ال..... ۳۷
- شکل ۲-۱۱. ساختار جملات آر.دی.اف..... ۳۸
- شکل ۲-۱۲. مدل مرجع کتابخانه دیجیتال دلاس..... ۴۳
- شکل ۲-۱۳. معماری پروژه تی.اس.آی.ام.آی.اس..... ۴۵
- شکل ۲-۱۴. مولفه‌های معماری ان.اس.دی.ال..... ۴۷
- شکل ۴-۱. میزان بهره‌گیری از روش‌های فهرست‌نویسی بنیادی و نسخه‌برداری توسط کتابخانه‌های..... ۷۱
- شکل ۴-۲. نمودار پایگاه‌های اطلاعاتی مورد استفاده برای نسخه‌برداری جهت توصیف و تحلیل..... ۷۱
- شکل ۴-۳. نمودار پایگاه‌های اطلاعاتی غیر فارسی مورد استفاده برای نسخه‌برداری جهت توصیف و تحلیل..... ۷۲
- شکل ۴-۴. نمودار میزان اصلاح فیلدهای اطلاعاتی توسط کتابخانه‌های دیجیتال مورد بررسی..... ۷۲
- شکل ۴-۵. ابزارهای مورد استفاده در نمایه‌سازی کنترل شده در کتابخانه‌های دیجیتال مورد بررسی..... ۷۳
- شکل ۵-۱. روش استفاده از مترجم‌ها و ماشین استنتاج..... ۸۳
- شکل ۵-۲. معماری مبتنی بر ثبت پیشینه جستجو..... ۸۴
- ۳-۵. الگوی پیشنهادی یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال ایران..... ۸۹

فصل ۱

معرفی پژوهش

۱-۱. مقدمه

دستیابی سریع و یکپارچه به اطلاعات مناسب از جمله آرزوهای انسان حاضر در قرن بیست و یک است. کتابخانه‌ها از نخستین گام‌های پیدایش تا به امروز در پی تحقق چنین آرزویی هستند و از ابزارها و فناوری‌های روز در هر عصری برای وصول به این هدف بزرگ استفاده کرده و می‌کنند. قوانین پنج‌گانه رانگانان نیز برای پوشش همین آرزو مطرح شده‌اند (مزینانی، ۱۳۸۲). وب و ابزارهای جستجو و بازیابی اطلاعات وب بستری را فراهم کرده‌اند که افراد برای دستیابی به اطلاعات مورد نیاز خود به آن مراجعه و اقدام به ایجاد دستیابی به محتواهای مورد نیاز خود می‌کنند. با توجه به حجم عظیم اطلاعات موجود در محیط وب و ضعف‌های ناشی از نقصان ابزارهای جستجو و بازیابی و مهمتر از همه مسائل مرتبط با ذخیره و سازماندهی اطلاعات همواره موضوع دستیابی سریع و یکپارچه به محتوای وب را نیز با چالش همراه ساخته‌اند (ساریان قلی، ۱۳۸۴).

محیط یکپارچه وب و ابزارهای جستجو و بازیابی اطلاعات در محیط وب تأثیر مستقیمی بر رفتار اطلاع‌یابی کاربران گذاشته است. به نحوی که کاربران برای دستیابی به اطلاعات مورد نیاز خود به عنوان اولین منبع اطلاعاتی، به وب مراجعه می‌کنند. در صورتی که همین امر در مورد کتابخانه‌ها، پیش از ظهور وب صادق بود. همین امر گویای ظهور رقیب قدرتمندی برای کتابخانه‌ها به عنوان مراجع اصلی ذخیره و بازیابی اطلاعات بود. این رویداد عظیم و تأثیرگذار در فعالیت‌های روزمره انسان‌ها، وب، در روند توسعه کتابخانه‌ها نیز تأثیر گذار شد. از این رو، کتابخانه‌ها اقدام به استفاده از تهدید به عمل آمده از طرف وب و تبدیل آن به فرصت کردند. بنا براین کتابخانه‌ها با حضور در محیط وب اقدام به ارائه اطلاعات کتابشناختی خود کردند و این اطلاعات را در دسترس عموم کاربران قرار دادند. در این راستا کاربران نیز با حضور در محیط وب قادر به دستیابی به اطلاعات کتابشناختی منابع کتابخانه-ای شدند. به تدریج کتابخانه‌ها با الگو گرفتن از وب و استفاده از قابلیت‌های به عمل آمده اقدام به ارائه اشیاء دیجیتال^۱ به کاربران و دسترس‌پذیر ساختن آنها برای همگان در محیط وب کردند. همین امر باعث ظهور نسل جدیدی از کتابخانه‌ها با عنوان کتابخانه-های دو وجهی^۲، کتابخانه‌های دیجیتال و کتابخانه‌های مجازی در فضای پیوسته وب شد. بنا براین کاربران با حضور در محیط وب و اتصال به این کتابخانه‌ها قادر به دستیابی به محتوای منابع اطلاعاتی شدند.

هرچند این حرکت در نوع خود مفید و حائز اهمیت فراوانی بود، ولی همچنان کاربران از عملکرد کتابخانه‌ها خوشنود نبودند. کاربرانی که برای یافتن اطلاعات مورد نیاز خود با استفاده از ابزارهای جستجو و بازیابی اطلاعات در محیط وب و به صورت یکپارچه اقدام می‌کردند، علاقه‌مند به جستجو در کتابخانه‌های دیجیتال به صورت منفرد نبودند. به بیانی دیگر، کاربران در محیط وب با یک‌بار جستجو قادر به بازیابی تمامی موارد مرتبط بودند. این مسئله انتظارات افراد را از سیستم‌های اطلاعاتی بالا برده است. لذا کاربران خواهان دستیابی یکپارچه به محتواهای اطلاعاتی مورد نیاز خود هستند که کتابخانه‌های دیجیتال بدون ایجاد همکاری متقابل قادر به ارائه چنین سرویسی نیستند. کاربران در استفاده از کتابخانه‌های دیجیتال باید شرایط زیر را احراز می‌کردند:

- شرایط مورد نیاز برای عضویت در کتابخانه‌های دیجیتال در تک تک موارد؛
- اتصال به کتابخانه‌های دیجیتال به صورت منفرد؛
- جستجو و در نهایت بازیابی اطلاعات مورد نیاز خود در تک تک کتابخانه‌های دیجیتال؛

^۱ منظور محتوای منابع اطلاعاتی کتابخانه‌ای می‌باشد.

^۲ Hybrid libraries

حال فرض کنید کاربری برای انجام یک جستجوی ساده ترکیبی در یک کتابخانه چند بار باید جستجو کند. این تعداد را در تعداد کتابخانه‌های دیجیتال مرتبط، پایگاه‌های اطلاعاتی مرتبط، وب و سایر موارد دیگر ضرب کنید. به عنوان مثال جستجو برای مبادله اطلاعات در کتابخانه‌های دیجیتال فقط به زبان انگلیسی نیاز به ده جستجوی ترکیبی دارد. حال اگر این جستجو در ۲۰ پایگاه مختلف مرتبط انجام پذیرد، کاربر مجبور به تکرار جستجو به تعداد ۲۰۰ مورد خواهد شد (Alipour-Hafezi, Horri, Shiri, & Ghaebi, 2010). مثال ذکر شده نشان از میزان اتلاف وقت در جستجو و بازیابی اطلاعات در کتابخانه‌های دیجیتال دارد. در چنین شرایطی کتابخانه‌های دیجیتال بسان جزایری هستند که ارتباطی با سایر کتابخانه‌های دیجیتال و پایگاه‌های علمی دیگر ندارند (علیپور حافظی ۱۳۸۸). سکون بودن در چنین شرایطی باعث از دست دادن کاربران کتابخانه‌ها حتی کتابخانه‌های دیجیتال خواهد شد. از این رو، کتابخانه‌ها باید به دنبال راهکارهایی باشند تا آنها را از شرایط فعلی مذکور رها سازند.

۲-۱. بیان مسئله

یکپارچه‌سازی دستیابی به اطلاعات امروزه جزء ضروری‌ترین نیازهای انسانی در دنیای فناورانه است. محیط وب بستر مورد نیاز برای برقراری ارتباط بین شبکه‌های اطلاعاتی را فراهم کرده است. کتابخانه‌ها نیز که با حضور در محیط وب قابلیت دستیابی به محتواها و خدمات اطلاعاتی را فراهم نموده‌اند باید در پی استفاده از قابلیت‌های محیط وب در ایجاد دستیابی یکپارچه به محتواها و خدمات اطلاعاتی خود باشند. در غیر این صورت جایگاه پژوهشی و اطلاعاتی خود را از دست خواهند داد. از این رو، یکپارچه‌سازی محتوا و خدمات از جمله ضروریات حضور مستمر کتابخانه‌های دیجیتال در عصر حاضر قلمداد می‌شود.

یکپارچه‌سازی سیستم‌های اطلاعاتی به ویژه سیستم‌های اطلاعاتی کتابخانه دیجیتال نیازمند یکپارچه‌سازی در سطح نحوی^۱ و معنایی^۲ است (Ding & Sølberg, 2004); (Shen R. , Applying the 5S Framework To Integrating Digital Libraries, 2006); در یکپارچه‌سازی نحوی که زیربنای یکپارچه‌سازی سیستم‌های اطلاعاتی نیز قلمداد می‌شود، به لحاظ عملیاتی، سیستم‌های اطلاعاتی با یکدیگر مرتبط می‌شوند و اقدام به مبادله اطلاعات با یکدیگر می‌نمایند. در این سطح از یکپارچه‌سازی اطلاعات سه سطح توافقنامه محتوا، فنی و سازمانی مطرح هستند. توافقنامه فنی شامل فرمت‌ها، تفاهم‌نامه‌ها، و سیستم‌های امنیتی است. به این ترتیب محتواهای اطلاعاتی می‌توانند مبادله شوند. توافقنامه محتوا شامل داده و فراداده است که در بر گیرنده توافقنامه‌های معنایی^۳ در تفسیر پیام‌ها است؛ توافقنامه سازمانی در بر گیرنده قواعد زیربنایی برای دسترسی، جایگزینی مجموعه‌ها و خدمات، پرداخت‌ها، شناسایی و غیره است (NSDL technical infrastructure white paper version 2.0, 2004; Shi, 2005). در این سطح از یکپارچه‌سازی بحث‌هایی مانند مدل‌های یکپارچه‌سازی، تفاهم‌نامه‌های مبادله اطلاعات، استانداردهای فراداده مورد استفاده در مبادله اطلاعات، و استانداردها در لایه زمینه در لایه توافقنامه‌های فنی و محتوایی مطرح هستند. مباحث فرهنگی و سازمانی مانند ضرورت توجه به یکپارچه‌سازی از منظر مدیران و تصمیم‌سازان، توجیه کتابداران و کاربران و غیره نیز در لایه توافقنامه سازمانی مطرح هستند (علیپور حافظی، ۱۳۹۰).

¹ Syntactic

² Semantic

³ Semantic agreements

بحث بعدی در یکپارچه‌سازی محتوا و خدمات در کتابخانه‌های دیجیتال پس از یکپارچه‌سازی نحوی، ایجاد روابط معنایی بین منابع اطلاعاتی است که موضوع مورد نظر در یکپارچه‌سازی در سطح معنایی است. هدف از این سطح از تعامل تسهیل در بازیابی معنایی اطلاعات در سیستم‌های اطلاعاتی کتابخانه‌های دیجیتال است تا کاربران به سهولت بتوانند با هر کلیدواژه جستجویی به منابع اطلاعاتی مرتبط با نیاز اطلاعاتی خود دست یابند. مسئله اصلی در این سطح از یکپارچه‌سازی برقراری روابط معنایی بین محتواهای اطلاعاتی است که از سیستم‌های اطلاعاتی مختلف گردآوری شده و امکان دستیابی یکپارچه یافته‌اند. یکی از ویژگی‌های محتواهای گردآوری شده از سیستم‌های اطلاعاتی مختلف این است که محتواهای اطلاعاتی به شیوه‌های مختلف و با استفاده از اصول و استانداردهای مختلفی تهیه و ذخیره شده‌اند. همین امر باعث عدم یکدستی در محتواهای اطلاعاتی مبادله شده می‌شود. از آنجایی که هدف اصلی در یکپارچه‌سازی سیستم‌های اطلاعاتی تسهیل در دستیابی به اطلاعات مرتبط در کمترین زمان و با کمترین تلاش است، یکپارچه‌سازی در سطح معنایی از اهمیت بسیار بالایی برخوردار خواهد شد. بنا براین مسئله اصلی در پژوهش حاضر شناسایی روش‌ها و ابزارهای یکپارچه‌سازی معنایی در سیستم‌های اطلاعاتی کتابخانه دیجیتال و ارائه الگوی پیشنهادی مناسب با توجه به وضعیت موجود سیستم‌های اطلاعاتی کتابخانه دیجیتال مورد استفاده در کتابخانه‌های دیجیتال ایران می‌باشد. در این راستا با توجه به اهمیت استانداردهای مرتبط موجود در این زمینه سعی شده است تا الگوی پیشنهادی به استانداردهای مناسب این حوزه و قابلیت کاربرد آنها در تعامل معنایی توجه داشته و مبتنی بر آن عمل نماید.

۳-۱. تعریف عملیاتی اجزاء مسئله

• **کتابخانه‌های دیجیتال.** برخی محققان بر این باورند که اصطلاح کتابخانه دیجیتال، همواره مقوله‌ای بسیار بحث انگیز و ابهام برانگیز بوده است (Borgman, 2000). به دلیل اینکه کتابخانه دیجیتال مفهومی بین‌رشته‌ای است، متخصصان حوزه‌های موضوعی مختلف نیز بر آن نظر داشته‌اند و بدین ترتیب تعاریفی نیز از دیدگاه خود برای آن ارائه کرده‌اند. از این رو، در این پژوهش به تعاریفی که از دید کتابداران ارائه شده تأکید گردیده است. دیگان و تانر^۱ در کتاب خود با عنوان آینده دیجیتال: راهبردهایی برای عصر اطلاعات^۲، با بررسی تعاریف مختلف، ویژگی‌های زیر را برای یک کتابخانه دیجیتال برشمرده‌اند:

- کتابخانه دیجیتال مجموعه‌ای از اطلاعات سازماندهی شده در قالب دیجیتالی است؛
- اطلاعات دیجیتال باید مبتنی بر اصول مجموعه‌سازی، گردآوری یا تولید شده باشند؛
- اطلاعات دیجیتال باید به صورت منسجم و مناسبی نظیر دیگر منابع اطلاعاتی در کتابخانه‌ها توسط استفاده کنندگان قابل دسترس و بازیابی باشند؛
- اطلاعات دیجیتال باید در مدت زمان طولانی به مثابه منابع اطلاعاتی پایدار قابل دسترس باشند (Deegan & Tanner, 2000).

با توجه به تعاریف پیشگفته و نیز این که در حال حاضر کتابخانه‌های زیادی در ایران حرکت به سمت کتابخانه‌های دیجیتال را آغاز کرده‌اند و برخی از آنها گام‌های مهمی را در این زمینه برداشته‌اند ولی هنوز عنوان کتابخانه دیجیتال برای آنها مناسب نیست و از طرفی نیز برخی از کتابخانه‌ها بدون کسب حداقل قابلیت‌های مورد نیاز عنوان کتابخانه دیجیتال را بر خود نهاده‌اند، لذا معیارهایی جهت انتخاب جامعه پژوهش در تعریف کتابخانه دیجیتال بر اساس تعاریف ذکر شده

¹ Deegan and Tanner

² Digital futures: strategies for the information age

مشخص شدند و در تعریف عملیاتی پژوهش حاضر لحاظ شدند تا بتوان جامعه پژوهشی یکدست و معتبری را برگزید. از این رو در پژوهش حاضر انتخاب، سازماندهی و وجود محتوای دیجیتال و ارائه آن‌ها به کاربران از طریق شبکه جزو ضرورت‌های اولیه کتابخانه‌های دیجیتال برای ارائه اطلاعات به کاربران و مبادله اطلاعات بین سیستم‌های اطلاعاتی قلمداد می‌شوند. بنابراین در تعریف عملیاتی اصطلاح کتابخانه دیجیتال سعی شده تا تعریفی ارائه شود تا بتوان سیستم‌هایی را که حائز این حداقل‌ها هستند در زمره جامعه مطالعاتی پژوهش حاضر قلمداد کرد. لذا سیستم‌هایی که با تعریف ذیل مطابقت داشته باشند جزو جامعه مطالعاتی پژوهش حاضر قلمداد می‌شوند:

کتابخانه‌های دیجیتال سازمان‌هایی هستند که در آن: (۱) منابع اطلاعاتی دیجیتال باید مبتنی بر اصول مجموعه‌سازی، گردآوری یا تولید شده باشند؛ (۲) کارکنان متخصصی به انتخاب و سازماندهی منابع اطلاعاتی پردازند؛ (۳) اطلاعات تمام متن دیجیتال به جامعه کاربران آن کتابخانه ارائه شود، (۴) مدیریت منابع اطلاعاتی با استفاده از سازوکارهای مدیریت محتوا، مانند مدیریت در نمایش و ارائه اطلاعات به کاربران، و مدیریت سطح دسترسی صورت پذیرد؛ و (۵) اطلاعات دیجیتالی به مدت زمان طولانی به مثابه منابع اطلاعاتی پایدار قابل دسترس باشند. کتابخانه‌هایی که همه این عناصر را دارا باشند، در این پژوهش کتابخانه دیجیتال نامیده شده‌اند.

- **سیستم‌های اطلاعاتی نامتجانس.** منظور سیستم‌های اطلاعاتی هستند که از نرم‌افزارهای متنوعی برای کتابخانه دیجیتال خود استفاده می‌کنند. نرم‌افزارهای مذکور دارای ساختارها و قابلیت‌های مختلفی می‌باشند که وجه ممیزه آنها با سیستم‌های اطلاعاتی دیگر است.
- **یکپارچه‌سازی.** منظور یکپارچه ساختن دستیابی به محتواهای اطلاعاتی کتابخانه‌های دیجیتالی است که از نرم‌افزارهای متنوعی استفاده می‌کنند. در این فرآیند لازم است تا زیرساخت‌های لازم تهیه شود تا امکان دستیابی یکپارچه به محتوا فراهم شود.
- **یکپارچه‌سازی معنایی.** منظور برقراری شبکه معنایی واژه‌ها و دستیابی یکپارچه معنایی در سیستم‌های اطلاعاتی یکپارچه کتابخانه دیجیتال ایران است.

۴-۱. پرسش‌های اساسی

۱. عناصر دخیل در یکپارچه‌سازی معنایی کتابخانه‌های دیجیتال کدام‌ها هستند؟
۲. از چه روش‌هایی برای یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال استفاده می‌شود؟
۳. از چه ابزارهایی در یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال استفاده می‌شود؟
۴. وضعیت موجود محتواهای پیشینه‌های کتابشناختی کتابخانه‌های دیجیتال مورد بررسی چگونه است؟
۵. وضعیت موجود خروجی‌های فراداده نرم‌افزارهای کتابخانه دیجیتال مورد بررسی چگونه است؟
۶. الگوی مناسب برای یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال ایران چیست؟

۵-۱. هدف و فایده

۱-۵-۱. هدف

هدف اصلی پژوهش حاضر ارائه الگوی مناسب پیشنهادی برای یکپارچه‌سازی معنایی کتابخانه‌های دیجیتال ایران است. در این راستا اهداف جزئی زیر نیز در این پژوهش قابل حصول خواهند بود:

- شناسایی عناصر دخیل در یکپارچه‌سازی معنایی کتابخانه‌های دیجیتال
- شناسایی روش‌های مورد استفاده برای یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال
- شناسایی ابزارهای مورد استفاده در یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال
- تبیین وضعیت موجود محتواهای کتابشناختی قابل یکپارچه‌سازی در کتابخانه‌های دیجیتال ایران

۲-۵-۱. فایده.

فواید پژوهش حاضر به شرح زیر می‌باشند:

- ارتقاء سیستم‌های اطلاعاتی کتابخانه دیجیتال در سطح بازیابی معنایی
- ارتقاء سطح رضایتمندی کاربران در استفاده از سیستم‌های اطلاعاتی (لازم به ذکر است که این فایده به صورت پیش فرض لحاظ شده است و نیاز به پژوهشی در این زمینه وجود دارد تا این موضوع مورد بررسی قرار گیرد).
- ارتقاء سطح بازیابی اطلاعات در سیستم‌های اطلاعاتی کتابخانه دیجیتال

۶-۱. جنبه‌های نو بودن پژوهش

موضوع کتابخانه‌های دیجیتال و ظهور آنها مبحث جدیدی است. از طرفی نیز بررسی و پژوهش در این زمینه موضوع برخی از پژوهش‌های اخیر قرار گرفته است. پژوهش در زمینه کتابخانه‌های دیجیتال به حوزه‌های متنوع و گوناگونی اختصاص دارد. از جمله این حوزه‌ها می‌توان به نرم‌افزار، سخت‌افزار، راه‌اندازی کتابخانه دیجیتال، سازماندهی اطلاعات و استانداردها و قوانین مرتبط، مجموعه‌سازی، خدمات و غیره اشاره کرد. یکی از مهم‌ترین جنبه‌ها در مطالعه و بررسی در زمینه کتابخانه‌های دیجیتال، یکپارچه‌سازی خدمات و محتواهای کتابخانه‌های دیجیتال می‌باشد. این حوزه به طور مشترک در زمینه قابلیت‌های نرم‌افزارهای کتابخانه دیجیتال و خدمات کتابخانه‌های دیجیتال می‌باشد. به طور کلی یکپارچه‌سازی در دو سطح نحوی و معنایی انجام می‌گیرد. تاکنون توجهی به قابلیت‌های یکپارچه‌سازی خدمات و محتوا در نرم‌افزارهای نامتجانس کتابخانه‌های دیجیتال ایران پرداخته نشده است، و تنها پژوهش‌های انجام گرفته در این زمینه توسط پژوهشگر به عنوان پایان‌نامه دکتری وی^۱ و نیز طرح پژوهشی بررسی هزینه-سازگاری نظام‌های اطلاعاتی نامتجانس در یکپارچه‌سازی اطلاعات: پژوهشی در نرم‌افزارهای کتابخانه‌های دیجیتال ایران در سال ۱۳۹۰ انجام گرفته است. پژوهش حاضر در ادامه پژوهش‌های مذکور و در تکمیل آن‌ها انجام می‌گیرد. لذا حاوی یافته‌ها و جنبه‌های نویی مانند یکپارچه‌سازی معنایی می‌باشد که تاکنون به آنها پرداخته نشده است.

۷-۱. محدودیت‌های پژوهش

^۱ بررسی نحوه مبادله اطلاعات بین سیستم‌های اطلاعاتی کتابخانه‌های دیجیتال در ایران و ارائه الگوی پیشنهادی

با توجه به انجام نخستین بار پژوهشی در این سطح، و نو بودن مبحث کتابخانه‌های دیجیتال در ایران به نظر می‌رسد شناسایی کتابخانه‌های دیجیتال از سایر محصولات می‌تواند چالش برانگیز باشد. از طرفی نیز در طول انجام پژوهش به نظر می‌رسد که شرکت‌های نرم‌افزاری کتابخانه دیجیتال و نیز خود کتابخانه‌های دیجیتال شاید به علت منافع مالی خود و یا دلایل دیگر سعی بر وارونه جلوه دادن و یا ارائه پاسخ‌های نادرست به پرسش‌های طرح شده در پرسشنامه داشته باشند. علاوه بر این ممکن است بنا به دلایل مختلف حاضر به شرکت در فرآیند پژوهش نباشند. از این رو، حین کسب اطلاعات از مسئولان مربوطه سعی شده است تا قابلیت‌های مورد ادعا در نرم‌افزار مورد بررسی و نیز کتابخانه‌های دیجیتال شناسایی و مورد ارزیابی قرار گیرد. تأخیر در ارائه پاسخ به پرسشنامه و سیاهه‌های واری و نیز هماهنگی‌ها در این زمینه نیز به عنوان محدودیتی دیگر در این پژوهش تأثیر گذار بوده است. از طرفی نیز با توجه به منحصر به فرد بودن موضوع پژوهش و نیز اهمیت آن در بستر فرهنگی و علمی ایران پژوهشگر همواره در مستندسازی برخی از محتواها به دلیل فقدان منابع اطلاعاتی مناسب در زمینه یکپارچه‌سازی معنایی مواجه بوده است.

فصل ۲

مبانی نظری و ادبیات پژوهش

۲-۱. مقدمه

دنای دیجیتال مدرن، با تناقضی بزرگ روبروست که این تناقض را بسیاری از کاربران متوجه نمی‌شوند، اما طراحان و معماران نظام‌ها اهمیت این مساله را به‌خوبی درک نموده‌اند. در جهان به‌طور روزافزونی هم‌زمان با افزایش ارتباطات، چندپارگی نیز گسترش پیدا می‌کند (Hart, 2010). امروزه با وضعیتی روبرو هستیم که در آن نظام‌های اطلاعاتی مانند جزایر کوچک جداگانه‌ای هستند که نسبتاً باهم سازگاری داشته و از طریق آبراهه‌های مختلف با هم‌دیگر در ارتباط هستند، افراد برای یافتن اطلاعات و یا منابع اطلاعاتی مورد نیاز خود باید به‌صورت جداگانه به نظام‌های اطلاعاتی مختلف سر بزنند. تعداد بسیار زیاد این جزایر، اطلاعات بسیار زیاد، و انواع مختلف اطلاعات (به‌صورت زبان طبیعی، تصویر، داده و...) ما را بر آن می‌دارند تا به فکر روش‌های ارتباط دادن این جزایر برای یکپارچگی در دستیابی محتواها با هم باشیم (Robertson, 2010). با توجه به رشد فناوری‌ها و الزامات جدید، نظام‌های اطلاعاتی امروزه بسیار پیچیده شده‌اند. مدیریت این پیچیدگی تبدیل به موضوع مورد علاقه برای پژوهشگران شده و آثار قابل توجهی توسط جوامع، پژوهشگران، و شرکت‌های مختلف از اواسط دهه هشتاد میلادی به این موضوع پرداخته‌اند. این امر موجب شده تا نظام‌های اطلاعاتی بتوانند حجم گسترده‌ای از داده‌ها و تراکشن‌های پیچیده را مدیریت کرده و باعث ارائه برنامه‌های کاربردی زیادی در شبکه‌های جهانی شوند. با این وجود، این روزها با چالشی جدید روبرو هستیم که در پی رفع ناهم‌آهنگی^۱ یا غیریکنواختی در نظام‌های اطلاعاتی می‌باشند. ناهم‌آهنگی باعث شده تا یکپارچه‌سازی نظام‌های اطلاعاتی مشکل شود و این مساله در بسیاری از سطوح یک نظام اطلاعاتی می‌تواند واقع شود. مهم‌ترین منبع عدم تجانس در مفهوم، مدل‌سازی و ساختار اطلاعات است به‌طوری که از یک نوع منبع اطلاعاتی مشخص (مانند مقاله، کتاب و غیره)، نمونه‌های مختلفی از مناظر مختلف فرهنگی، منطقه‌ای، و سازمانی ارائه می‌گردد (Kalfoglou, Preface, 2010).

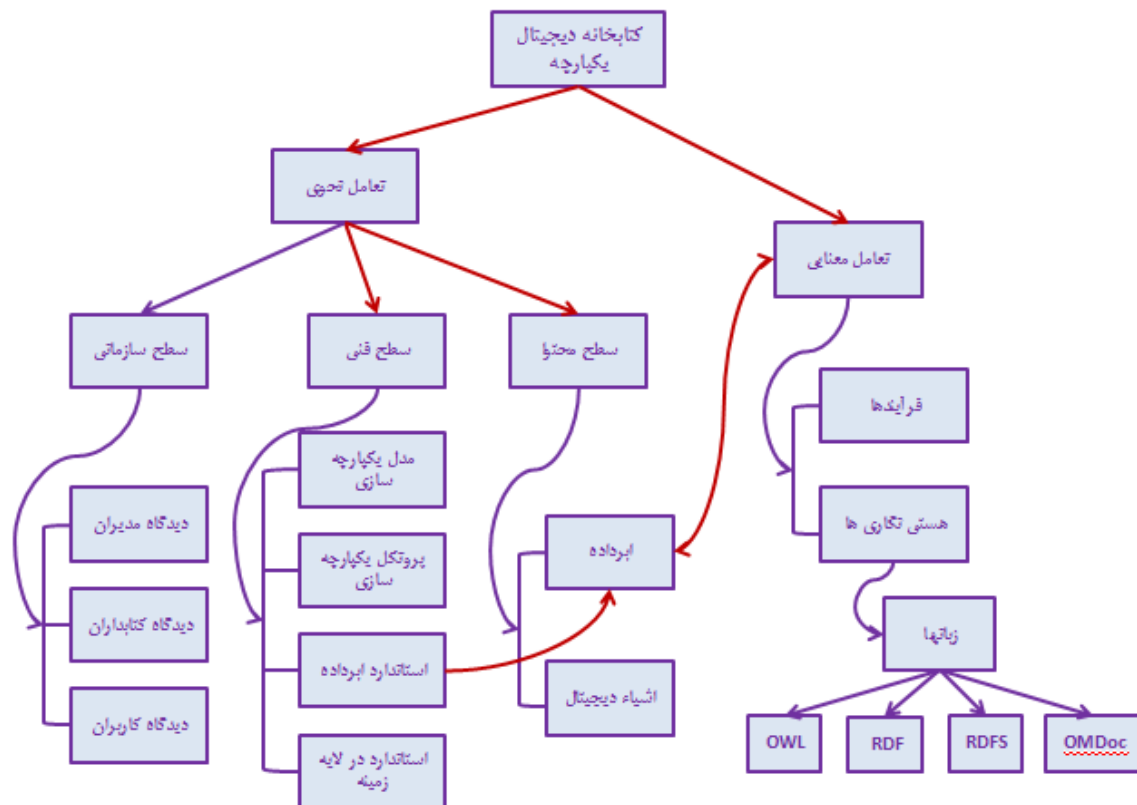
یکپارچه‌سازی در سطوح مختلف در برگزیده اجزاء و عناصری است که قابلیت یکپارچه‌سازی را فراهم می‌کنند. این موضوع در سیستم‌های اطلاعاتی و به‌ویژه سیستم‌های اطلاعاتی نامتجانس دارای پیچیدگی خاص خود است. یکپارچه‌سازی کتابخانه‌های دیجیتال در برگزیده عناصر، روابط و فرآیندهای مشخص شده در شکل ۱-۲ می‌باشد. همان‌گونه که از شکل نیز بر می‌آید یکپارچه‌سازی کتابخانه‌های دیجیتال از دو منظر تعامل نحوی و تعامل معنایی قابل اجرا است. اولویت در یکپارچه‌سازی با تعامل نحوی است. پس از پیاده‌سازی تعامل نحوی است که تعامل معنایی قابلیت اجرایی پیدا می‌کند. تعامل نحوی به برقراری شبکه ارتباطی بین کتابخانه‌های دیجیتال و امکان مبادله اطلاعات بین کتابخانه‌های دیجیتال عضو گفته می‌شود.^۲ در این سطح از تعامل سه سطح محتوا، فنی و سازمانی مورد نظر هستند که باید در پیاده‌سازی تعامل نحوی در یکپارچه‌سازی کتابخانه‌های دیجیتال مورد توجه قرار گیرند. سطح فنی از نظر اهمیت و اولویت اجرا می‌تواند در سطح بالاتری نسبت به سایر سطوح قرار گیرد^۳ (علیپور حافظی، ۱۳۸۸).

^۱Heterogeneity

^۲ لازم به ذکر است که با توجه به تاکید متن حاضر این مفهوم فقط در محدوده سیستم‌های اطلاعاتی کتابخانه دیجیتال قرار گرفته است. با این وجود مفاهیم

ذکر شده می‌تواند محدوده گسترده‌ای از سیستم‌های اطلاعاتی مانند پایگاه‌های اطلاعاتی، سیستم‌های مدیریت محتوا، و نظایر اینها را نیز شامل شود.

^۳ برای کسب اطلاعات بیشتر در این زمینه به پایان نامه علیپور حافظی (۱۳۸۸) مراجعه نمایید.



شکل ۱-۲. یکپارچه‌سازی کتابخانه‌های دیجیتال

پس از ایجاد یکپارچه‌سازی نحوی، برای ارتقاء سطح دستیابی به محتوای اطلاعاتی و یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال، برقراری تعامل معنایی اولویت ویژه‌ای می‌یابد. در حقیقت تعامل معنایی به برقراری شبکه مفهومی و معنایی میان اصطلاح-های مورد استفاده و همچنین محتواهای فیلدهای اطلاعاتی در کتابخانه‌های دیجیتال عضو در نظام، یکپارچه‌سازی گفته می‌شود. در این سطح از تعامل با استفاده از فرآیندهای مورد نیاز و نیز ابزارهای مناسب برای ایجاد شبکه معنایی بین اصطلاح‌ها اقدام به عملی ساختن شبکه معنایی و مفهومی بین منابع اطلاعاتی می‌شود.

در این فصل پس از بررسی مفهوم، روش‌ها و ابزارهای یکپارچه‌سازی مفهومی و معرفی نمونه‌ای از پروژه‌های انجام شده در این زمینه، پیشینه پژوهش در داخل و خارج از کشور ارائه شده است.

۲-۲. یکپارچه‌سازی محتوا و خدمات در کتابخانه‌های دیجیتال

نظام‌های کتابخانه‌ای دیجیتال از بدو پیدایش تا کنون به لحاظ تعداد و استفاده از برنامه‌های کاربردی تخصصی و پیاده‌سازی ساختارهای داده‌ای بسیار پیچیده، گسترش پیدا کرده‌اند (McDaniel & Kruk, 2009). در نتیجه این گسترش در تعداد کتابخانه‌های دیجیتال همچنین تنوع ساختارها و نرم‌افزارهای مورد استفاده در آنها، افراد برای پیدا کردن اطلاعات مورد نیاز خود ناگزیر از مراجعه به کتابخانه‌های دیجیتال مختلف هستند بطوری که هرکدام از این کتابخانه‌ها ممکن است ملزومات نرم‌افزاری خاص خود را داشته و همچنین اطلاعات را در قالب‌های متنوع و مخصوص به خود به افراد ارائه نمایند. راه‌حلی که به منظور فائق آمدن بر این مشکل به نظر می‌رسد، تعامل^۱ کتابخانه‌های دیجیتال مختلف از طریق یکپارچه‌سازی^۲ امکانات و ساختارهای اطلاعاتی

^۱interoperability

^۲integration

موجود در آنها است. تعامل مساله بسیار مهمی است که تمامی انواع سیستم‌های اطلاعاتی دیجیتالی را تحت تاثیر خود قرار داده اما شهرت خود را مدیون گسترش وب است. در تعامل، این پتانسیل وجود دارد تا بسیاری از کارهایی که در حال حاضر به صورت دستی انجام می‌گیرند، به صورت خودکار انجام پذیرند (Patel, Koch, Doerr, & Tsinaraki, 2005)

کار بر روی یکپارچه‌سازی کتابخانه‌های دیجیتال تا حدی بر روی سه مساله زیر متمرکز است (Shen R. , Applying the 5S Framework To Integrating Digital Libraries, 2006)

(۱) پراکندگی^۱: گستره جغرافیایی

(۲) ناهماهنگی^۲: تفاوت در سطوح فنی (مانند سخت‌افزار، سیستم عامل، زبان برنامه نویسی، و غیره) و سطح مفهومی (مانند فهم و مدل سازی مختلف از یک موجودیت واقعی)

(۳) استقلال^۳: میزان خودکفایی اجزا^۴

از طریق یکپارچه‌سازی کتابخانه‌های دیجیتال قصد داریم که پراکندگی و ناهماهنگی‌ها را مخفی کنیم و هم‌زمان (تا حدودی) باعث ممکن شدن و روشن ساختن استقلال اجزا شویم (همان).

اوکسل و شت^۵ چهار نوع ناهماهنگی (عدم تجانس) که زمینه‌ساز چهار نوع تعامل بالقوه است، را شناسایی کردند (Patel, Koch, Tsinaraki & Doerr, 2005) که عبارتند از:

– سیستمی: ناسازگاری بین سخت‌افزار و سیستم‌های عامل

– نحوی^۶: تفاوت در کدگذاری‌ها و بازنمایی

– ساختاری^۷: تنوع در مدل‌های داده‌ای، ساختارها و طرح‌های داده‌ای

– معنایی^۸: ناسازگاری در اصطلاحات به کار رفته و معانی

یکپارچه‌سازی کتابخانه دیجیتال می‌تواند سطوح مختلفی داشته باشد، مانند سطح اطلاعات و سطح خدمات. اطلاعات یکپارچه باعث می‌شود که مجموعه‌های پراکنده و نامتجانس منابع اطلاعاتی به عنوان یک واحد منسجم به نظر بیاید. خدمات یکپارچه نیز به کاربران این توانایی را می‌دهد تا از طریق واسطه‌های کاربری همسان‌تر و سهل‌الاستفاده‌تری که تفاوت‌های معنایی و نحوی را در کتابخانه‌های دیجیتال یکپارچه‌شده، پنهان می‌نمایند، استفاده جامع‌تری از منابع کتابخانه دیجیتال نمایند. با اینکه تلاش‌های بسیاری در زمینه مساله یکپارچه‌سازی کتابخانه‌های دیجیتال صورت گرفته، بسیاری نیز رویکردهای خود را به صورت موقتی و تدریجی ایجاد نموده‌اند (Shen R. , Applying the 5S Framework To Integrating Digital Libraries, 2006).

اگر منطقه یا ملتی به دنبال بهترین روش برای استفاده از دارایی‌های اطلاعاتی خود و کاهش چندبارگی در جمع‌آوری داده‌ها است، اشتراک اطلاعات در میان بخش‌های عمومی ضروری است (Gottschalk & Solli-Sæther, 2009). بوکوئت^۹ و دیگران (۲۰۱۰) بیان می‌کنند که به صورت کلی، تعامل در سطح اطلاعاتی به دو دلیل مشکل است. این دو دلیل عبارتند از:

¹distribution

²Heterogeneity

³Autonomy

^۴منظور از اجزا، هر یک از کتابخانه‌های دیجیتالی هستند که به صورت یکپارچه با هم عمل می‌کنند

⁵Ouksel & Sheth

⁶Syntactic

⁷Structural

⁸Semantic

⁹Bouquet

۱. از یک سو، اطلاعات ممکن است در سطح ساختاری^۱ نامتجانس باشند (برای مثال پایگاه اطلاعاتی افراد و پروژه‌ها در یک شرکت ممکن است بر اساس طرح^۲‌های مختلف باشد و از این رو یکپارچه‌سازی آنها ممکن است مستلزم نوعی نگاشت^۳ موقت بین دو طرح باشد)؛
 ۲. از سوی دیگر یک موجودیت^۴ واقعی ممکن است به صورت‌های مختلف از جمله کلیدهای اولیه^۵ مختلف در پایگاه‌های اطلاعاتی مختلف، تنوع نام در زبان طبیعی^۶، توصیف‌های مختلف، یو.آر.آی. (شناسگر متحدالشکل منابع) مختلف در منابع اطلاعاتی وبی (مانند پایگاه‌های دانشی آر.دی.اف./او.دبلیو.ال.^۷) ثبت شده باشند.
- بوکوئث و همکاران دسته اول را ناهماهنگی در سطح طرح و دسته دوم را ناهماهنگی در سطح موجودیت می‌نامند. دسته دوم ناهماهنگی برخلاف دسته اول که تعداد پژوهش‌های مهم در آن زیاد است، پژوهش‌های بسیار کمتری را به خود اختصاص داده است (Bouquet, Stoermer, Barczyński, & Bocconi, 2010)
- به منظور یکپارچه‌سازی داده‌ها، پیشینه‌هایی (و پیروان شناساگرهایی) که به شیء یکسان دلالت دارند باید با هم منطبق شوند. این البته یک جنبه از یکپارچه‌سازی داده‌ها است، یعنی همان جنبه نحوی، اما این جنبه، گامی اساسی در رسیدن به فهمی مشترک در مورد معنای پاره‌های مختلف داده‌ها و به عبارت دیگر همان یکپارچه‌سازی معنایی داده‌ها، است (همان).
- در کنار انتقال داده‌ها، کتابخانه دیجیتال در مقیاس وسیع بعید نیست که شامل چندین واسط کاربری یا چندین نوع ارائه از یک خدمت واحد باشد. به علاوه، این واسط‌های کاربری و خدمات در مکان‌های جغرافیایی مختلف قابل ارائه هستند که از طریق یک شبکه به هم متصل می‌شوند. به خاطر اینکه به کاربران نمای یکپارچه مجموعه داده‌ها و مجموعه ملموس خدمات را ارائه کنیم، نیازمند سازوکارهایی برای دسترسی از راه دور به خدمات کتابخانه دیجیتال هستیم. ساده‌ترین شکل این سازوکارها، مفهوم جستجوی از راه دور^۹ است که در آن به جای واسط انسانی یا در کنار آن، واسط کاربری ماشینی^{۱۰} برای موتور جستجو قابل دسترس بوده و از این رو یکپارچه‌سازی موتورهای جستجوی راه دور را در یک واسط کاربری و جریان‌های کاری^{۱۱} محلی مقدور می‌سازد. پروتکل Z39.50 به صورت خاص در راستای همین عملکرد بازایی از راه دور طراحی شده و در نظام‌های کتابخانه‌ای طرفدار است (Fox, Suleman, Gaur, & Devika, 2005).
- یکی از روش‌های یکپارچه‌سازی محتوا و خدمات در کتابخانه‌های دیجیتال، استانداردسازی است. استانداردسازی یکی از رویکردهای سنتی به مساله تعامل است که در این رویکرد، تمامی مشارکت کنندگان در راستای استفاده از استانداردهای یکسان مانند استانداردهای فراداده‌ای^{۱۲} و پروتکل‌های انتقال داده^{۱۳} به توافق برسند. یکی از مزایای استاندارد سازی این است که اطلاعات به سرعت قابل انتقال و یکپارچه‌سازی است، بدون اینکه تغییری در آن رخ دهد (Shen R. , Applying the 5S Framework To Integrating Digital Libraries, 2006). با وجود این راهکارهای دیگری نیز می‌توانند برای این

¹Structural

²Schema

³mapping

⁴entity

⁵Primary key

⁶Natural language

⁷URI(Uniform Resource Identifier)

⁸RDF/OWL

⁹Remote searching

¹⁰Machine interface

¹¹workflows

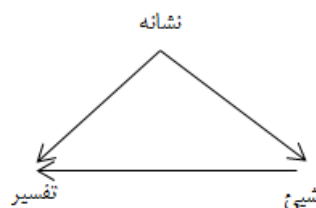
¹²metadata

¹³Transaction protocol

منظور مورد استفاده قرار گیرند که از محدوده این بحث خارج هستند. با توجه به اهمیت معنا و محتوا در مبحث یکپارچه‌سازی کتابخانه‌های دیجیتال در ادامه به این موضوع در سیستم‌های بازیابی اطلاعات می‌پردازیم.

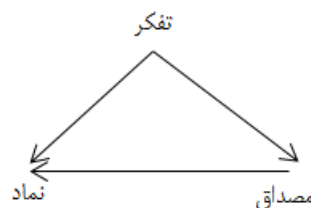
۲-۳. معنا^۱ در بازیابی اطلاعات

گفتار و نوشتار شامل علائم و نمادهای قراردادی هستند که با در کنار هم قرار گرفتن آنها محتوا شکل می‌گیرد. با درک محتوا توسط افراد معنا در افراد گیرنده محتوا شکل می‌گیرد. بر این اساس زیربنای تمام این فرآیند در بستری به نام ارتباط قرار دارد. بنا بر این افراد به واسطه ارتباط و تعامل با یکدیگر و با سیستم‌های اطلاعاتی به رفع نیازهای خود اقدام می‌کنند. پیرس^۲ با توجه به این مفهوم بود که در سال ۱۹۸۵ الگوی سه گانه نشانه، شیء و مخاطب یا تفسیرگر خود را به صورت زیر بیان کرد.



شکل ۲-۲. الگوی نشانه‌شناسی پیرس (نشاط، ۱۳۹۰)

همین الگو را اگدن و ریچاردز^۳ در سال ۱۹۲۳ به شکل زیر نشان دادند:



شکل ۲-۳. مثلث معنایی اگدن و ریچاردز (نشاط، ۱۳۹۰)

بنابر نمودارهای مذکور در عمل با نشانه‌ها یا نمادهایی مواجه هستیم که نشانگر شیء یا مصداقی هستند و برداشت‌های افراد از این مصداق‌ها یا اشیاء شکل دهنده تفسیر آنها یا تفکر آنها است و معنا را در ذهن افراد شکل می‌دهد. بنابراین تجزیه و تحلیل معناشناختی متون بر این اصل استوار است که متن محل تجمع نشانه‌ها یا نمادها نیست، بلکه محل شکل‌گیری فرآیندی معنایی است (نشاط، ۱۳۹۰). شکل‌گیری معنا در ذهن با ارتباط نمادهای دریافتی جدید با نمادهای پیشین فرد اتفاق می‌افتد. بنابراین فرد همواره با ذهنی مملو از مفاهیم و پیش‌داشته‌ها و پیش‌داوری‌ها سراغ متن یا موضوعی خاص می‌رود و با پرسش‌ها و انتظارات خاصی به کاوش می‌پردازد. بر این اساس است که گادامر^۴ آغاز هر گونه فهمی را پیش‌داوری می‌داند (نشاط، ۱۳۹۰).

با توجه به مطالب پیش‌گفته بحث در زمینه بازیابی اطلاعات شکل پیچیده‌ای به خود می‌گیرد. به ویژه زمانی که بخواهیم نظام‌های بازیابی اطلاعات را مبتنی بر شبکه معنایی مورد نیاز مخاطبان ارتقاء دهیم. بازیابی اطلاعات زمانی مطرح می‌شود که ما پایگاه(های) اطلاعاتی شامل محتوا داریم و قصد بازیابی برخی از اطلاعات از میان حجم انبوه اطلاعات ذخیره شده را داریم. از منظر بازیابی شرط لازم برای موفقیت در جستجو بستگی به میزان ارتباط میان واژگان به کار رفته در سیستم و واژگان مورد استفاده توسط مخاطبان سیستم دارد. واژگان بکار رفته در سیستم نیز بستگی به واژگان مورد استفاده توسط تولید کنندگان محتوا و نظام ذخیره و

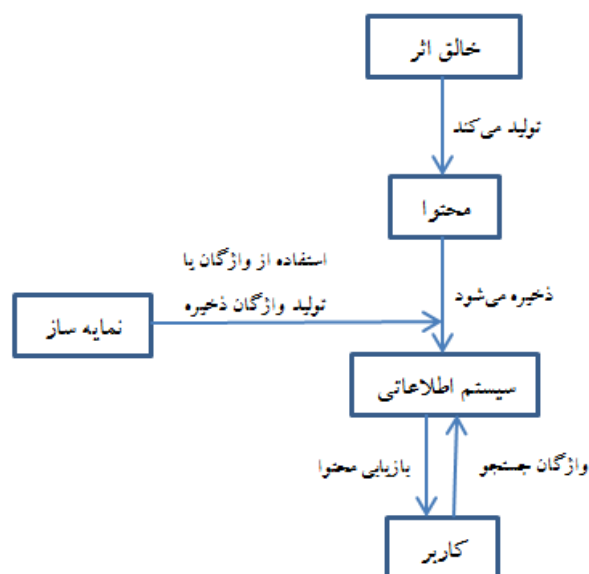
^۱ semantics

^۲ Peirce

^۳ Ogden, Richards

^۴ Gadamer

بازیابی مورد استفاده در سیستم اطلاعاتی دارد. در چنین شرایطی با شبکه‌ای متشکل از تولید کننده محتوا، محتوا، سیستم اطلاعاتی، مفسر و نمایه‌ساز (بسته به سیاست سیستم اطلاعاتی)^۱ و مخاطب یا کاربر سیستم به شکل زیر مواجه هستیم.



شکل ۲-۴. شبکه تولید و استفاده از محتوا

بازنمایی اطلاعات در قالب الکترونیک، ذخیره‌سازی آن در سیستم‌های پایگاه اطلاعاتی و بازیابی اطلاعات از این پایگاه‌ها یکی از مهم‌ترین زمینه‌های مورد پژوهش در علوم کامپیوتر است. امروزه بازیابی اطلاعات اهمیت فراوانی یافته چرا که دیگر محدود به بررسی نحوه پیدا کردن منابع در پایگاه‌های اطلاعاتی نیست. بازیابی اطلاعات مدرن، به زمینه‌های واسطه‌های کاربری هوشمند^۲، بازخورد کاربر^۳، و تا حدی به برجسب‌زنی^۴ به منظور نامگذاری نیز می‌پردازد. اصطلاح پایگاه اطلاعاتی خود نیز به حوزه‌های جدید نظیر کتابخانه‌های دیجیتال و وب گسترش یافته است (Linckels & Meinel, 2011).

تا حد زیادی کتابخانه‌های دیجیتال از طریق ساختار راهنما^۵ یا از طریق نمایه تمام‌متنی^۶ که جستجوی محتوای منابع اطلاعاتی را مقدور می‌سازند، جستجوپذیر هستند. وجود امکان بررسی راهنما، سازوکاری برای پاسخ به سؤالاتی نظیر "تمامی منابع را به-صورت الفبایی فهرست کن" یا "تمامی منابعی را که از تاریخی مشخص تا کنون ایجاد شده‌اند را فهرست کن" می‌باشد. جستجوی نمایه تمام متن این حالت را گسترش داده و به پرسش‌هایی نظیر "تمامی منابعی که حاوی کلماتی خاص هستند را فهرست کن" یا "تمامی منابعی را که حاوی اسم یا کلمه‌ای خاص هستند را حذف کن"، پاسخ می‌دهد. اما روشی برای پاسخ به سؤالاتی نظیر "تمامی منابعی را که مرتبط با موضوعی خاص هستند را فهرست کن" وجود ندارد به‌ویژه اگر اصطلاح مشخص

^۱ با توجه به سیاست مورد استفاده در سیستم اطلاعاتی این امر می‌تواند مرتبط با جستجوی تمام متن، جستجوی آزاد (استخراج واژگان بکار رفته توسط تولید کننده محتوا) و جستجوی مبتنی بر نمایه‌سازی و اصطلاح‌های پذیرفته شده در قالب اصطلاحنامه‌ها و آنتولوژی‌ها باشد.

^۲ Intelligent user interfaces

^۳ User feedback

^۴ tagging

^۵ Directory structure

^۶ fulltext

کننده آن موضوع در منابع اطلاعاتی یک کتابخانه دیجیتال موجود نباشد. در اینجا اضافه نمودن معنا به کتابخانه دیجیتال می‌تواند مفید باشد (McDaniel & Kruk, 2009)

معنا تا حد زیادی به بافتی که فرد در آن به جستجوی اطلاعات می‌پردازد بستگی دارد (Guha, 2006) محدودیت بزرگ بسیاری از مدل‌ها و نظام‌های موجود بازیابی اطلاعات این است که بازیابی اطلاعات تنها بر اساس سوال^۱ کاربر و مجموعه اسناد صورت می‌گیرد؛ اطلاعاتی در مورد کاربری خاص و جستجوی بافت^۲ تا حد زیادی نادیده گرفته می‌شود. بازیابی حساس به بافت به‌عنوان چالشی عظیم در پژوهش‌های بازیابی اطلاعات تبدیل شده است (Shen, Tan, & Zhai, 2005).

اطلاعات معنایی از طریق فراداده‌هایی که به هر منبع اطلاعاتی اختصاص داده می‌شود و یک یا چند هستی‌نگاری^۳ جهت ارائه بافت معنایی برای جستجوها، بازنمایی می‌گردد (همان)، کاربران می‌توانند در چنین نظام‌های اطلاعاتی علاوه بر جستجوی عادی (که به معنای بازیابی منابع اطلاعاتی بر اساس انطباق کلیدواژه‌های ورودی با کلیدواژه‌های موجود در منابع اطلاعاتی) به جستجوی معنایی (که در آن لزومی بر وجود کلیدواژه یا اصطلاح مورد نظر کاربر در منابع اطلاعاتی وجود ندارد)، پرداخته و به این ترتیب جامعیت و مانعیت جستجوی خود را افزایش دهند.

بوترام^۴ (۲۰۱۰) چهار مدل بازیابی اطلاعات معنایی را برمی‌شمارد که عبارتند از:

۱. انطباق الگوی ساده و مدل‌های بنیادی تعامل معنایی: بنیادی‌ترین روش جستجو بر اساس انطباق واژه‌های مختلف با هم است.

۲. سوالات مفهومی و تعامل مفهومی: در این حالت فرایند بازیابی به‌جای واژه‌های مختلف بر روی سطح مفهومی تمرکز می‌کند که در این حالت روابط مفاهیم، مترادف‌ها، و سایر نقاط دسترسی مفید به واژه ورودی کاربر که بازنمون مفهوم ذهنی است متصل می‌شوند. در این حالت تمامی مترادف‌های مربوط به هم یکپارچه شده و از این رو جامعیت جستجو را بالا می‌برد.

۳. اکتشاف معنایی و تعامل مفهومی بهبود یافته^۵: در این حالت به کاربر محیطی مفهومی ارائه می‌شود که در آن روابط بین مفاهیم و حدود و ثغور آنها نشان داده شده و کاربر می‌تواند در این محیط به کنکاش پرداخته و اشکال مختلف محاوره‌ای و پیش‌رونده پرسش خود را ببیند (اصلاح پرسش)، طبیعی است که چنین حالتی نیازمند ساختار رابطه‌ای مفصلی است.

۴. اکتشاف موضوعی^۶ و تعامل موضوعی پیشرفته. در این حالت نه تنها به کاربران در کشف زمینه‌های موضوعی کمک می‌کند، بلکه به آنها کمک می‌کند تا نیاز اطلاعاتی خود را روشن کرده، تعریف نماید و در قالب الفاظ بیان کند (روشن ساختن سوال) (Boteram, 2010).

بر این اساس متوجه می‌شویم که بازیابی اطلاعات در اصل به دنبال معنا بوده و همواره تلاش بر این بوده تا در نظام‌های بازیابی اطلاعات کاربر به منابع و اطلاعات مورد نیاز و مرتبط با نیاز اطلاعاتی خود دست یابد. از طرفی نیز اکثر نظام‌های بازیابی اطلاعات موجود در حال حاضر چنین قابلیت‌هایی را ندارند ولی به تدریج در این مسیر حرکت می‌کنند. همانگونه که بوترام (۲۰۱۰) نیز اشاره کرده مدل‌ها و سطوح مختلفی در بازیابی معنایی اطلاعات وجود دارند که نگاه عمیق به آنها نشانگر بلوغ آنها در طی زمان است. همین امر بیانگر این حقیقت است که رویکرد توسعه سیستم‌های اطلاعاتی و بازیابی اطلاعات رویکردی معنایی است و به دنبال آن

¹Query

²context

³ontology

⁴Boteram

⁵Semantic exploration and enhanced conceptual interoperability

⁶Thematic

سیستم‌های اطلاعاتی هوشمند را خواهیم داشت. بر این اساس سیستم‌های اطلاعاتی یکپارچه نیز باید دارای قابلیت‌های بازیابی معنایی اطلاعات باشند.

۴-۲. یکپارچه‌سازی معنایی و روش‌های انجام آن

راه‌حل‌های ارائه شده برای مساله عدم تجانس، اسامی مشترکی دارند، یکپارچه‌سازی در سطح سیستم یا تعامل در سطح وظایف^۱. وقتی که راه‌حلی "معنا^۲های" اطلاعاتی را که در راستای تعامل کتابخانه‌های دیجیتال نیازمند یکپارچه‌سازی آنها هستیم در نظر می‌گیرد، ما معمولاً اصطلاح مورد نظر خود را بیشتر اصلاح کرده و از اصطلاح یکپارچه‌سازی معنایی^۳ یا تعامل معنایی^۴ استفاده می‌کنیم (Kalfoglou & Hu, 2010).

شن (۲۰۰۶) معتقد است که در دنیای واقعی، باید ابتدا تعامل نحوی صورت بگیرد تا تعامل معنایی بتواند محقق شود. هرگونه طرح تعامل نحوی در کتابخانه‌های دیجیتال نیازمند عناصری در ۴ سطح می‌باشد که عبارتند از: مدل‌ها^۵، پروتکل‌ها^۶، استانداردهای فراداده‌ای^۷، و سطح بافت محور^۸ (Alipour-Hafezi, Horri, Shiri, & Ghaebi, 2010).

عامل اول که در ایجاد شبکه یکپارچه کتابخانه‌های دیجیتال نقشی بسیار مهم دارد، انتخاب مدل مناسب است. مدل‌های مختلفی برای تعامل نحوی موجودند که عبارتند از مدل جستجوی هم‌زمان^۹، برداشت اطلاعات^{۱۰}، ترکیبی^{۱۱} (استفاده هم‌زمان از مدل جستجوی هم‌زمان و برداشت اطلاعات) و مدل جمع‌آوری^{۱۲}، که هر کدام بسته به شرایط کتابخانه‌های دیجیتال موجود قابل استفاده هستند. در هر یک از این مدل‌ها از پروتکل‌هایی برای تبادل اطلاعات استفاده می‌شود. اطلاعات مبادله شده در قالب فراداده‌ای، با توجه به وجود استانداردهای مختلف فراداده‌ای می‌تواند به دو صورت باشد. رویکرد اول استفاده از استاندارد فراداده‌ای واحد در کتابخانه‌های دیجیتال (که با توجه به سادگی و سهولت، دو استاندارد دابلین‌گر^{۱۳} و یا مودس^{۱۴} توصیه می‌شوند) و رویکرد دوم استفاده از مترجم^{۱۵} برای تبدیل استانداردهای فراداده‌ای مختلف به یکدیگر یا یک یا دو استاندارد فراداده‌ای مشترک می‌باشد. پس از تصمیم‌گیری در زمینه رویکرد مناسب برای یکپارچه‌سازی استانداردهای فراداده‌ای، نوبت بافتی است که این استانداردها در آن ارائه می‌گردند. دو بافت رایج برای ارائه استانداردهای فراداده‌ای، بافت متنی ساده^{۱۶} و بافت سلسله مراتبی^{۱۷} هستند که با توجه به کارایی بالای ایکس‌ام‌ال^{۱۸} زمینه بافت سلسله مراتبی، استفاده از آن توصیه می‌گردد (Alipour-Hafezi, Horri, Shiri, & Ghaebi, 2010) شکل ۲-۵ این عوامل و روابط میان آنها را به تصویر می‌کشد.

¹Interoperability at a task level

²meaning

³Semantic integration

⁴Semantic interoperability

⁵ Models

⁶ Protocols

⁷ Metadata standards

⁸ Context-based levels

⁹ Federated or On-demand

¹⁰ Harvested or data warehousing

¹¹ Hybrid

¹² Gathering

¹³ Dublin core

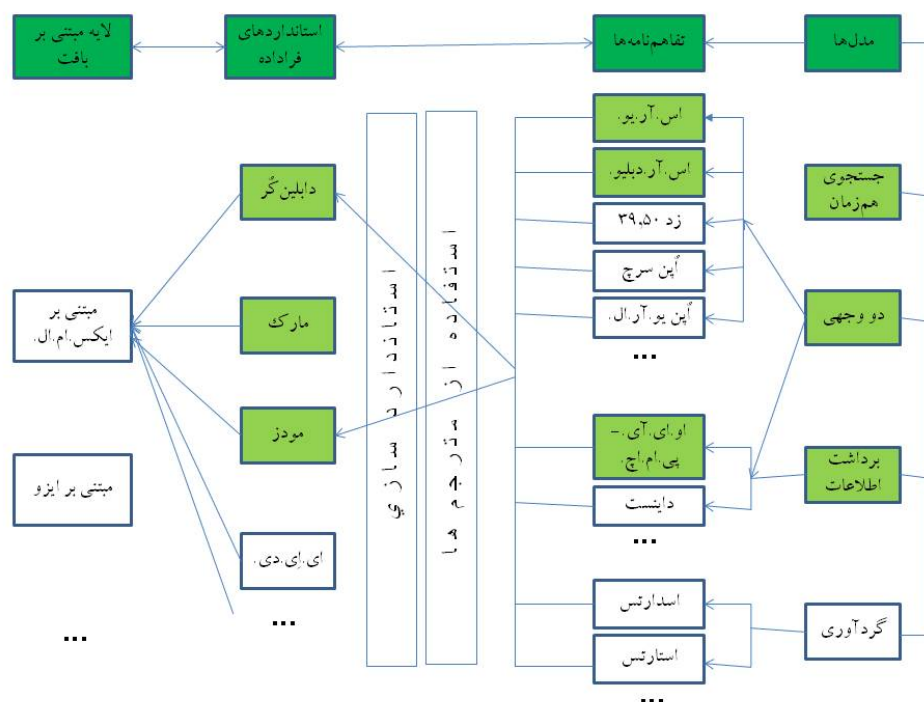
¹⁴ MODS

¹⁵ Translator

¹⁶ Simple text

¹⁷ Hierarcal

¹⁸ XML



شکل ۲-۵. عوامل دخیل در یکپارچه‌سازی نحوی و روابط آنها با هم

منظور از معنا همان معنای اصطلاحات و عبارات مورد استفاده در توصیف یا محتوای منابع اطلاعاتی است. از این رو تعامل معنایی "توانایی نظام‌های اطلاعاتی در تبادل اطلاعات بر اساس معناهای مشترک، از پیش تعیین شده و مذاکره شده اصطلاحات و عبارات" است و به منظور محقق ساختن سایر انواع تعامل‌ها (نحوی، میان فرهنگی، بین‌المللی و غیره) به آن نیازمندیم (Veltman, 2009)

تعامل معنایی چالش اصلی در جمع‌آوری و واسطه‌گری اطلاعات است. عدم تجانس معنایی نشان دهنده عدم تطابق در معناها است و به علت یکپارچه‌سازی سیستم‌های اطلاعاتی که به صورت جداگانه طراحی شده‌اند، رخ می‌دهد. انواع مختلف ناسازگاری‌ها موجب ناهماهنگی معنایی می‌شوند مانند نام، ساختار، ویژگی‌ها، قابل تقسیم بودن ارزش‌ها و غیره. قضیه بدیهی ریاضیات $X=X$ در دنیای واقعی عبارت^۱ طبیعی (که در وب شایع است) نامعتبر است و در آن افراد ممکن است منظور خود را به صورت بسیار موثر بیان کنند. از این رو نیازمند تعامل معنایی هستیم تا فراسوی فهم ساده اصطلاحات رفته و معنای عمیق آنها را درک کنیم (Lakshminarayan, 1998)

یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتالی به معنای توانایی سیستم‌های اطلاعاتی مختلف در انتقال اطلاعات یکسان با معنای مورد نظر است. یکپارچه‌سازی اطلاعات تنها یکی از نتایج ارتباط موفق است. به خاطر وجود زبانهای گویشی مختلف، ارتباط به دو شیوه قابل حصول است: ۱) اجبار افراد برای یادگیری و استفاده از زبان واحد ۲) پیدا کردن مترجمانی که می‌دانند چگونه به‌طور مناسب اطلاعات یک فرد را به فرد دیگر تفسیر نمایند. رویکرد اول استاندارد سازی پیشگیرانه^۲ بوده و دومی تفسیر واکنشی^۳ است (Shen R. , 2006). در حالت اول در مورد استانداردهایی به توافق می‌رسیم (مانند استانداردهای فراداده‌ای) که پیش از وقوع هر گونه عدم تجانس، همه کتابخانه‌های دیجیتال به منظور یکپارچه‌سازی معنایی از آن استفاده کنند و در حالت دوم ساختارها و

^۱expression

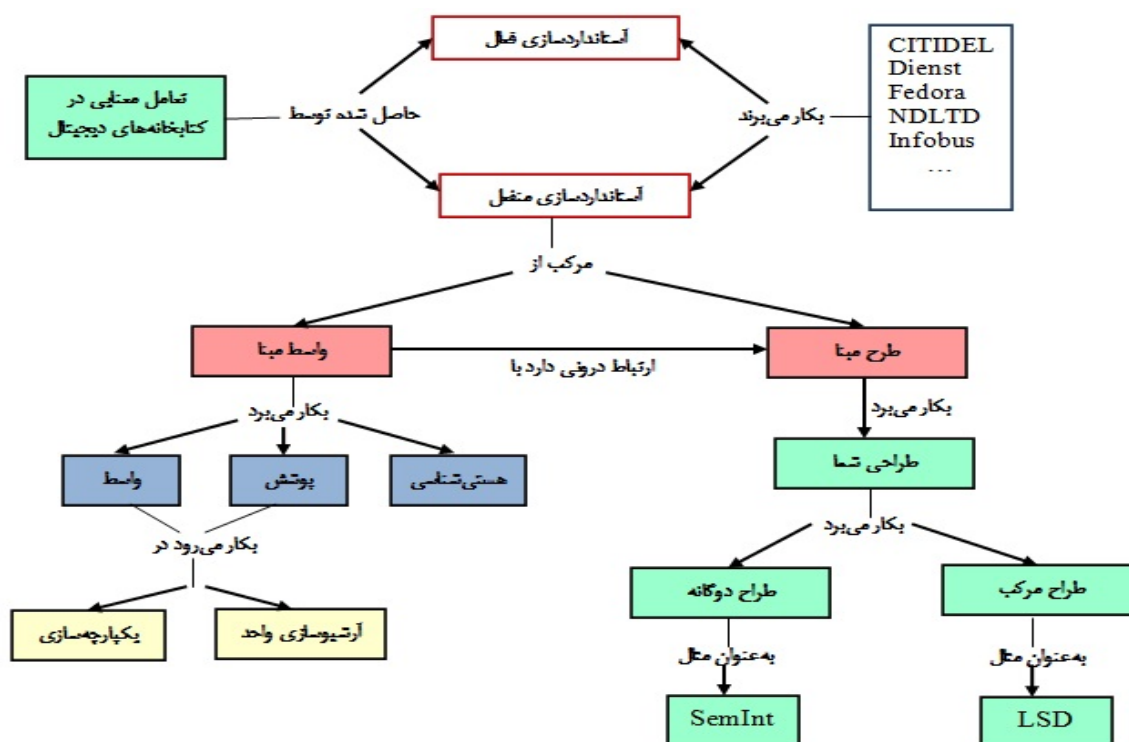
^۲Proactive standardization

^۳Reactive interpretation

قالب‌هایی که هریک از کتابخانه‌های دیجیتال از آن استفاده کرده‌اند به منظور یکپارچه‌سازی به هم تبدیل می‌کنیم مانند تبدیل قالب‌های فراداده‌ای به یکدیگر (همان). بنابراین رویکرد اول حالتی پیشگیرانه داشته و رویکرد دوم واکنشی است.

استاندارد سازی در حالتی که میزان تنوع کمتر و ثبات طولانی مدت^۱ بالایی وجود دارد مناسب است. استاندارد سازی از بروز تنوع سودمند و دگرگون شده جلوگیری می‌کند و باعث فروکاهی اطلاعات می‌شود. در راستای کاربست آن، ممکن است نیازمند مفسرانی باشیم که ورودی‌ها را در قالب استاندارد تولید نمایند در مقابل رویکرد تفسیری در محیط‌هایی که میزان تنوع بیشتر بوده و ثبات طولانی مدت کمتری وجود دارد مناسب است. مفسران می‌توانند در درون یک محیط یکپارچه بر اساس نیازهای گروه‌های فرعی و کوچک‌تر، منعطف و گزینشی عمل نمایند (Patel, Koch, Doerr, & Tsinaraki, 2005). در رویکرد تفسیری نیز امکان فروکاهی اطلاعات وجود دارد (Shen R. , 2006)

در شکل ۶-۲ نقشه محتوایی فعالیت‌های مرتبط در زمینه یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال ارائه شده است.



شکل ۶-۲. یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتالی (علیپور حافظی، ۱۳۹۰)

۲-۵. وب معنایی^۲

رشد سریع وب منجر به تکثیر منابع اطلاعاتی و محتوای قابل دسترس از طریق اینترنت شده است. با اینکه پیشرفت‌های صورت گرفته در زمینه توانمندی‌های سخت‌افزاری به سرعت به جریان اطلاعات در کامپیوترهای شبکه‌ای کمک می‌کند، با این حال مشکل اصلی برای کاربر انسانی در جهت همگام شدن با گسترش سریع فضای اطلاعاتی وب پابرجا است. هرچند بخشی از اتاق که به کامپیوترها اختصاص داده شده تا به افراد جهت کشف، راهبری، و یکپارچه‌سازی اطلاعات در این فضای اطلاعاتی وسیع کمک

^۱Long term stability

^۲Semantic web

کنند، اما شیوه‌ای که امروزه از طریق آن اطلاعات در وب بازنمایی و ساختار بندی می‌شود به راحتی برای کامپیوترها قابل خواندن نیست. برای پرداختن به این مساله، وب معنایی ظهور کرده است. وب معنایی ساختار اطلاعاتی جدیدی را متصور می‌شود که رایانه‌ها را قادر می‌سازد بهتر به نیازهای اطلاعاتی کاربران انسانی بپردازند (Baker & Cheung, 2007).

وب معنایی امتداد رو به گسترش وب جهانی^۱ است که در آن محتوای وب نه تنها قابل بیان در قالب زبان طبیعی است، بلکه به صورت قابل خواندن و فهم توسط عامل نرم‌افزاری^۲ نیز هست، از این رو عوامل نرم‌افزاری را مجاز به پیدا کردن، اشتراک، و یکپارچه‌سازی اطلاعات به صورت خیلی ساده می‌نماید. وب معنایی حاصل برداشت تیم برنرز لی^۳ از وب به عنوان رسانه‌ای جهانی برای داده‌ها، اطلاعات و تبادل دانش حاصل شده است (Linckels & Meinel, 2011, p. 15).

وب معنایی در پی دو چیز است: اول اینکه قالب‌های مشترک برای یکپارچه‌سازی و ترکیب داده‌های حاصل از منابع نامتجانس، در صورتی که در وب اصلی اکثراً بر روی تبادل اسناد تمرکز می‌شود. از این رو، وب معنایی اغلب به عنوان یک پایگاه اطلاعاتی عظیم یا مجموعه‌ای از پایگاه‌های اطلاعاتی پراکنده نشان داده می‌شود. دوم، این که وب معنایی در پی زبانی برای ثبت چگونگی ارتباط داده‌ها با اشیای واقعی جهان است. این عامل به یک فرد یا ماشین این امکان را می‌دهد که از یک پایگاه اطلاعاتی شروع کرده و سپس از طریق مجموعه‌ای بی‌پایان از پایگاه‌های اطلاعاتی به مسیر خود ادامه دهد، این پایگاه‌های اطلاعاتی از طریق سیم‌ها به هم متصل نیستند، بلکه عامل اتصالشان این است که در مورد یک چیز هستند (Linckels & Meinel, 2011, p. 16).

لینکلز و میل (Linckels & Meinel, 2011, p. 17) اصولی را که وب معنایی از ابتدا بر روی آن تکیه دارد را به صورت زیر بر می‌شمارند:

- وب معنایی در مورد منابع است (و محدود به صفحات وب مانند WWW کلاسیک نیست). یک منبع هر چیزی می‌تواند باشد مانند یک صفحه وب، یک کتاب، یا یک شخصیت انسانی.
- هر منبعی با شناساگر متحدالشکل منابع (یو.آر.آی). مشخص می‌شود. یو.آر.آی.ها رشته‌های اسکی^۵ کوتاهی هستند که منابع را مشخص می‌کنند. از این رو یک یو.آر.آی. می‌تواند آدرس یک وبگاه باشد (یو.آر.ال.)، آی.اس.بی.ان.^۶ یک کتاب باشد، یا کد ملی یک شخص. شناساگرهای بین‌المللی منابع (آی.آر.ای.^۷ها) مکملی برای یو.آر.آی.ها هستند. یک آر.آی.آی. توالی کاراکترهای برگرفته از مجموعه کاراکترهای بین‌المللی (یونی‌کد^۸) هستند که برای مثال جهت استفاده به جای نمادهای چینی در یو.آر.آی. مورد استفاده قرار می‌گیرند.
- منابع موجود در وب معنایی با اطلاعاتی اضافی که فراداده نامیده می‌شوند، توصیف می‌شوند. فراداده‌ها در قالب زبانی رسمی مانند ایکس.ام.ال.^۹، آر.دی.اف.^{۱۰}، یا او.دبلیو.ال.^{۱۱} توصیف می‌شوند.

۶-۲. ابزارهای یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال

¹World wide web

²Software agents

³Tim Berners-Lee

⁴Format

⁵ASCII

⁶ISBN

⁷Internationalized resource identifier (IRI)

⁸Unicode, Universal character set

⁹XML

¹⁰RDF

¹¹OWL

پیش‌نیازهای زیادی از جمله زیر ساخت فناوانه (مانند ساختارها^۱، پروتکل‌ها، روش‌های ترکیبی^۲، طرح‌های کدگذاری^۳، و نظام‌های شناساگر^۴)، استانداردها، مسائل سازمانی و قانونی، خدمات حمایتی مانند ثبت‌های^۵ مختلف و پایگاه‌های دانشی معنایی^۶ و سیستم سازماندهی دانش^۷ (از جمله هستی‌نگاری‌های بنیادی^۸ و هسته‌ای^۹) (Patel, Koch, Doerr, & Tsinaraki, 2005)، برای یکپارچه‌سازی معنایی و محقق ساختن یکپارچه‌سازی معنایی مورد نیاز هستند. از آنجایی که پرداختن به تمامی این ملزومات از اهداف این پژوهش خارج است، در زیر به دو نمونه از مهم‌ترین ملزومات یکپارچه‌سازی معنایی یعنی هستی‌نگاری‌ها و سیستم‌های سازماندهی دانش پرداخته می‌شود.

۱-۶-۲. هستی‌نگاری‌ها

چئونگ و دیگران (Cheung, Smith, Yip, Baker, & Gerstein, 2007) برخی از چالش‌های پیش‌روی یکپارچه‌سازی اطلاعات از منابع مختلف را بر می‌شمارند که از جمله آنها موارد زیر هستند:

- مترادفات^{۱۰}: وجود چندین مترادف برای یک مفهوم
- ابهام^{۱۱}: یک اصطلاح می‌تواند نشان‌دهنده چندین مفهوم باشد برای مثال انسولین می‌تواند جهت بازنمایی مفاهیمی مانند ژن، پروتئین، دارو و غیره به کار رود.

آنها یکی از رویکردهای پرداختن به این مسائل را رویکرد وب معنایی می‌دانند که بر خلاف رویکردهای سنتی (مانند انبارسازی داده‌ها^{۱۲} و نظام اطلاعاتی یکپارچه^{۱۳}) که بر اساس تبدیل مدل‌های داده‌ای مختلف به مدل‌های داده‌ای مشترک بهره می‌گیرند، از هستی‌نگاری استاندارد نیز جهت یکپارچه‌سازی نظام‌های اطلاعاتی مختلف بهره می‌گیرند (Cheung, Smith, Yip, Baker, & Gerstein, 2007).

هستی‌نگاری‌ها به عنوان جزء کلیدی یکپارچه‌سازی معنایی به حساب می‌آیند. آنها به عنوان قالبی برای بازنمایی دانش در مورد جهان یا بخشی از آن به کار می‌روند. سرچشمه اصطلاح هستی‌نگاری در فلسفه است که در آن هستی‌نگاری نام یک شاخه اصلی متافیزیک است که با تحلیل انواع یا اشکال وجود^{۱۴} سروکار دارد (Linckels & Meinel, 2011, p. 18). اساساً، یک هستی‌نگاری به طور رسمی مفاهیم و روابطی که ممکن است بین آنها در برخی جوامع وجود داشته باشد، توصیف می‌کنند. به عبارت دیگر یک هستی‌نگاری بخشی از جهان را توصیف می‌نماید (Synak, Dabrowski, & Kruk, 2009, p. 41).

دانش ساختاری که در درون هستی‌نگاری‌ها قرار گرفته است به بازیابی و تبادل اطلاعات کمک می‌کند. هستی‌نگاری‌ها همچنین به بررسی تطابق بین عناصر منابع داده‌ای نامتجانس کمک می‌کنند. پژوهش‌های صورت گرفته در زمینه هستی‌نگاری‌ها پیشنهاد می‌کنند که اگر در توسعه اجزای نظام‌های اطلاعاتی به هنگام پیاده‌سازی هستی‌نگاری‌ها، آنها را در قالب سلسله مراتب‌های

¹Architecture

²Syntactic solution

³Encoding schemes

⁴Identifier systems

⁵Registries

⁶Semantic knowledge bases

⁷Knowledge organization system(KOS)

⁸Fundamental ontology

⁹Core ontology

¹⁰Synonyms

¹¹Ambiguity

¹²Data warehousing

¹³Federated

¹⁴existence

شیء^۱ گرا^۱ در آوریم، منجر به اعتلای مفهوم سیستم‌های اطلاعاتی مبتنی بر هستی‌نگاری می‌شود (Shen R. , Applying the 5S Framework To Integrating Digital Libraries, 2006)
هر هستی‌نگاری از بخش‌های زیر تشکیل شده است (Synak, Dabrowski, & Kruk, 2009, p. 42):

- رده^۲: که گاهی نیز مفهوم نامیده می‌شود، از طریق آن در هستی‌نگاری کیفیات و ویژگی‌های عمومی گروهی از اشیاء نشان داده می‌شوند. اگر گروهی از اشیاء خصایص یکسان داشته باشند، این حقیقت باید شناخته شده و رده‌ای جهت بازنمایی این خصلت‌ها ایجاد شود. ایده خوب این است که اول گروه‌های اشیاء را شناخته و سپس رده‌هایی برای آنها ایجاد نماییم. برای مثال اگر قد برای ما مهم باشد، می‌توانیم تعریف کنیم که تمامی افرادی که قدشان بالاتر از ۱۸۰ است در رده‌ای تحت عنوان بلند قد دسته‌بندی شوند.
- رده فرعی^۳: بخشی از گروه اشیاء را نشان می‌دهد که ویژگی‌هایی دارند که این ویژگی‌ها در تمامی اعضای گروه مشترک نیست. برای مثال اگر ما رده‌ای از "اعضای کارمندان" داشته باشیم که متشکل از تمامی افرادی که در شرکت ما کار می‌کنند، باشد، ما می‌توانیم یک رده فرعی تحت عنوان "مدیریت" ایجاد کنیم که شامل همه افرادی است که در بخش مدیریت کار می‌کنند.
- فرد یا شیء^۴: هر فرد یا شیء یک مورد مجرد در دنیای ماست که ممکن است متعلق به برخی یا همه رده‌ها باشد. برای مثال رونالد ریگان می‌تواند فردی از رده رئیس‌جمهوران آمریکا باشد. دو فرد مجزا دو شیء مجزا هستند. ما باید توجه داشته باشیم که هر مورد ممکن است بازنمون مفاهیم انتزاعی (مانند فعالیت‌ها) یا اشیاء عینی باشند.
- ویژگی^۵: جهت توصیف کیفیات یکسان همه موارد افراد موجود در یک رده به کار می‌رود. ویژگی‌ها بازنمون روابط در هستی‌نگاری‌ها هستند. وقتی که ویژگی‌ای را به یک رده نسبت می‌دهیم آن رده حیطه^۶ آن ویژگی می‌شود که گستره آن اشیاء موجود در آن رده است.
- تحدید ویژگی^۷: می‌تواند حالت‌های بیشتری به ویژگی‌ها دهد. برخی از تحدیدهای ویژگی رایج تحدیدهای ارزشی و اهمیتی باشند. همچنین ممکن است بخواهیم تعیین کنیم که برخی ویژگی‌ها مورد نیاز هستند و غیره.

برحسب گستره و حوزه کاربرد، هستی‌نگاری‌ها قابل تقسیم به دو دسته عمومی^۸ و تخصصی^۹ هستند. دسته اول هستی‌نگاری عمومی (یا هستی‌نگاری جهانی^{۱۰}، هستی‌نگاری برتر^{۱۱}، هستی‌نگاری بنیادی^{۱۲}) و مدلی از اشیاء یکسان است که به‌صورت کلی در بسیاری از حوزه‌ها قابل کاربرد است. این نوع هستی‌نگاری حاوی واژگان اساسی است که در آنها اصطلاحات در مجموعه‌ای از حوزه‌ها قابل تعریف هستند برخی از انواع این نوع هستی‌نگاری‌ها اوپن سایک^{۱۳}، سومو^{۱۴}، جی.او.^{۱۵}، وُرد نت^۱ می‌باشند. دسته دو

¹Object oriented hierarchies

²class

³Sub class

⁴Individual or objec

⁵Property or attribute

⁶Domain

⁷Property restriction

⁸Upper ontology

⁹Domain ontology

¹⁰World ontology

¹¹Top ontology

¹²Foundation

¹³Open Cyc

¹⁴Suggested upper merged ontology(SUMO)

¹⁵Gene ontology(GO)

هستی‌نگاری تخصصی^۲ می‌باشد که حوزه‌ای خاص یا بخشی از جهان را مدل‌سازی می‌کند. این نوع هستی‌نگاری معنای خاص اصطلاحاتی را نشان می‌دهد که در یک حوزه کاربرد دارند. هستی‌نگاری موسیقی، اصطلاحنامه موسسه ملی سرطان^۳، و هستی‌نگاری فوتبال^۴ نمونه‌هایی از این نوع هستی‌نگاری‌ها هستند (Linckels & Meinel, 2011, p. 20). گزارش پروژه دلاس (۲۰۰۵) دسته‌سومی نیز از هستی‌نگاری‌ها بر می‌شمارد که عبارتند از هستی‌نگاری‌های هسته‌ای (یا میانی^۵) که در اصل همان هستی‌نگاری عمومی هستند ولی در حوزه کاربردی گسترده (مانند حوزه سمعی بصری^۶). این نوع هستی‌نگاری متشکل از اصطلاحات و روابطی است که به نظر می‌رسد که در بافت حوزه کاربردی گسترده (مانند واقع^۷ در حوزه سمعی بصری) ضروری هستند.

یکی از فنون رایج در پردازش زبان طبیعی^۸ استفاده از مجموعه آموزشی^۹ یا کورپوس^{۱۰} می‌باشد. در کورپوس‌ها مقادیر زیادی از متون مانند روزنامه‌ها مورد بررسی قرار گرفته و جهت ایجاد داده‌هایی آماری در مورد استفاده از کلمات به کار می‌روند. این امر ممکن است شناسایی الگوهای حاکم در جملات را که به تفسیر صحیح زبان کمک می‌کند، مقدور نماید. چنین رویکردهای آماری به پردازش زبان طبیعی در کانون تمرکز زبان‌شناسی رایانشی^{۱۱} قرار دارد (Linckels & Meinel, 2011, p. 62). با توجه به ماهیت کورپوس‌ها، آنها را نیز می‌توان به نوعی هستی‌نگاری قلمداد نمود و در تسهیل تعامل معنایی کتابخانه‌های دیجیتال مورد استفاده قرار داد.

۲-۶-۲. سیستم‌های سازماندهی دانش (کی.او.اس.)^{۱۲}

افراد اغلب به جستجوی موضوعات، مفاهیم، یا ایده‌ها می‌پردازند و این نوع جستجوها سخت‌ترین انواع جستجو هستند. اغلب کاربر تنها می‌خواهد با تایپ چند کلمه منظور خود را برساند، اما این کلمات همیشه گویای آن چیزی نیستند که در ذهن کاربر است. سایر اوقات، کاربر با کمبود کلمه مواجه می‌شود. کتابخانه‌های دیجیتال باید در این درخواست کاربر برای معنی، پشتیبان وی باشند. این امر یکی از نقش‌های اصلی سیستم‌های سازماندهی دانش است. در برخی از سیستم‌های اطلاعاتی چنین جستجوهای به-وسیله فهرست‌نویسی یا نمایه‌سازی موضوعی دستی، رایانه‌ای یا خودکار پشتیبانی می‌شوند. در این حالت واژگان کنترل شده برای اصطلاحات یا سایر مفاهیم اختصاص یافته و در سیستم سازماندهی دانش گنجانده می‌شود (Soergel, 2009, p. 27). اصطلاح سیستم‌های سازماندهی دانش به واژگان کنترل شده و نیز سیستم‌ها، ابزارها یا خدماتی اشاره دارد که به منظور سازماندهی دانش ایجاد شده‌اند (ابزارهایی که تفسیر سازماندهی شده‌ای از ساختار دانش ارائه می‌نمایند) (Patel, Koch, Doerr, & Tsinaraki, 2005).

ایجاد و بهبود تعامل معنایی همواره نیازمند استفاده از واژه‌ها یا بهتر است بگوییم سیستم‌های سازماندهی دانش است. گاهی نیازمندیم که ابتدا این سیستم‌ها را ایجاد (یا استخراج) نماییم. در سایر موارد واژگان موجود را باید تغییر شکل، مطابقت، ترکیب یا

¹WordNet

²Domain ontology

³National cancer institute thesaurus

⁴Soccer ontology

⁵intermediate

⁶Audio visual

⁷event

⁸Natural language processing (NLP)

⁹Training set

¹⁰corpus

¹¹Computational linguistics

¹² Knowledge Organization Systems (KOS)

شبیه‌سازی نماییم تا به این واژگان و یا اسناد یا پایگاه‌های اطلاعاتی قابلیت تعامل معنایی بدهیم. این امر بویژه زمانی اهمیت دارد که سیستم‌های سازماندهی دانش از نظر ساختار، حوزه، زبان یا تکینگی^۱ با یکدیگر متفاوت باشند (Patel, Koch, Doerr, & Tsinaraki, 2005).

برخی از کارکردهای سیستم‌های سازماندهی دانش عبارتند از (Soergel, 2009, p. 29):

- نقشه راه معنایی برای رشته‌های جداگانه و روابط میان رشته‌ها هستند.
- باعث بهبود ارتباطات شده و در امر یادگیری مفید بوده و اطلاعات را تلفیق می‌نمایند.
- پایه مفهومی برای طراحی پژوهش خوب و پیاده‌سازی آن فراهم می‌نمایند.
- رده‌بندی را در عمل و با اهداف سیاسی و اجتماعی ارائه می‌کنند.
- به بازیابی و تحلیل اطلاعات کمک می‌کنند.
- هستی‌نگاری‌ای برای تعریف اجزای داده‌ها و فرهنگی برای آنها هستند.
- پایه مفهومی برای سیستم‌های دانش‌بنیان فراهم می‌کنند.
- کارکردهای آن به چندین زبان قابل ارائه هستند.
- فرهنگی یک یا دو یا چند زبانه برای استفاده افراد ارائه می‌کنند.
- پایگاه دانشی لغوی برای پردازش زبان طبیعی هستند.

سیستم‌های سازمان دهی دانش انواع مختلفی دارند که عبارتند از (Patel, Koch, Doerr, & Tsinaraki, 2005):

• فهرست‌های اصطلاحات^۲

منظور از فهرست‌های اصطلاحات، فهرست‌ها و پایگاه‌هایی هستند که اصطلاحات و واژه‌های مورد نیاز سیستم از قبل در سیستم اطلاعاتی دانش استقرار می‌یابند و بستر اصلی برقراری روابط معنایی را فراهم می‌نمایند. انواع این فهرست‌ها به شرح زیر هستند:

0 فایل‌های مستند (مانند فایل مستند نام کتابخانه کنگره^۳)

0 واژه‌نامه‌ها^۴ (مانند واژه‌نامه‌های موجود در انتهای منابع مرجع یا کتاب‌ها که واژه‌های مورد نیاز را تعریف می‌کنند. این واژه‌ها در قالب پایگاه اطلاعاتی در سیستم اطلاعاتی دانش ارائه می‌شوند).

0 جاینامه‌ها^۵ (مانند انواع جاینامه‌های موجود نظیر جاینامه‌های جغرافیایی)

0 فرهنگ‌های لغت (مانند انواع فرهنگ‌های لغت یک زبانه یا چندزبانه موجود)

• رده‌بندی و دسته‌بندی

0 سرعنوان‌های موضوعی (مانند سرعنوان‌های موضوعی پزشکی مش^۶)

0 طرح‌های رده‌بندی، تاکسونومی‌ها^۷ و طرح‌های دسته‌بندی (مانند رده‌بندی دهدهی دیویی^۸، تاکسونومی‌های

علمی)

• طرح‌های ارتباط^۱ (عنوان بهتر طرح‌های رابطه‌ای)

¹Granularity

²Term lists

³LC name authority file

⁴Glossary

⁵Gazetteer

⁶Medical subject headings(MESH)

⁷Taxonomies

⁸Dewey Decimal Classification(DDC)

0 اصطلاح نامه‌ها^۲ (مانند اصطلاحنامه‌های موجود که روابط بین اصطلاح‌ها و اصطلاح‌های مرجع و نامرجع را نشان می‌دهند).

0 شبکه‌های معنایی^۳ (مانند وردنت که روابط بین کلمات را بیان می‌کنند. کورپوس‌ها نیز که مجموعه‌ای از کلمات و روابط آنها را با استفاده از تلفیق محتواهای بسیار زیادی مانند مقالات و روزنامه‌ها نشان می‌دهند).

0 هستی‌نگاری‌ها (که نشانگر انواع روابط بین کلمات و اصطلاحات هستند).

به طور کلی سیستم‌های اطلاعاتی که با بهره‌گیری از قابلیت‌های مذکور تهیه می‌شوند هدفشان نشان دادن روابط معنایی بین محتواها و ارائه قابلیت‌هایی برای بازیابی معنایی و خدمات دانشی است. برای این منظور ساختار یک سیستم سازماندهی دانش جامع در برگیرنده دو سطح زیر است (Soergel, 2009, p. 31)

سطح ۱. روابط مفهوم-اصطلاح^۴

سطح ۲. ساختار مفهومی^۵

۲-۱. تحلیل معنایی و چهریزه‌ها^۶

۲-۲. سلسله مراتب^۷

۲-۳. تعامل سلسله مراتب و چهریزه‌ها^۸

۲-۴. روابط مفهومی متمایز شده^۹ (یا پالایش شده)

عناصر موجود در سطح یک نیاز به کنترل واژگان در سطح نمایه‌سازی (استفاده از واژگان کنترل شده) و در سطح جستجو (گسترش اصطلاح درخواستی (کلیدواژه)) به ویژه در جستجوی آزاد و برچسب‌هایی که کاربران تخصیص می‌دهند (برچسب‌های اجتماعی) دارند. در این سطح لازم است تا واژگان انتخاب شده جهت استفاده در سطح دوم مورد بررسی و تایید قرار گیرند. بر این اساس واژه‌ها و اصطلاحات موجود مورد ممیزی قرار می‌گیرند. این سطح به ویژه در سیستم‌های دانشی که مبتنی بر نمایه تمام متن و یا جستجوهای آزاد هستند اهمیت ویژه‌ای دارد.

عناصر سطح دوم به سازماندهی مفهومی می‌پردازند. دو اصلی که برای تعامل ساختار مفهومی مورد نیازند یعنی تحلیل چهریزه‌ها (تحلیل اجزاء، تحلیل ویژگی، تحلیل جنبه، عاملیت معنایی) و سلسله مراتب، در این سطح مورد توجه قرار می‌گیرند.

یکی از سیستم‌های سازماندهی دانش شناخته شده سیستم ساده سازماندهی دانش^{۱۰} است. سیستم ساده سازماندهی دانش به عنوان پل ارتباطی بین سیستم‌های سنتی سازماندهی دانش مانند اصطلاحنامه‌ها، طرح‌های رده‌بندی و نظایر آنها و ساختارهای مبتنی بر وب مانند دایرکتوری‌های باز^{۱۱} محسوب می‌شود (Miles, Matthews, Wilson, & Brickley, 2005).

کارگروه وب معنایی کنسرسیوم وب، سیستم ساده سازماندهی دانش (اسکاس) را در سال ۲۰۰۱ با هدف توزیع و انتشار اصطلاحنامه‌ها، کنترل واژگان‌ها، سرعنوان‌های موضوعی، رده‌بندی‌های کتابخانه‌ای و سایر طبقه‌بندی‌های استاندارد در وب بنیان نهاد. اسکاس از آر.دی.اف. استفاده می‌کند. بر این اساس هر نوع واژگانی که از اسکاس استفاده می‌کند در محیط وب می‌تواند با

¹Relationship schemes

²Thesauri

³Semantic networks

⁴Concept-term relationship

⁵Conceptual structure

⁶Facet

⁷Hierarchy

⁸Interaction of hierarchy and facets

⁹Differentiated (refined) concept relationship

¹⁰ Simple Knowledge Organization System (SKOS)

¹¹ Open directories

هر نوع منبع اطلاعاتی دیگر که از آر.دی.اف. استفاده می‌کند ارتباط برقرار کند. این امر، بازیابی اطلاعات در پایگاه‌های اطلاعاتی مختلف را بدون نیاز به یکپارچه‌سازی فیزیکی بانک‌های اطلاعاتی آنها به آسانی میسر می‌کند (حسن زاده & حسینی، ۱۳۹۱). اسکاس از چهار ساختار اصلی آر.دی.اف.، ایکس.ام.ال.، یو.آر.آی. و عناصر فراداده‌ای تشکیل شده است. عناصر فراداده‌ای اسکاس در ۶ بخش عناصر مفهومی^۱، برچسب‌های خصایص^۲، خصائص سند^۳، روابط معنایی^۴، مجموعه معنادار مفاهیم^۵، و خصائص نمایه‌های موضوعی^۶ دسته‌بندی می‌شود.

اسکاس، طبقاتی می‌سازد و در این طبقات روابط هر کدام را قرار می‌دهد و بر اساس آن خصائص آنها را مشخص می‌کند و هر کدام از این خصائص را به واژگانی از مفاهیم اختصاص می‌دهد و چارچوب‌ها و قلمرو علاقه‌مندی آنها را مشخص می‌کند. در ساختار اسکاس هر طبقه با برچسب‌هایی خاص تعریف می‌شود. ارتباط سلسله مراتبی در اصطلاحنامه با مفاهیم اعم و اخص و مرتبط برقرار می‌شود. در این ساختار برای هر مفهوم برچسب‌های معادل ایجاد شده است. برخی از این برچسب‌ها عمان‌هایی هستند که در اصطلاحنامه‌ها نیز مورد استفاده قرار می‌گیرند اما برخی دیگر برای سازگاری و کاربردپذیری بیشتر ساختار اصطلاحنامه با و معنایی ایجاد شده‌اند. عمده‌ترین این برچسب‌ها به شرح جدول ۱-۲ هستند:

جدول ۱-۲. عمده‌ترین برچسب‌های اسکاس و توضیحات آنها (حسن زاده & حسینی، ۱۳۹۱)

کاربرد	برچسب اسکاس
مفهوم اصطلاحنامه‌ای	Skos:Concept
اصطلاح مرجع در اصطلاحنامه	SKOS:PrefLable
اصطلاح بکار رفته در اصطلاحنامه	SKOS:AltLable
نماد مرجع در اصطلاحنامه	SKOS:PrefSymbol
نماد بکار رفته در اصطلاحنامه	SKOS:AltSymbol
توضیحات مربوط به اصطلاح	SKOS:ScopeNote
تعاریف اصطلاح	SKOS:Definition
اصطلاح اعم	SKOS:Broader
اصطلاح اخص	SKOS:Narrower
اصطلاح مرتبط	SKOS:Related
منبع اصطلاحنامه	SKOS:InScheme

۲-۲. زبان‌های مورد استفاده در یکپارچه‌سازی معنایی

¹ Conceptual elements: SKOS: concept

² Labeling properties: SKOS: HiddenLable, SKOS:AltLable, SKOS:PrefLable, SKOS: PrefSymbol, SKOS:AltSymbol

³ Documentation properties: SKOS:Note, SKOS:Definition, SKOS:ScopeNote, SKOS:Example, SKOS:HistoryNote, SKOS:EditorialNote, SKOS:ChangeNote

⁴ Semantic relationship: SKOS:Related, SKOS:Narrower, SKOS:Broader

⁵ Meaningful collection of concepts: SKOS:Collection

⁶ Subject indexing properties: SKOS:Subject, SKOS:PrimarySubject

برای بیان روابط معنایی و ایجاد آنتولوژی‌ها نیاز به استفاده از زبان‌های ماشینی و بهره‌گیری از آنها برای برقراری روابط معنایی در سیستم‌های اطلاعاتی و ابزارهای مورد استفاده در آنها است. برای این منظور تا کنون زبان‌های متعددی با کاربردها و کارایی‌های متعددی مورد استفاده قرار گرفته شده‌اند که در ادامه به برخی از مهم‌ترین و کارآمدترین آنها می‌پردازیم.

۱-۲-۲. زبان توصیفی ایکس.ام.ال.^۱

زبان‌های نشانه‌گذاری عمومیت‌یافته^۲ در دهه‌های ۶۰ و ۷۰ میلادی پا به عرصه وجود نهادند. هدف از این زبان‌ها، سازماندهی اطلاعات اسناد بر اساس نشانه‌گذاری‌های انجام گرفته توسط نویسندگان بود. لذا تمامی این اسناد برای شروع سند، نیاز به تعریفی داشتند تا به استفاده‌کننده از سند خواه انسان، خواه ماشین، اطلاعات لازم برای شروع رمزگشایی نشانه‌ها را بدهند. با گذشت زمان و گسترش حوزه نشر، نیاز به توسعه این زبان‌ها و استانداردسازی آنها پیش آمد. از این رو، زبان نشانه‌گذاری عمومیت‌یافته استاندارد (اس.جی.ام.ال.)^۳ پا به عرصه وجود نهاد. اس.جی.ام.ال. زبان توصیفی^۴ است که چگونگی نشانه‌گذاری یک سند را نشان نمی‌دهد، اما به برنامه‌های کاربردی و محیط‌های مستقل اجازه کار می‌دهد. بنابراین هدف اولیه آن ارائه زبانی استاندارد برای نشانه‌گذاری ساختار منطقی مدارک مستقل از سخت‌افزار و نرم‌افزار بود، با این هدف که مبادله مدارک را در حلقه مراحل انتشار تسهیل کند. اسناد با استفاده از این زبان توصیفی، قابل انتقال گشته و امکان اشتراک داده‌ها و بسته‌بندی چندباره^۵ اطلاعات فراهم می‌گردد. با این همه، بسیار پیچیده بوده و اجرا و نگهداری اسناد با استفاده از آن بسیار گران است. از طرف دیگر، نرم‌افزارهای پشتیبان این استاندارد هم گران و پیچیده‌اند و نیز بدلیل عدم پشتیبانی از روند کلی مرورگرها، با محیط وب سازگار نیستند (زارع زاده، ۱۳۸۳). لذا نیاز به زبانی بود تا عاری از این محدودیت‌ها و مشکلات باشد. به این منظور، زبان نشانه‌گذاری گسترش‌پذیر (ایکس.ام.ال.) بوجود آمد.

ایکس.ام.ال. برای انتشار و مبادله اطلاعات ساخت‌یافته در اینترنت و اینترنت مورد استفاده قرار گرفته است و همچنین به طراحان امکان می‌دهد تا برچسب‌های خود را تعریف کنند. بنابراین هدف اولیه آن را تسهیل اشتراک اطلاعات میان سیستم‌های مختلف، بویژه سیستم‌هایی که به اینترنت متصل هستند، و مهم‌تر از همه کاربرد در برقراری روابط معنایی می‌دانند. از این رو، می‌توان گفت که ایکس.ام.ال. برای تسهیل امکان مبادله اطلاعات (نحوی و معنایی) بین برنامه‌های کاربردی در وب بوجود آمده است. کنسرسیوم وب گسترده جهانی، اهداف ده‌گانه زیر را برای طراحی ایکس.ام.ال. بیان می‌کند:

۱. ایکس.ام.ال. باید مستقیماً از طریق اینترنت قابل استفاده باشد.
۲. ایکس.ام.ال. باید انواع مختلف برنامه‌های کاربردی را پشتیبانی کند.
۳. ایکس.ام.ال. باید با اس.جی.ام.ال. سازگار باشد.
۴. ایکس.ام.ال. باید برای نوشتن برنامه‌هایی که اسناد ایکس.ام.ال. را پردازش می‌کنند، آسان باشد.
۵. تعداد اجزای اختیاری در ایکس.ام.ال. باید به حداقل برسد.
۶. اسناد ایکس.ام.ال. باید واضح و قابل خواندن توسط انسان باشند.

^۱ XML (eXtensible Markup Language)

^۲ Generalized Markup Languages (GML)

^۳ Standard Generalized Markup Language (SGML)

^۴ زبان توصیفی (metalanguage) زبانی است که یک زبان دیگر را توصیف می‌کند. مثلاً فعل، اسم، صفت و قید کلماتی هستند که جزو دستور زبان محسوب می‌شوند و برای توصیف کلمات دیگر بکار می‌روند. مزیت یک زبان توصیفی در این است که می‌توان از آن برای توصیف زبانهای مختلف استفاده کرد (اشباکر، ۱۳۸۰).

^۵ Repackage

۷. فایل ایکس.ام.ال. باید به سرعت طراحی شود.

۸. فایل ایکس.ام.ال. باید دقیق و ساده باشد.

۹. ایجاد اسناد ایکس.ام.ال. باید آسان باشد.

۱۰. اختصار در نشانه گذاری ایکس.ام.ال. از حداقل اهمیت برخوردار است.

با توجه به موارد فوق می‌توان به سادگی، دقت، قابلیت سازگاری با اس.جی.ام.ال. و برنامه‌های کاربردی، و قابل استفاده بودن از طریق اینترنت اشاره کرد.

ایکس.ام.ال. برای حروف از استانداردهای یونی کد و ISO/IEC 10646، برای برچسب‌های شناسایی زبان از استاندارد IETF RFC 3066، برای کدهای نام زبان‌ها از استاندارد ISO 639 و برای کد نام کشورها از استاندارد ISO 3166 استفاده می‌کند.

نهایت اینکه با توجه به موارد اشاره شده می‌توان اظهار داشت که ایکس.ام.ال. یک زبان توصیفی مورد حمایت کنسرسیوم وب گسترده جهانی است که در نوامبر سال ۱۹۹۶ به جامعه کاربران معرفی شد. این زبان شکل محدود شده اس.جی.ام.ال. است که با هدف انتشار، ذخیره و مبادله اطلاعات ساخت یافته در اینترنت و یا اینترنت بوجود آمده است و به طراحان امکان تعریف و استفاده از برچسب‌های خود را می‌دهد.

برخی آن را مانند چتری می‌دانند که برنامه‌های کاربردی دیگر نظیر ایکس.اس.ال.^۱، ایکس.لینک^۲، ایکس.پس^۳ و غیره را احاطه کرده است (Brown, Byard, & Hart, 2001) (Desmarais, 2000) (Gunn, 1999). به هر حال ایکس.ام.ال. دارای ویژگی‌هایی از قبیل خود مستندی^۴، توصیف داده‌ها در اسناد به صورت سلسله مراتبی، ایجاد برچسب‌ها توسط کاربر، توانایی انطباق با سایر برنامه‌های کاربردی و غیره است که آن را نسبت به سایر محصولات منحصر به فرد می‌سازد.

همان‌طور که می‌دانیم ایکس.ام.ال. قادر است محتوای اسناد و نیز خود اسناد را با ساختار سلسله مراتبی ذخیره و نمایش دهد.

```
<book>
  <booktitle> The Selfish Gene </booktitle>
  <author id = "dawkins">
    <name>
      <firstname> Richard </firstname>
      <lastname> Dawkins </lastname>
    </name>
    <address>
      <city> Timbuktu </city>
      <zip> 99999 </zip>
    </address>
  </author>
</book>
```

شکل ۲-۷. چیدمان سلسله مراتبی ایکس.ام.ال.

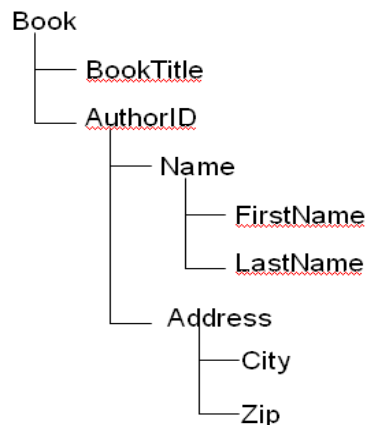
بنابراین چیدمان سلسله مراتبی مثال فوق را می‌توان به شکل زیر نمایش داد:

^۱ XSL

^۲ Xlink

^۳ Xpath

^۴ یکی از مزایای مهم ایکس.ام.ال. است که با یک نگاه می‌توان پی برد که چه کاری انجام می‌دهد. (اشباکر، ۱۳۸۰، ص ۱۹)



شکل ۲-۸ چیدمان درختی ایکس.ام.ال.

با توجه به این که کاربر خود می‌تواند برچسب‌ها را در ایکس.ام.ال. تعریف کند، لذا نیاز به چیزی دارد تا بتواند این برچسب‌ها را تعریف نماید. در این صورت است که برنامه‌های کاربردی می‌توانند این برچسب‌ها و چیدمان آن سند را شناسایی و مورد استفاده قرار دهند. دی.تی.دی. ابزاری است که این وظیفه را بر عهده دارد. دی.تی.دی. چیدمان یک سند ایکس.ام.ال. را توصیف می‌کند و چیدمان یک عنصر ایکس.ام.ال. را با مشخص کردن نام عناصر فرعی و خصیصه‌های آنها مشخص می‌کند (Waldt, 2005).

```

<!ELEMENT book (booktitle, author)
<!ELEMENT article (title, author*, contactauthor)>
<!ELEMENT contactauthor EMPTY>
<!ATTLIST contactauthor authorID IDREF IMPLIED>
<!ELEMENT monograph (title, author, editor)>
<!ELEMENT editor (monograph*)>
<!ATTLIST editor name CDATA #REQUIRED>
<!ELEMENT author (name, address)>
<!ATTLIST author id ID #REQUIRED>
<!ELEMENT name (firstname?, lastname)>
<!ELEMENT firstname (#PCDATA)>
<!ELEMENT lastname (#PCDATA)>
<!ELEMENT address ANY>
  
```

شکل ۲-۹. دی.تی.دی. یک فایل ایکس.ام.ال.

دی.تی.دی. می‌تواند به عنوان جزئی از یک فایل ایکس.ام.ال. باشد. بدین معنی که دی.تی.دی. مربوط به آن سند در ابتدای سند قرار می‌گیرد و برنامه‌های کاربردی برای نمایش یا مبادله سند ابتدا دی.تی.دی. مربوطه را ترجمه می‌کنند. در مراکز اطلاعاتی بزرگ مانند کتابخانه‌های دیجیتال که اسناد بسیار زیادی در آنها نگهداری و یا مبادله می‌شوند نیاز به تعریف یک دی.تی.دی. مشترک برای اسناد وجود دارد تا همه اسناد بتوانند از آن استفاده کنند. لذا برای این کار می‌توان دی.تی.دی. را در یک فایل خارجی ذخیره کرد و در مواقع مورد لزوم در فایل‌های ایکس.ام.ال. آن را فراخوانی کرد.

حال که با چیدمان فایل‌های ایکس.ام.ال. آشنا شدیم باید بدانیم که برای استفاده از آنها در محیط وب نیاز به استفاده از برنامه‌ها و زبان‌های کاربردی دیگری است تا بتوان این اسناد را به شکل مطلوب در اختیار کاربر قرار داد. سال‌ها پیش وقتی که وب هنوز جوان بود، و کاربران نیز توقع کمی داشتند، صفحات ایستا^۱ همه را راضی می‌کرد. اما اکنون برای جلب (و حفظ) مخاطب، پویایی به یک عنصر حیاتی تبدیل شده است. لذا برای نمایش فایل‌های ایکس.ام.ال. نیاز به کمک گرفتن از سایر برنامه‌ها و زبان‌ها مثل ایکس.اس.ال.^۲، اچ.تی.ام.ال.^۳، جاوا اسکریپت^۴ و غیره است. ایکس.اس.ال. می‌تواند با چند کد، فایل، پایگاه داده ایکس.ام.ال. را خوانده و سپس با استفاده از این اطلاعات صفحات اچ.تی.ام.ال. پویا ایجاد کرده و به رایانه کاربر ارسال کند. بدین ترتیب هر برنامه سه فایل خواهد داشت: یک فایل داده ایکس.ام.ال.، یک فایل الگوی ایکس.اس.ال. و یک فایل اچ.تی.ام.ال. و یا جاوا اسکریپت برای نمایش داده‌ها.



شکل ۲-۱۰. مجموعه عناصر مورد استفاده در استفاده از ایکس.ام.ال.

ایکس.ام.ال. زبانی است که با توجه به چیدمان سلسله مراتبی که دارد بیشترین کاربرد را در برقراری روابط معنایی می‌تواند ایجاد کند. این زبان همچنین به خاطر اینکه زبانی توصیفی است می‌تواند بیشترین استفاده را در ابزارهای واسط برقراری روابط معنایی و ایجاد هستی‌نگاری‌های داشته باشد. از این رو به جرأت می‌توان گفت که تمامی ابزارهای مبادله اطلاعات و بویژه مبادله معنایی اطلاعات مانند استانداردهای فراداده، ابزارهای تولید هستی‌نگاری مانند پروتزه^۵ از این زبان برای اهداف خود استفاده می‌کنند.

۲-۷-۲. آر.دی.اف.

آر.دی.اف. (چارچوب توصیف منبع)^۶ زبانی برای شناسایی روابط بین اشیاء با استفاده از یو.آر.آی. مبتنی بر وب برای شناسایی اشیاء اطلاعاتی و اصطلاح‌های بکار رفته برای توصیف آنها است (Baker T., 2012). هدف آر.دی.اف. ایجاد فرآیندی برای توصیف منابع به شیوه‌ای ماشین خوان است. مبنای کار آر.دی.اف. ارائه الگویی برای بازنمود ویژگی‌های منابع و مقادیر آنها است. الگوی آر.دی.اف. سه جزء دارد که به شرح زیر هستند:

¹ Static

² XSL

³ HTML

⁴ Java script

⁵ Protege

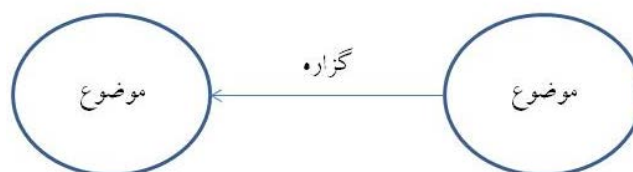
⁶ RDF (Resource Description Framework)

۱. منابع. منابع همه عناصری هستند که توسط آر.دی.اف. توصیف می‌شوند. منبع می‌تواند یک کتاب، فصلی از یک کتاب، بخشی از یک فصل، یک پاراگراف، یک جمله و ... باشد. این کتاب در محیط وب از طریق یو.آر.ال.ها یا سایر شناسه‌های منحصربفرد مانند دی.او.آی.^۱ شناسایی می‌شوند.

۲. ویژگی‌ها. ویژگی‌ها عبارت است از وجه‌میزه، صفت، خاصیت یا رابطه که برای توصیف یک منبع استفاده می‌شود. هر ویژگی معنایی مشخصی دارد که مقدار یا ارزش مجاز آن و نوع منابعی که می‌تواند توصیف کند و نیز رابطه آن را با سایر ویژگی‌ها تعریف می‌کند. این ویژگی‌ها شبیه فراداده‌ای هستند که برای توصیف و تحلیل منبع اطلاعاتی مورد استفاده قرار می‌گیرند.

۳. جملات. یک منبع خاص به همراه ویژگی مشخص به انضمام مقدار آن ویژگی برای آن منبع شکل دهنده یک جمله آر.دی.اف. است. جمله آر.دی.اف. سه جزء مبتدا (فاعل)، خبر (مسند) و مفعول دارد. این سه جزء به ترتیب همان منبع، ویژگی و مقدار یا ارزش ویژگی هستند.

در آر.دی.اف. آن چیزی مهم است که جمله آر.دی.اف. به صورتی بیان شود که برای همه یک معنا را تداعی کند. مثلاً کلمه خالق ممکن است برای افراد مختلف مفاهیم متعددی داشته باشد. برای رفع این مشکل از فرامای آر.دی.اف.^۲ استفاده می‌شود. معنای کلمه مورد استفاده در آر.دی.اف. از طریق ارجاع به یک فرامای بیان می‌شود و می‌توان فرامای آر.دی.اف. را همانند یک واژه‌نامه در نظر گرفت که واژه‌های به کار رفته در آر.دی.اف. را تعریف کرده و به آنها معنای خاصی می‌بخشد. فرامای آر.دی.اف. دارای چیدمان سلسله‌مراتبی و رده‌ای است. رده‌ها از طریق زیررده‌ها قابل بسط هستند (یوسفی راد، ۱۳۸۸).



شکل ۲-۱۱. چیدمان جمله‌های آر.دی.اف.

مدل داده‌ای آر.دی.اف. شبیه رویکردهای کلاسیک مدل‌سازی مفهومی مانند موجودیت-رابطه^۳ است که مبتنی بر ایده اولیه جملات درباره منابع به شکل عبارات موضوع-گزاره-شیء است. این عبارات به عنوان سه گانه‌ها در آر.دی.اف. شناخته می‌شوند. موضوع نشان دهنده منبع، گزاره نشانگر صفات یا جنبه‌های منابع و نشانگر رابطه میان موضوع و شیء است. به عنوان مثال یک روش برای نمایش مفهوم "آسمان به رنگ آبی است" در آر.دی.اف. به این شکل است: موضوع نشانگر "آسمان"، گزاره نشانگر "به رنگ" و شیء نشانگر آبی. بنابراین آر.دی.اف. ارتباط شیء را با موضوع برقرار می‌کند (Manola & Miller, 2013). واژگان آر.دی.اف. به شرح زیر است:

Classes

- **rdf:XMLLiteral** - the class of XML literal values
- **rdf:Property** - the class of properties
- **rdf:Statement** - the class of RDF statements
- **rdf:Alt**, **rdf:Bag**, **rdf:Seq** - containers of alternatives, unordered containers, and ordered containers (rdfs:Container is a super-class of the three)
- **rdf:List** - the class of RDF Lists
- **rdf:nil** - an instance of rdf:List representing the empty list

¹ DOI (Digital Object Identifier)

² RDFS (Resource Description Framework Schema)

³ Entity-Relationship

- **rdfs:Resource** - the class resource, everything
- **rdfs:Literal** - the class of literal values, e.g. [strings](#) and [integers](#)
- **rdfs:Class** - the class of classes
- **rdfs:Datatype** - the class of RDF datatypes
- **rdfs:Container** - the class of RDF containers
- **rdfs:ContainerMembershipProperty** - the class of container membership properties, `rdf:_1`, `rdf:_2`, ..., all of which are sub-properties of `rdfs:member`

Properties

- **rdf:type** - an instance of `rdf:Property` used to state that a resource is an instance of a class
- **rdf:first** - the first item in the subject RDF list
- **rdf:rest** - the rest of the subject RDF list after `rdf:first`
- **rdf:value** - idiomatic property used for structured values
- **rdf:subject** - the subject of the subject RDF statement
- **rdf:predicate** - the predicate of the subject RDF statement
- **rdf:object** - the object of the subject RDF statement
- **rdfs:subClassOf** - the subject is a subclass of a class
- **rdfs:subPropertyOf** - the subject is a subproperty of a property
- **rdfs:domain** - a domain of the subject property
- **rdfs:range** - a range of the subject property
- **rdfs:label** - a human-readable name for the subject
- **rdfs:comment** - a description of the subject resource
- **rdfs:member** - a member of the subject resource
- **rdfs:seeAlso** - further information about the subject resource
- **rdfs:isDefinedBy** - the definition of the subject resource (Manola & Miller, 2013).

زبان پرس و جوی مورد استفاده در آر.دی.اف. نیز اسپارکل^۱ است که مبتنی بر اس.کیو.ال. کار می‌کند و از سلا ۲۰۰۸ توسط کنسرسیوم وب پیشنهاد شده است.

۳-۲-۲. دَمَل+اوایل^۲

دَمَل سرنام زبان نشانه گذاری دارپا^۳ است که به طور رسمی فعالیت خود را از آگوست سال ۲۰۰۰ آغاز کرد. هدف اصلی دَمَل تلاش برای توسعه زبان و ابزاری برای تسهیل مفهوم وب معنایی ذکر شده است. مایکل پاگلز^۴ مدیر برنامه‌های دارپا^۵ در پروژه دَمَل است. پروژه دَمَل در سال ۲۰۰۶ خاتمه یافت و این پروژه با او.دبلیو.ال.^۶ ادامه یافت (DAML, 2006).

وب شامل محتوایی است که توسط اچ.تی.ام.ال. نمایش داده می‌شوند. این محتواها در قالب اچ.تی.ام.ال. برای انسان جهت استفاده توسط مرورگرها در دسترس هستند. در صورتی که اچ.تی.ام.ال. به افراد اجازه بصری سازی اطلاعات را در محیط وب می‌دهد، و نمی‌تواند قابلیت‌های بالایی را در توصیف اطلاعات به روشی فراهم کند که با استفاده از برنامه‌های کامپیوتری امکان پیدا کردن و تفسیر آنها باشد. برای این منظور کنسرسیوم وب زبان توصیفی ایکس.ام.ال. را معرفی کرد. ایکس.ام.ال. همانگونه که پیشتر نیز ذکر شد با استفاده از برچسب‌هایی توصیف اطلاعات را توسط ماشین و به صورت ماشین فهم معجاز می‌کند. به عنوان مثال کلمه الگول^۷ در وبگاه شاید نشانگر زبان برنامه‌نویسی، نام یک ستاره یا کشتی اقیانوس پیمای تحقیقاتی باشد. با استفاده از ایکس.ام.ال. فراداده، مانند الگول، معنی اثر را آشکار می‌سازد. به هر حال، ایکس.ام.ال. قابلیت محدودی برای توصیف روابط (شماها یا هستی

¹ SPARQL

² DAML+OIL

³ DARPA Agent Markup Language (DAML)

⁴ Michael Pagels

⁵ DARPA (Defense Advanced Research Projects Agency)

⁶ OWL (Web Ontology Language)

⁷ Algol

نگاری‌ها) با توجه به اشیاء اطلاعاتی دارد. استفاده از هستی‌نگاری‌ها روش توانمندی را برای توصیف اشیاء و روابط آنها با سایر اشیاء فراهم می‌کند. زبان دَمَل به عنوان توسعه‌ای برای ایکس.ام.ال. و آر.دی.اف. ایجاد شده است. آخرین نسخه آن با عنوان دَمَل+اوایل^۱ مجموعه غنی از ساختارهای مورد نیاز برای ایجاد هستی‌نگاری‌ها را فراهم کرده است و اطلاعات را ماشین فهم می‌کند (DAML, 2006).

دَمَل+اوایل رویکردی شیئی گرا^۲ دارد و ساختار یک شیئی اطلاعاتی را با استفاده کلاس و ویژگی توصیف می‌کند. یک هستی‌نگاری مرکب از بدیهیاتی^۳ مانند روابط تابع میان کلاس‌ها و ویژگی‌ها است که اظهار می‌شوند. نحوه بیان و توصیف منابع در این زبان نیز مانند آر.دی.اف. است. به عنوان مثال زمانی که یک منبع (۲) نمونه‌ای از کلاس (c) است می‌گوییم ۲ از جنس c است. جدول ۲-۲ اجزای تشکیل دهنده دَمَل+اوایل را نشان می‌دهد. در نحو آر.دی.اف.، به عنوان مثال $\geq 2 \text{hasChild.Lawyer}$ به شکل زیر نوشته می‌شود:

```
<daml:Restriction daml:minCardinalityQ="2">
<daml:onProperty rdf:resource="#hasChild"/>
<daml:hasClassQ rdf:resource="#Lawyer"/>
</daml:Restriction>
```

جدول ۲-۲. اجزاء تشکیل دهنده دَمَل+اوایل

Constructor	DL Syntax	Example
intersectionOf	$C_1 \cap \dots \cap C_n$	Human $\cap$ Male
unionOf	$C_1 \cup \dots \cup C_n$	Doctor $\cup$ Lawyer
complementOf	$\neg C$	$\neg$ Male
oneOf	$\{x_1 \dots x_n\}$	{John,mary}
toClass	$\forall P.C$	$\forall \text{hasChildDoctor}$
hasClass	$\exists P.C$	$\exists \text{hasChildLawyer}$
hasValue	$\exists P.\{x\}$	$\exists \text{citizenOf}.\{\text{USA}\}$
minCardinalityQ	$\geq nP.C$	$\geq 2 \text{hasChildLawyer}$
maxCardinalityQ	$\leq nP.C$	$\leq 1 \text{hasChild.Male}$
cardinalityQ	$= nP.C$	$= 1 \text{hasParent.Female}$

معنی سه ساختار اول (intersectionOf, unionOf and complementOf) نسبتاً مشخص است: آنها عملگرهای بولی استاندارد هستند که کلاس را مجاز می‌کنند تا از تلفیق، متحد کردن و حذف طبقات دیگر تشکیل شود. ساختار OneOf به کلاس اجازه می‌دهد تا تعریف شود. ساختار HasClass و ToClass وابسته به حافظه‌ای هستند که به زبان مبتنی بر چارچوب ایجاد شده است. کلاس $\forall P.C$ کلاسی است که تمامی اجزاء آن مرتبط با ویژگی P است که تنها منبع از نوع C است،

¹ DAML+OIL (Ontology Inference Layer)

² Object Oriented

³ Axioms

در حالی که کلاس $P.C$ ^۳ کلاسی است که تمام اجزاء آن با استفاده از ویژگی P به حداقل یک منبع از نوع C مرتبط هستند. ساختار $HasValue$ مختصر شده ترکیب $HasClass$ و $OneOf$ است.

ساختار $minCardinalityQ$ و $maxCardinalityQ$ و $CardinalityQ$ تعمیم یافته $ToClass$ و $HasClass$ هستند. کلاس $\geq nP.C$ ($\leq nP.C$, $=nP.C$) کلاسی است که تمام نمونه‌های آن از طریق ویژگی P به حداقل به n منبع مختلف از نوع C مرتبط باشند.

جدول ۲-۳ به طور خلاصه قواعدی را که توسط دمل+اوایل پشتیبانی می‌شوند را نشان می‌دهد. این قواعد قابلیت اظهار توابع و هم ارزی را با توجه به کلاس و ویژگی‌ها می‌دهد. ویژگی مهم دمل+اوایل این است که قواعد $SubClassOf$ و $SameClassAs$ می‌توانند بجای کلاس نیز استفاده شوند. نتیجه این نوع بیان‌ها این است که تمام کلاس‌ها و قواعد می‌توانند تعداد زیرکلاس‌ها و کلاس‌های هم سطح را کاهش دهند.

جدول ۲-۳. قواعد دمل+اوایل (Horrocks, 2002)

Axiom	DL Syntax	Example
subClassOf	$C_1 \sqsubseteq C_2$	Human $\sqsubseteq$ Animal $\sqcap$ Biped
sameClassAs	$C_1 \equiv C_2$	Man $\equiv$ Human $\sqcap$ Male
subPropertyOf	$P_1 \sqsubseteq P_2$	hasDaughter $\sqsubseteq$ hasChild
samePropertyAs	$P_1 \equiv P_2$	cost $\equiv$ price
disjointWith	$C_1 \sqsubseteq \neg C_2$	Male $\sqsubseteq \neg$ Female
sameIndividualAs	$\{x_1\} \equiv \{x_2\}$	{President_Bush} $\equiv$ {G_W_Bush}
differentIndividualFrom	$\{x_1\} \sqsubseteq \neg\{x_2\}$	{john} $\sqsubseteq \neg$ {peter}
inverseOf	$P_1 \equiv P_2^-$	hasChild $\equiv$ hasParent ⁻
transitiveProperty	$P^+ \sqsubseteq P$	ancestor ⁺ $\sqsubseteq$ ancestor
uniqueProperty	$\top \sqsubseteq \leq 1P$	$\top \sqsubseteq \leq 1$ hasMother
unambiguousProperty	$\top \sqsubseteq \leq 1P^-$	$\top \sqsubseteq \leq 1$ isMotherOf ⁻

همانگونه که مشاهده می‌شود دمل+اوایل به ویژگی‌ها اجازه می‌دهد تا اظهار شوند (Horrocks, 2002).

۴-۷-۲. او.دبلیو.ال.

او.دبلیو.ال. سرنام زبان هستی نگاری وب^۱ است. این زبان برای استفاده توسط نرم‌افزارهایی طراحی شده که نیاز به پردازش محتوای اطلاعات بجای نمایش اطلاعات به انسان را دارند. او.دبلیو.ال. تعامل ماشینی محتوای وب را تسهیل می‌کند که توسط ایکس.ام.ال.، آر.دی.اف. و آر.دی.اف.اس. با افزودن واژگان جدید به همراه معنای رسمی پشتیبانی می‌شوند. او.دبلیو.ال. دارای سه زبان فرعی به شرح زیر است: او.دبلیو.ال. لایت^۲، او.دبلیو.ال. دی.ال.۲، و او.دبلیو.ال. فول^۴. (McGuinness & Harmelen, 2009). این زبان توسط کنسرسیوم وب در اکتبر ۲۰۰۷ معرفی شد. متعاقب ویرایش اول، ویرایش جدید این زبان با عنوان او.دبلیو.ال. ۲ در سال ۲۰۰۹ معرفی شد. در حال حاضر نرم‌افزارهای ویرایش معنایی و ساخت هستی نگاری مانند پروتزه^۵ و نرم‌افزارهای تحلیل منطقی^۶ مانند پلت^۷، ریسرپرو^۸، فکت++^۱ و هریمیت^۲ از این زبان استفاده می‌کنند.

^۱ Web Ontology Language

^۲ OWL Lite

^۳ OWL DL

^۴ OWL Full

^۵ Protégé

نرم‌افزار منبع باز ویرایشگر هستی نگاری و سیستم اکتساب دانش است. این نرم‌افزار با جاوا نوشته شده و از سوینگ (Swing) برای ایجاد رابط کاربری بسیار پیچیده آن استفاده شده است.

^۶ Semantic reasoners

^۷ Pellet

^۸ RacerPro

همانگونه که در بالا نیز اشاره شد او.دبلیو.ال. دارای سه زبان فرعی به شرح زیر است:

- او.دبلیو.ال. لایت. از کاربرانی پشتیبانی می‌کند که به طور مقدماتی نیاز به طبقه‌بندی سلسله مراتبی و محدودیت‌های ساده هستند. او.دبلیو.ال. لایت مسیر انتقالی سریع به اصطلاحنامه‌ها و تاکسونومی‌ها^۳ است.
- او.دبلیو.دی.ال. از کاربرانی پشتیبانی می‌کند که خواهان قابلیت نمایشی حداکثری هستند در حالی که محاسبات کامل و دقیق و سریع نیز مراعات شود. او.دبلیو.ال. دی.ال. شامل تمام ساختارهای زبان او.دبلیو.ال. است، در حالی که می‌توانند تحت محدودیت‌های خاصی کار کنند (به عنوان مثال زمانی که یک کلاس ممکن است زیر کلاس بسیاری از کلاس‌های دیگر باشد، یا اینکه یک زیر کلاس نتواند نمونه‌ای از کلاس دیگر باشد).
- او.دبلیو.ال. فول مورد استفاده کاربرانی قرار می‌گیرد که حداکثر انتظارات و آزادی نحوی از آر.دی.اف. را بدون تضمین محاسباتی دارند. به عنوان مثال، در او.دبلیو.ال. فول یک کلاس می‌تواند به طور هم زمان به عنوان مجموعه‌ای از واحدها و به عنوان یکی در محل خود بکار رود. او.دبلیو.ال. فول به آنتولوژی اجازه تقویت معنی پیش تعریف واژگان آر.دی.اف. یا او.دبلیو.ال. را می‌دهد.

بر این اساس توسعه دهندگان هستی نگاری‌ها با استفاده از او.دبلیو.ال. باید توجه داشته باشند که کدام زبان فرعی برای برآوردن نیازشان مناسب است. انتخاب میان او.دبلیو.ال. لایت و او.دبلیو.ال. دی.ال. بستگی به میزان نیاز کاربران به سازه رسا و نمایشی که توسط او.دبلیو.ال. دی.ال. فراهم شده دارد. انتخاب میان او.دبلیو.ال. دی.ال. و او.دبلیو.ال. فول بستگی به میزان نیاز کاربران به تسهیلات فرامدل سازی شمای آر.دی.اف. دارد (McGuinness & Harmelen, 2009).

همانگونه که پیشتر نیز ذکر است این زبان برای تولید هستی نگاری‌ها مورد استفاده قرار می‌گیرد. هستی نگاری‌ها نیز برای برقراری روابط معنایی در سیستم‌های اطلاعاتی می‌توانند کاربرد داشته باشند. بنابراین در برقراری تعامل معنایی در سیستم‌های اطلاعاتی متجانس و حتی نامتجانس می‌تواند برای برقراری روابط معنایی مورد استفاده قرار گیرد.

۸-۲. نمونه‌های یکپارچه‌سازی معنایی

در زیر برخی از پروژه‌های انجام شده در زمینه یکپارچه‌سازی معنایی که در منابع مورد استفاده جهت نگارش این بخش مورد اشاره قرار گرفته‌اند، معرفی می‌شوند.

۸-۲-۱. پروژه دلایس: شبکه کتابخانه دیجیتالی است که با بودجه کمیسیون اروپا^۴ در قالب برنامه فناوری جامعه اطلاعاتی^۵ در سال ۲۰۰۵ در قالب ۷ پروژه اجرا شده است. ۵۹ موسسه بین‌المللی و ملی در پروژه دلایس همکاری دارند که برخی از آنها به شرح زیر هستند: دانشگاه کوینزلند استرالیا^۶، فناوری اطلاعات سلامت تیرول اتریش^۷، موسسه سیستم‌های اطلاعاتی و رسانه کامپیوتری اتریش^۸، دانشکده علوم اتریش^۹، کتابخانه ملی اتریش^{۱۰}، دانشگاه فناوری وین^{۱۱}، دانشگاه علوم بهداشت فناوری و

^۱ FaCT++

^۲ HermiT

^۳ Taxonomies

^۴ European Commission

^۵ Information Society Technologies Programme (IST)

^۶ University of Queensland

^۷ Health Information Technologies Tyrol

^۸ Institute for Information Systems and Computer Media

^۹ Austrian Academy of Science

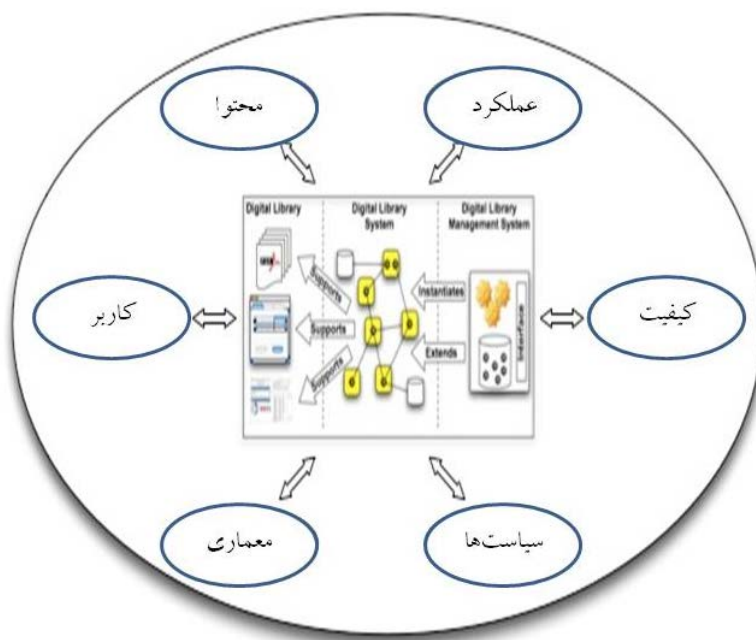
^{۱۰} Austrian National Library

^{۱۱} Vienna University of Technology

انفورماتیک پزشکی اتریش^۱، دانشگاه ماساریک برنو جمهوری چک^۲، مرکز مهندسی سیستم‌های شناختی دانمارک^۳، موسسه اشتراک دانش دانمارک^۴، آزمایشگاه ملی ریسو دانمارک^۵ ([Partners, 2008]). دو هدف عمده پروژه شبکه دلاس عبارتند از:

- تعریف نظریه‌ها و چارچوب‌های واحد و جامع چرخه حیات اطلاعات کتابخانه دیجیتالی؛
- ایجاد خدمات مبادله چندزبانی و مدیریت محتوای یکپارچه و انتشار عمومی آن. این شبکه در نظر دارد فناوری کتابخانه دیجیتال منحصر بفردی را توسعه دهد. این هدف ذکر شده در شبکه دلاس به اهداف پروژه حاضر مرتبط است.

یکی از زیر پروژه‌های اجرا شده در این پروژه میانکنش پذیری معنایی و استخراج دانش^۶ در سال‌های ۲۰۰۴ و ۲۰۰۵ بود که مرتبط‌ترین پروژه نزدیک به موضوع این پژوهش است. راهکار ارائه شده در این پروژه استفاده از استانداردها، اصطلاحنامه‌ها، هستی‌نگاری‌ها، سیستم‌های سازماندهی دانش و اجرای شمای فراداده بود. یکی از محصولات این پروژه مدل مرجع^۷ کتابخانه دیجیتال است که در شکل ۲-۱۲ نشان داده شده است.



شکل ۲-۱۲. مدل مرجع کتابخانه دیجیتال دلاس

بررسی شکل ۲-۱۲ نشان می‌دهد که محتوای کتابخانه دیجیتالی شامل انواع منابع اطلاعاتی است که می‌تواند توزیع یافته در نظر گرفته شود. با توجه به اینکه شکل مذکور مدل مرجع کتابخانه دیجیتال است در آن شرح جزئیات ملاحظه نمی‌شود.

¹ University for Health Sciences Medical Informatics and Technology

² Masaryk University of Brno

³ Center of Cognitive Systems Engineering

⁴ Institute of Knowledge Sharing

⁵ Risø National Laboratory

⁶ Knowledge extraction and semantic interoperability (WP5)

⁷ Reference Model

گزارش این پروژه با عنوان تعامل معنایی در سیستم‌های کتابخانه دیجیتال در سال ۲۰۰۵ منتشر شده است. این گزارش در پنج فصل تنظیم شده و به مسائلی نظیر اهمیت تعامل معنایی در سیستم‌های کتابخانه دیجیتال، انواع و سطوح مختلف تعامل معنایی، ارتباط بین تعامل معنایی و نحوی، توصیف روش‌ها، پیش‌نیازها و استانداردها و خدمات معنایی، چگونگی بهبود تعامل معنایی در سیستم‌های کتابخانه دیجیتال و مسائل مرتبط با پروژه دلاس و جامعه کتابخانه دیجیتال بصورت کلی می‌پردازد.

۲-۸-۲. **پروژه تی.اس.آی.ام.ام.آی.اس.^۱**: این پروژه با همکاری دانشگاه استنفورد و شرکت آی.بی.ام. با حمایت مالی دارپا^۲ از سال ۱۹۹۴-۱۹۹۶ به اجرا درآمد (Rys, 1998). هدف این پروژه توسعه نسل بعدی سیستم‌های مدیریت اطلاعاتی است که دستیابی یکپارچه‌ای را به سطح گسترده‌ای از منابع اطلاعاتی و دانشی فراهم کند (Chawathe, et al., [1996]). این سیستم دستیابی به متون ساختار نیافته مانند متون نوشتاری و فایل‌های متنی، پایگاه‌های داده ساخت یافته و داده‌های خام مانند تصاویر ماهواره‌ای را ایجاد و مدیریت می‌کند. از طریق این سیستم، تصمیم ساز قادر خواهد بود تا اطلاعات مورد نیاز خود را بیابد، با اطلاعات منابع مختلف درآمیزد، و آن را پردازش کند. در مجموع این سیستم برای پشتیبانی از استخراج اطلاعات، ابزاری را برای نظارت و اعمال محدودیت در سراسر مخازن اطلاعاتی ناهمگن ایجاد می‌کند (Garcia-Molina, 1994). رویکرد این پروژه ایجاد سه مولفه زیر بود: مترجم‌ها^۳، واسطه‌ها^۴، و مدیران محدودیت^۵. مترجم‌ها پوشش‌هایی^۶ هستند که دستورات و اطلاعات را به مدل شیئی مشترک^۷ و زبان دستوری تبدیل می‌کنند. واسطه‌ها قادر به تلفیق اطلاعات از منابع مختلف هستند. مدیران محدودیت، محدودیت‌هایی را اعمال یا نقیض آنها را در حالت توزیع یافته شناسایی می‌کنند. در این پروژه مدل او.ای.ام.^۸ به انضمام زبان پرس و جوی اس.کیو.ال.^۹ ایجاد شد. معماری مورد استفاده در این پروژه در شکل ۲-۱۳ نشان داده شده است. همانگونه که در معماری زیر نیز مشاهده می‌شود سه مولفه مذکور محور اصلی معماری پروژه تی.اس.آی.ام.ام.آی.اس. را تشکیل می‌دهد. کاربر نهایی نیز از طریق نرم‌افزاری که قادر به ارائه پرس و جو از طریق زبان او.ای.ام. یا هر مرورگر تحت وبی باشد می‌تواند با این سیستم تعامل داشته و اطلاعات مورد نیاز خود را در سیستم‌های اطلاعاتی توزیع یافته بازیابی کند. طبقه‌بندی کننده/استخراج کننده^{۱۰} در پایین شکل ۲-۱۳ وظیفه استخراج و دسته بندی داده‌ها از منابع ساختار نیافته را دارد. به عنوان مثال داده‌ها را به فایل‌های برگرفته از پست الکترونیکی، فایل‌های متنی، فایل‌های تصویری و غیره دسته بندی می‌کند و اطلاعات کلیدی مانند تاریخ تولید، نویسنده و غیره را استخراج می‌کند (Chawathe, et al., [1996]).

¹TSIMMIS (The Stanford-IBM Manager of Multiple Information Sources)

² DARPA

³ Translators

⁴ Mediators

⁵ Constraint managers

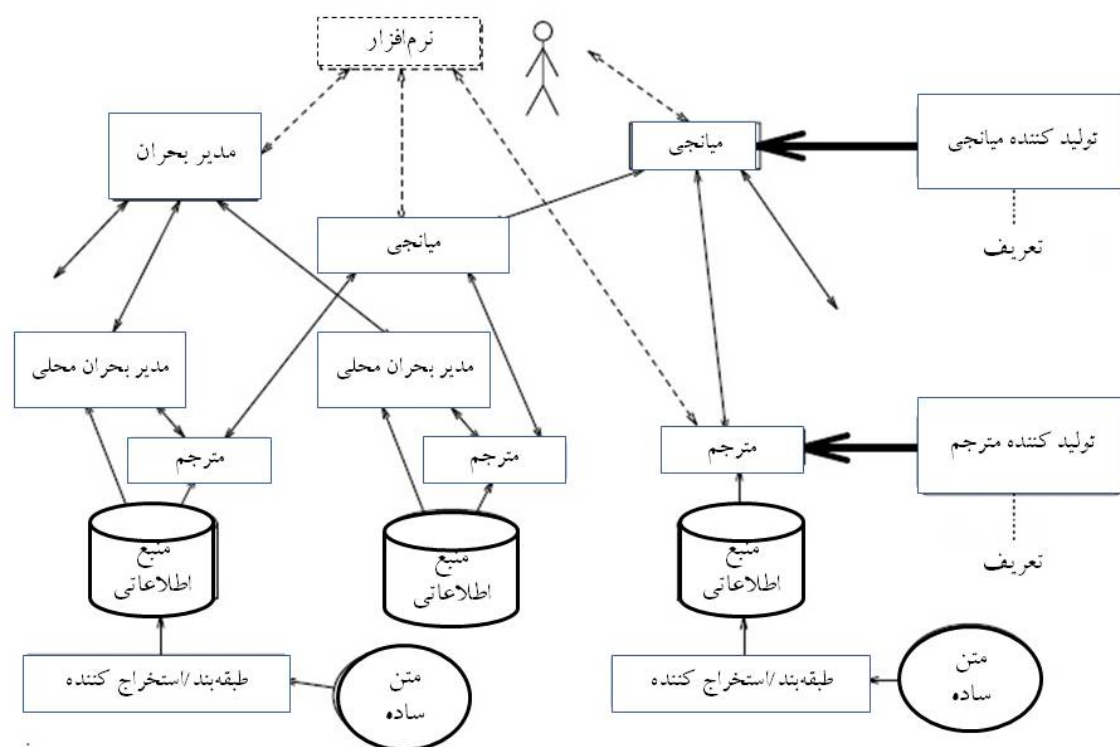
⁶ Wrappers

⁷ Common Object Model

⁸ OEM (Object Exchange Model)

⁹ SQL-like query language

¹⁰ Classifier/Extractor



شکل ۲-۱۳ معماری پروژه تی.اس.آی.ام.آی.اس. (Chawathe, et al., [1996])

به طور کلی این پروژه به دنبال ایجاد سیستمی برای یکپارچه‌سازی اطلاعات است. در این پروژه مدل داده‌ای و یک زبان درخواست مشترک جهت کمک به ترکیب اطلاعات از چند منبع مختلف طراحی و ارائه شده است. در این سیستم ابزارهایی نیز برای تولید خودکار اجزای مورد نیاز برای ایجاد سیستم‌های یکپارچه‌سازی اطلاعات ارائه شده است. تاکید این سیستم بر روی ایجاد خودکار ترجمه و واسطه‌هایی برای دسترسی و ترکیب اطلاعات در منابع داده‌ای نامتجانس است.

۳-۸-۲. پروژه یورویانا^۱: یورویانا پرتال اینترنتی برای دستیابی به میلیون‌ها کتاب، نقاشی، فیلم، اشیاء موزه‌ای و منابع آرشیوی است که در اروپا دیجیتالی شده‌اند. بیش از ۲۰۰۰ موسسه در اروپا با این پروژه همکاری دارند. علت اصلی شروع این پروژه مربوط به نامه‌ای از ژاک شیراک^۲، رئیس جمهور وقت فرانسه، در سال ۲۰۰۵ بود. نامه مذکور پیشنهاد ایجاد کتابخانه مجازی اروپا را داده بود. نسخه اولیه کتابخانه دیجیتالی در سال ۲۰۰۹ راه اندازی شد. یورویانا امکان دستیابی به محتواهای با فرمت‌های مختلفی را همانگونه که پیشتر نیز ذکر شد فراهم کرده است. این منابع در سیستم‌های اطلاعاتی مختلف در مراکز مختلف نگهداری می‌شوند ولی امکان دستیابی به آنها از طریق درگاه یورویانا فراهم است. یورویانا فراداده‌های مربوط به منابع اطلاعاتی را گردآوری می‌کند. جستجوی کاربران در این فراداده‌ها انجام می‌گیرد و برای دستیابی به منابع اطلاعاتی به منبع اصلی متصل می‌شوند. برای این منظور فراداده‌های مختلف به یک فراداده استاندارد با عنوان عناصر معنایی یوپانا^۳ ترجمه می‌شوند. قالب مورد استفاده در یکپارچه‌سازی معنایی آر.دی.اف. بر مبنای استاندارد فراداده دابلین‌گر در این پروژه استفاده می‌شود (Introduction, 2011). به

¹Europeana

² Jacques Chirac

³ Europeana Semantic Elements

طور کلی این پروژه اشیاء منابع اطلاعاتی نامتجانس (مانند کتابخانه‌ها، موزه‌ها و آرشیوها) را به صورت یکپارچه ارائه می‌نماید. تکیه بیشتر این پروژه بر روی استانداردهای وب و اینترنت و عناصر ساختاری است که از قبل موجود است.

۴-۸-۲. **ان.دی.ال.تی.دی.**^۱: این پروژه در سال ۱۹۸۷ توسط موسسه بین المللی میکروفیلم‌های دانشگاهی^۲ اجرا شد. دانشگاه ویرجینیا تک بین سال‌های ۱۹۸۰ تا ۱۹۹۰ اقدام به پژوهش در خصوص تبدیل پایان‌نامه‌های قابل ارائه در دیکست به یک فرمت الکترونیکی استاندارد کرد. سپس در سال ۱۹۹۲، دانشگاه ویرجینیا تک با همکاری نمایندگان از یو.ام.آی. و انجمن موسسات آموزش عالی آمریکا کارگاهی را برای شروع یک همکاری برای پروژه‌ای با محتوای اطلاعات شبکه‌ای برای توسعه مجموعه‌های پایان‌نامه‌ها برگزار نمود. در سال ۱۹۹۴ دانشگاه ویرجینیا تک سامانه مدیریت پایان‌نامه‌ها را به کار گرفت که دانشجویان موظف به ارائه پایان‌نامه‌های خود به صورت الکترونیکی از طریق آن بودند. برای توسعه رویکرد مشارکت در این حوزه، در سال ۱۹۹۵ کتابخانه دیجیتال شبکه‌ای پایان‌نامه‌ها با عنوان ان.دی.ال.تی.دی. به منظور توسعه برنامه‌ها، ابزارها، سامانه‌ها، استانداردها و راهنمای نرم‌افزاری الگو و بحث بر روی مشکلات این حوزه ایجاد گردید. این کتابخانه از اکتبر ۱۹۹۶ اقدام به جمع‌آوری نسخه‌های پایان‌نامه‌ها نموده است. بیش از ۱۴۰ دانشگاه از سراسر جهان در این موسسه عضو هستند (به روز شده از سایت ان.دی.ال.تی.دی.). این موسسه اقدام به طراحی و به کارگیری ای.تی.دی.-ام.اس.^۳، که یک استاندارد فراداده اختصاصی برای پایان‌نامه‌ها می‌باشد، نموده است. ای.تی.دی.-ام.اس. که خود یکی از بخش‌های یک مفهوم بزرگتر با نام او.ای.آی.-پی.ام.اچ.^۴ است. این مفهوم یک مکانیزم استاندارد برای جمع‌آوری فراداده‌ها در یک سایت مرکزی و نمایش فراداده‌های رکورد‌های پایان‌نامه‌ها برای دیگر موسسات می‌باشد. این مکانیزم دسترسی به مجموعه‌های دانشگاه‌ها را فراهم می‌نماید.

به طور کلی کتابخانه دیجیتال شبکه‌ای پایان‌نامه‌ها تلاشی مشترک توسط دانشگاه‌هایی از سرتاسر جهان است تا پایان‌نامه‌های الکترونیک را در دسترس محققان قرار دهد. زمانی که تعداد دانشگاه‌ها کم بود از رویکرد جستجوی یکپارچه^۵ استفاده می‌شد. ولی با گسترش عضویت‌ها، فهرستگانی تهیه شد که جستجو و بازیابی پایان‌نامه‌های الکترونیک را از سایت‌های مختلف متعلق به اعضای این پروژه، مقدور می‌سازد. در این پروژه از پروتکل او.ای.آی.-پی.ام.اچ. جهت یکپارچه‌سازی فراداده‌های مربوط به پایان‌نامه‌ها در هریک از دانشگاه‌های عضو استفاده می‌شود.

۵-۸-۲. **پروژه ان.اس.دی.ال.**^۶: کتابخانه دیجیتال ملی آموزش علم، فناوری، مهندسی و ریاضیات تلاشی است که توسط بنیاد ملی علوم ایالات متحده در جهت سازماندهی و دسترس‌پذیر نمودن منابع الکترونیک برای آموزش حوزه‌های علمی مذکور شکل گرفت. این پروژه تعامل را در سه سطح یکپارچه، گردآوری، و جمع‌آوری به انجام می‌رساند. در سطح سازمانی تعامل میان مجموعه‌ها مطابق با استانداردهایی نظیر Z39.50 برقرار می‌شود. ولی به‌خاطر محدودیت‌های جستجوی یکپارچه در سطح سازمانی، ان.اس.دی.ال. از امکان گردآوری اطلاعات از مخازن آرشیو باز^۷ و ایجاد پایگاه اطلاعاتی مرکزی استفاده می‌کند (Fulker &

¹ NDLTD (Networked Digital Library of Theses and Dissertations)

² University Microfilms International

³ Electronic Theses and Dissertations Metadata Standard

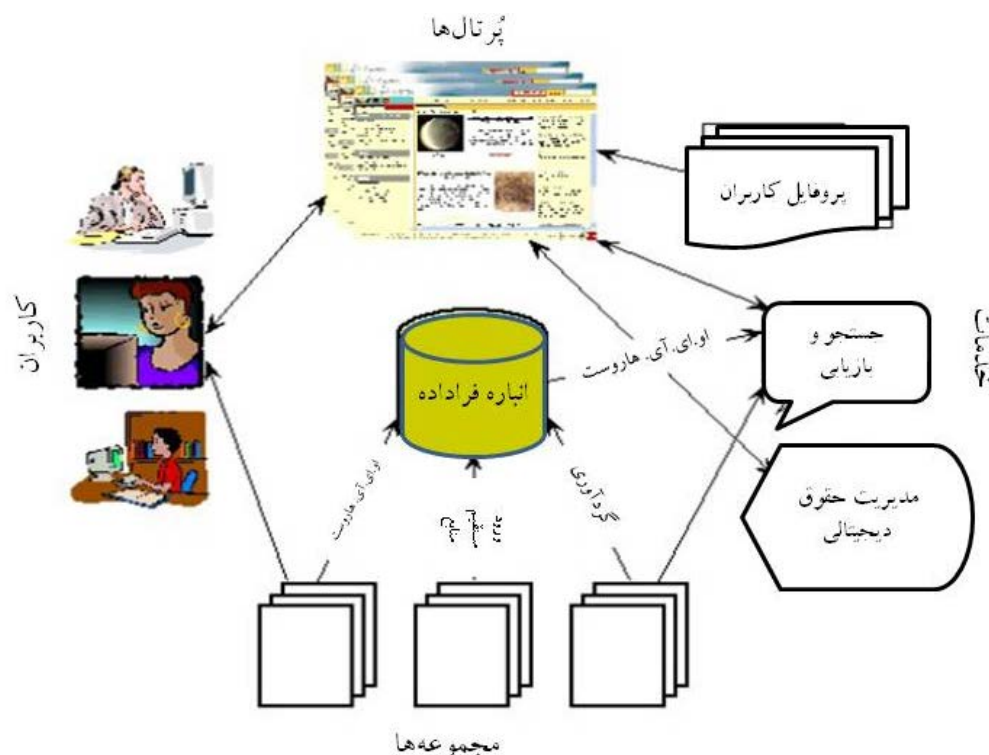
⁴ Open Archive Initiative/ Protocol of Metadata Harvesting

⁵ Federated search

⁶ NSDL

⁷ OAI repositories

(Janée, 2002) علاوه بر این دو سطح، خزنده‌ها^۱ و افراد به کار گرفته می‌شوند تا منابع جمع‌آوری شوند شبیه همان روشی که موتورهای جستجوی عمومی وب انجام می‌دهند. معماری مورد استفاده در این پروژه به شکل زیر است:



۲-۱۴. عناصر معماری ان.اس.دی.ال. (Lagoze, et al., 2002)

مرور معماری ذکر شده نشانگر استفاده از تفاهم‌نامه باز او.ای.آی. و استفاده از پایگاه اطلاعاتی واسط جهت ارائه خدمات یکپارچه به کاربران است.

۲-۹. نتیجه‌گیری

اینترنت و محیط وب بستری را فراهم ساخته‌اند تا افراد بتوانند از راه دور به اطلاعات مورد نیاز خود دست‌یابند. عمومیت یافتن وب باعث شد تا رفتار اطلاع‌جویی افراد نیز تغییر یابد و امروزه تقریباً برای دست‌یابی به هر نوع اطلاعاتی، افراد حداقل جستجویی نیز در محیط وب انجام می‌دهند. از این رو، کتابخانه‌ها نیز در مسیر رشد و بالندگی خود، حضور مستمر و پررنگ خود را در حضور در محیط وب و در دسترس بودن از طریق این مجرا دانسته و نسل‌های جدیدی از کتابخانه‌ها نمایان شده است. ابتدا اطلاعات کتابشناختی منابع اطلاعاتی کتابخانه‌ها از طریق وب در دسترس قرار گرفتند و به تدریج محتوای منابع اطلاعاتی نیز به این مجموعه‌ها افزوده شدند و دست‌یابی به منابع اطلاعاتی برای کاربران دوردست در کتابخانه‌ها فراهم شد. کارکرد مذکور مفهوم جدیدی با عنوان کتابخانه‌های دیجیتال را رواج داد که می‌توانند نقش پررنگ کتابخانه‌ها را در محیط وب ایفا نمایند.

¹ Crawlers

با وجود این همچنان دستیابی به اطلاعات در کتابخانه‌ها منوط به عضویت در تک تک کتابخانه‌ها و کسب شرایط عضویت بود. در اکثر موارد این مسئله مانع بزرگی در دستیابی به اطلاعات در کتابخانه‌های دیجیتال قلمداد می‌شود. این مسئله در سایر منابع اطلاعاتی خارج از کتابخانه‌ها، مانند پایگاه‌های اطلاعاتی تجاری ناشران، نیز به وضوح قابل مشاهده است. به عنوان مثال اگر نگاهی به وضعیت پایگاه‌های اطلاعاتی در چند ماه گذشته کنیم، مشاهده خواهیم کرد که فرد برای جستجو در هر یک از پایگاه‌های اطلاعاتی، علاوه بر پرداخت هزینه اشتراک تک تک پایگاه‌های اطلاعاتی، باید به صورت منفرد در تک تک پایگاه‌های اطلاعاتی اقدام به جستجو می‌نمود. فرآیندی که پایگاه‌های اطلاعاتی تجاری پیش گرفتند می‌توانست برای کتابخانه‌ها که منابع کاملاً غیرانتفاعی هستند راهگشا باشد. چرا که پایگاه‌های اطلاعاتی تجاری از روش‌های مختلفی به دنبال کسب درآمد اقتصادی خود هستند. اولین راهکاری که توسط پایگاه‌های اطلاعاتی تجاری شکل گرفت، استفاده از تلفیق پایگاه‌های اطلاعاتی با یکدیگر بود. در این روش چندین پایگاه اطلاعاتی کوچک یا بزرگ به جهت غنی‌تر ساختن مجموعه‌های خود با یکدیگر تلفیق شدند.

بنابراین کتابخانه‌های دیجیتال نیز با تأسی از راهکار مذکور باید سعی در ایجاد دستیابی یکپارچه به محتواهای خود داشته باشند. از این رو، فعالیت‌هایی از سال ۱۹۹۸ در این زمینه برای مبادله اطلاعات کتابشناختی و سپس یکپارچه‌سازی محتواها و خدمات کتابخانه‌های دیجیتال شکل گرفت (Alipour-Hafezi, Horri, Shiri and Ghaebi, 2010). در تمامی پروژه‌های مذکور یکپارچه‌سازی محتوا و خدمات بر مبنای موضوع فعالیت کتابخانه‌های عضو مورد نظر قرار گرفته‌اند. به این معنی که کتابخانه‌های دیجیتال هم موضوع اقدام به تشکیل شبکه‌های کتابخانه‌ای یکپارچه می‌کنند.

۱۰-۲. پیشینه پژوهش

هر پژوهشی بر پژوهش‌های پیش از خود متکی است و همواره در پژوهش‌ها سعی بر این است تا دانش جدیدی بر پایه دانش موجود تولید و افزوده شود. از این رو است که شعار ایستاده بر شانه غول‌ها^۱ در زیر جعبه جستجوی گوگل اسکالر^۲ ثبت شده است. شعار فوق به این مفهوم است که دانش و دانشمندان جدید بر شانه‌های دانش و دانشمندان پیش از خود ایستاده‌اند و بدون توسل به پیشینیان نمی‌توان دانش جدیدی را تولید کرده و افق‌های نوین را گشود. همین مسئله نشان از اهمیت بررسی دانش پیشین در پژوهش‌ها دارد. بنا بر این در این بخش به شناسایی و معرفی پژوهش‌های پیشین پرداخته شده است.

جهت شناسایی دانش پیشین موجود در زمینه یکپارچه‌سازی معنایی در کتابخانه دیجیتال ابتدا اقدام به شناسایی کلیدواژه‌های جستجو در پایگاه‌های اطلاعاتی مرتبط فارسی و لاتین گردید. در این زمینه کلیدواژه‌های فارسی و لاتین زیر با بهره‌گیری از اصطلاحنامه‌ها و مقالات دایرةالمعارفی و پایه مرتبط شناسایی گردیدند:

- یکپارچه‌سازی معنایی
- کتابخانه دیجیتال
- کتابخانه رقمی
- کتابخانه رقمی

^۱ Stand on the shoulders of giants

^۲ Google scholar

- مبادله اطلاعات
- میان‌کنش پذیری
- وب معنایی
- هستی‌شناسی
- Semantic Integration
- Semantic web
- Digital library
- Interoperability
- Semantic relationship
- Semantic digital library
- Ontology
- Integrated access
- Distributed digital libraries
- Semantic heterogeneity
- Semantic grid
- Linked Data

در راهبرد جستجوی مورد استفاده سعی شد تا جستجوهای ترکیبی با استفاده از ترکیب کلیدواژه‌های شناسایی شده بکار برده شوند. از طرف دیگر در انتخاب پایگاه‌های اطلاعاتی نیز تلاش شد تا پایگاه‌های اطلاعاتی مرتبط شناسایی و مورد جستجو قرار گیرند. از این رو پس از بررسی‌های به عمل آمده پایگاه‌های اطلاعاتی زیر برای جستجو و بازیابی اطلاعات مرتبط به زبان فارسی انتخاب شدند:

- پایگاه پایان‌نامه‌های ایران‌داک^۱
- پایگاه اطلاعاتی مگ‌ایران^۲
- پایگاه اطلاعاتی نورمگز^۳
- پایگاه‌های اطلاعاتی نشریات علمی-پژوهشی مرتبط، مانند فصلنامه علوم و فناوری اطلاعات، اطلاع‌شناسی، کتابداری و اطلاع‌رسانی آستان قدس رضوی، فصلنامه کتاب، و نظایر اینها.
- پایگاه‌های اطلاعاتی مرکز منطقه‌ای علوم و فناوری^۴
- همچنین وبگاه‌های برخی از دانشگاه‌هایی که در حوزه موضوعی پژوهش فعالیت دارند، مانند دانشگاه تهران، دانشگاه فردوسی، و غیره، نیز مورد جستجو قرار گرفتند.

^۱ <http://database.irandoc.ac.ir>

^۲ <http://www.magiran.com/>

^۳ <http://www.noormags.com/view/fa/default>

^۴ <http://www.ricest.ac.ir/Default.aspx?lan=fa>

برای جستجو به زبان انگلیسی نیز پایگاه‌های اطلاعاتی زیر مورد جستجو قرار گرفتند:

- پایگاه پایان‌نامه‌های پروکوئست^۱
- پایگاه اطلاعاتی امرالد^۲
- پایگاه اطلاعاتی ساینس دایرکت^۳
- پایگاه اطلاعاتی آی.ای.ای.ای.^۴
- پایگاه اطلاعاتی گوگل اسکالر^۵
- همچنین جستجوهای نیز به صورت عمومی در موتور جستجوی گوگل انجام گرفت.

جستجوهای انجام شده در پایگاه‌های اطلاعاتی مذکور منتج به یافته‌هایی شد که در ادامه به معرفی و تحلیل آنها پرداخته شده است.

۱-۱۰-۲. پیشینه پژوهش در داخل

جستجو در پایگاه‌های اطلاعاتی داخل با شناسایی پژوهشی که مستقیماً به موضوع پژوهش مرتبط باشد مواجه نشد. لذا اقدام به عام‌تر ساختن کلیدواژه‌های جستجو گردید در چنین شرایطی با ۴ پژوهش و چهار طرح اجرایی آشنا شدیم که دو مورد از پژوهش‌ها توسط پژوهشگر (مهدی علیپور حافظی) انجام گرفته است. در ادامه به معرفی موارد مذکور می‌پردازیم:

محقق (۱۳۷۴) در پایان‌نامه کارشناسی ارشد خود با عنوان "بررسی وضعیت مبادله اطلاعات و منابع میان کتابخانه‌های دانشگاهی شهر تهران" به بررسی وضعیت مبادله اطلاعات در کتابخانه‌های دانشگاهی پرداختند. هدف از این پژوهش شناسایی وضعیت مبادله اطلاعات و منابع اطلاعاتی در کتابخانه‌های دانشگاهی شهر تهران بود. پژوهش حاضر با روش پیمایشی و با استفاده از ابزار پرسشنامه انجام گرفت. براساس نتایج حاصل از این پژوهش ۴۹/۵ درصد از این کتابخانه‌ها با یکدیگر مبادله منابع اطلاعاتی دارند و کتابداران از این موضوع به عنوان ضروری‌ترین نیاز کتابخانه‌ها یاد می‌کنند. استفاده از فناوری اطلاعات در کتابخانه‌های دانشگاهی در اقلیت ناچیزی قرار گرفته که بیشتر تمایل به استفاده از دیسک‌های فشرده نوری و تکنولوژی ارتباط با بانک‌های جهانی وجود دارد. تعداد کمی از کتابخانه‌ها با مراکز علمی خارج از کشور ارتباط دارند که در اکثر این موارد از کتابخانه بریتانیا یاد شده است. استفاده از روش‌های قدیمی مبادله اطلاعات و منابع مانند پست و تهیه کپی از منابع در کتابخانه‌ها رواج داشته و بیشترین منابع مورد مبادله مقاله‌های مجلات، کتاب و سپس راهنمایی از طریق تلفن می‌باشد (محقق، ۱۳۷۴). در پژوهش حاضر به بحث در زمینه مبادله اطلاعات در کتابخانه‌های سنتی پرداخته شده است و تأکید اصلی بر ضرورت مبادله اطلاعات بوده و وضعیت موجود نشانگر استفاده از روال‌های سنتی در این راستا است.

علیپور حافظی (۱۳۸۸) در پایان‌نامه دکتری خود با عنوان "بررسی نحوه مبادله اطلاعات بین سیستم‌های اطلاعاتی کتابخانه‌های دیجیتال در ایران و ارائه الگوی پیشنهادی" به بررسی مبادله اطلاعات در سیستم‌های اطلاعاتی کتابخانه دیجیتال از منظر تعامل

¹ <http://proquest.umi.com/pqdweb?RQT=302&cfc=1>

² <http://www.emeraldinsight.com/>

³ <http://www.emeraldinsight.com/>

⁴ <http://ieeexplore.ieee.org/Xplore/guesthome.jsp>

⁵ <http://scholar.google.com/schhp?hl=en&tab=ws>

نحوی پرداخت. هدف از این پژوهش تعیین نحوه مبادله اطلاعات بین سیستم‌های اطلاعاتی کتابخانه‌های دیجیتال در ایران و ارائه الگوی پیشنهادی برای انجام بهینه آن بود. پژوهشگر در این پژوهش به بررسی ۱۱ سیستم اطلاعاتی کتابخانه دیجیتال فعال (در زمان انجام پژوهش، ۱۳۸۸) پرداخت. روش‌های پژوهش مورد استفاده در این پژوهش روش پیمایش تحلیلی جهت بررسی وضعیت موجود سیستم‌های اطلاعاتی کتابخانه‌های دیجیتال ایران و روش تحلیل سیستم نیز جهت تحلیل یافته‌ها و تهیه الگوی مناسب پیشنهادی برای مبادله اطلاعات بود. از این رو در گردآوری اطلاعات اقدام به استفاده از ابزار پرسشنامه شده بود. سه محور اصلی مورد بررسی در این پژوهش نحوه ذخیره‌سازی اطلاعات، ارائه خروجی‌های مناسب اطلاعاتی و شیوه مبادله اطلاعات به لحاظ نحوی بود. از این رو نتایج حاصل از این پژوهش نشان داد که کتاب به عنوان مهمترین ماده اطلاعاتی است که در قالب فایل‌های پی.دی.اف. در کتابخانه‌های دیجیتال مورد بررسی موجود است و در اختیار جامعه کاربران قرار می‌گیرد. همچنین قالب فراداده مارک بر مبنای یونی مارک نیز پر استفاده‌ترین قالب در ذخیره‌سازی اطلاعات می‌باشد. علاوه بر این تفاهم‌نامه Z39.50 سرویس گیرنده نیز توسط حدود ۲۷ درصد جامعه پژوهش مورد استفاده قرار می‌گیرد. متأسفانه هیچ یک از سیستم‌های اطلاعاتی بررسی شده قادر به ارائه خروجی‌های استاندارد نبودند. در نتیجه الگوی پیشنهادی پژوهش مذکور نیز به شرح زیر پیشنهاد شد: استفاده از مدل مجتمع، تفاهم‌نامه او.ای.آی.-پی.ام.اچ، فراداده ساختاری و مدیریتی متس با فراداده توصیفی مُدز، و ایکس.ام.ال. به عنوان فرمت مورد استفاده در لایه زمینه، و ایجاد ارتباط فراداده با محتوا به صورت درونه‌ای و ارائه محتوا پس از شناسایی کاربر (علیپور حافظی، ۱۳۸۸).

صنعت جو و فتحیان (۱۳۹۰) پژوهشی با عنوان "مقایسه کارآمدی اصطلاحنامه و هستی‌شناسی در بازنمون دانش (طراحی و ساخت نمونه هستی‌شناسی اصفا)" را با هدف شناسایی کارآمدی قابلیت‌های هستی‌شناسی در مقایسه با اصطلاحنامه در ساختار سنتی آن است. این پژوهش از نوع کاربردی بود و ابزارهای گردآوری اطلاعات، پرسشنامه ساختار یافته با روش انتومتریک، روش بلند فکر کردن در زمان جستجو و مصاحبه پس از انجام جستجو بود. یافته‌های پژوهش نشان داد میزان کارآمدی هستی‌شناسی در بازنمون دانش، از کارآمدی اصطلاحنامه اصفا بیشتر است. همچنین، با رویکرد تحلیل حوزه می‌توان ابزارهای معنایی کارآمدتری در بازنمون دانش قلمروهای موضوعی در مقایسه با ابزارهای سنتی مانند اصطلاحنامه‌ها طراحی نمود (صنعت جو & فتحیان، ۱۳۹۰). بنا براین پژوهش حاضر بر کارآمدی هستی‌شناسی‌ها بر اصطلاحنامه‌ها در بازنمون دانش تأکید داشت. این مسئله می‌تواند در ارتقاء سطح بازیابی سیستم‌های اطلاعاتی کتابخانه دیجیتال یکپارچه تأثیرگذار باشد. همچنین می‌تواند هستی‌نگاری را به عنوان ابزار مناسب برای یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال معرفی کند.

علیپور حافظی (۱۳۹۰) در پژوهش دیگری با عنوان بررسی هزینه-سازگاری نظام‌های اطلاعاتی نامتجانس در یکپارچه‌سازی اطلاعات: پژوهشی در نرم‌افزارهای کتابخانه‌های دیجیتال ایران به بررسی و شناسایی عناصر هزینه‌ای حاکم بر یکپارچه‌سازی سیستم‌های اطلاعاتی کتابخانه دیجیتال پرداخت. هدف از پژوهش حاضر تعیین هزینه-سازگاری نظام‌های اطلاعاتی نامتجانس در یکپارچه‌سازی اطلاعات کتابخانه‌های دیجیتال ایران بود. پژوهش مذکور از جمله پژوهش‌های کاربردی بوده و از روش پیمایش تحلیلی و ابزار پرسشنامه برای وصول به هدف خود بهره برده است. یافته‌های پژوهش حاضر نشان داد که عناصر هزینه‌ای در سه سطح فنی، محتوا و سازمانی وجود دارند و در این میان کتابخانه‌ها و سازمان‌های متبوع و نیز شرکت‌ها و مؤسسات تولیدکننده سیستم‌های اطلاعاتی کتابخانه دیجیتال از جمله سازمان‌های تأثیر پذیر و تأثیر گذار در فرآیند حرکت از وضعیت موجود به سمت وضعیت مطلوب هستند (علیپور حافظی، ۱۳۹۰).

همچنین علیپور حافظی (۱۳۹۰) در پایان‌نامه کارشناسی ارشد خود با هدف بررسی وضعیت سازماندهی منابع اطلاعاتی در کتابخانه‌های دیجیتال ایران به بررسی کتابخانه‌های دیجیتال ایران از منظر سازماندهی منابع اطلاعاتی دیجیتال پرداخت. پژوهش

مذکور با روش پژوهش پیمایشی و با استفاده از ابزار پرسشنامه انجام گردید. ایشان در پژوهش خود به این نتیجه رسیدند که کتابخانه‌های دیجیتال موجود تاکید اصلی خود را بر منابع اطلاعاتی متنی گذاشته‌اند و برای توصیف منابع اطلاعاتی بیشتر از استانداردهای فراداده دابلین‌گر و مارک استفاده می‌کنند. علاوه بر این کتابخانه‌های دیجیتال مورد بررسی تنها برخی فیلدهای ضروری را جهت توصیف منابع اطلاعاتی تکمیل می‌نمایند و سایر فیلدها خالی می‌مانند. در نهایت ایشان به این نتیجه رسیدند که کتابخانه‌های دیجیتال مورد بررسی در وضعیت مطلوبی از نظر سازماندهی قرار ندارند (علیپور حافظی، ۱۳۹۰).

۲-۱۰-۲. پیشینه‌های خارجی

پروژه وردنت از سال ۱۹۸۵ توسط آزمایشگاه علوم شناختی دانشگاه پرینستون^۱ با هدف تولید واژگان معنایی به زبان انگلیسی شروع شد (Szymański, 2011). در حال حاضر^۲ حدود ۱۵۵۲۸۷ کلمه در شبکه واژگانی آن وارد شده است. وردنت کلمات انگلیسی را در مجموعه‌هایی از مترادف‌ها گروه‌بندی می‌کند، تعاریف کوتاهی از کلمات ارائه می‌کند و روابط معنایی مختلف بین این مجموعه مترادف‌ها را برقرار می‌کند. هدف اصلی آن تلفیق واژه‌نامه با اصطلاحنامه و پشتیبانی از تحلیل متن خودکار و استفاده از آن در نرم‌افزارهای هوش مصنوعی ذکر شده است. وردنت امروزه می‌تواند به عنوان ابزار قدرتمندی در راستای اهداف هستی‌نگاری در علوم رایانه با هدف برقراری روابط معنایی بین واژگان بکار رود. امروزه نسخه مربوط به زبان فارسی وردنت توسط دانشگاه شهید بهشتی تهیه شده است که شامل ۱۸۰۰۰ کلمه فارسی است (شمس فرد، ۱۳۸۹).

اتحادیه اروپا در سال ۱۹۹۰ پروژه‌ای را با عنوان تله‌متیکس برای کتابخانه‌ها^۳ با هدف دستیابی یکپارچه به اطلاعات در کشورهای اروپایی آغاز کرد. پروژه حاضر شامل ۱۰۲ زیر پروژه بود^۴ که در سال ۱۹۹۸ به کار خود خاتمه داد. پروژه تله‌متیکس در سطوح چهارگانه زیر به اجرا درآمد:

- رایانه‌ای کردن اطلاعات کتابشناختی^۵. هدف از این سطح ایجاد، ارتقاء و هماهنگ ساختن فهرست‌ها و اطلاعات کتابشناختی ماشین خوان در اروپا بود. بنابراین بستر کارآمدی کتابخانه‌ها و توسعه اشتراک منابع بین آنها فراهم می‌شود.
- شبکه‌سازی کتابخانه‌ها و روابط درونی سیستم‌ها^۶. هدف از این سطح کمک به راه‌اندازی خدمات شبکه‌ای میان کتابخانه‌ها با کسب اطمینان از وجود قابلیت‌های فنی است که از طریق خدمات ارتباطی جدید و پیشرفت در تعامل سیستم‌های باز در دسترس هستند.
- خدمات نوین کتابخانه‌ای^۷. هدف از این سطح قادر ساختن کتابخانه‌ها به ارائه خدمات هزینه-سودمند از طریق استفاده از فناوری‌های ارتباطی و اطلاعاتی است.
- ابزارها و محصولات کتابخانه‌ای مبتنی بر فناوری^۸. هدف از این سطح فراهم نمودن انگیزه در بازار اروپا با تشویق بخش خصوصی جهت کار با کتابخانه‌ها برای تولید محصولات، ابزارها و خدمات تله‌متیک ماندگار تجاری که اختصاصاً برای

^۱ Cognitive Science Laboratory at Princeton University

^۲ آمار ذکر شده مربوط به ویرایش ۳ آن است که در سال ۲۰۰۷ ارائه شده است.

^۳ Telematics for libraries program

^۴ فهرست کامل پروژه‌ها در آدرس زیر قابل دسترس است:

<http://cordis.europa.eu/libraries/en/projects.html>

همچنین تحلیلی که توسط Brian Kelly روی کل این پروژه‌ها انجام گرفته در آدرس زیر قابل دسترس است:

<http://www.exploit-lib.org/issue7/webwatch/>

^۵ Computerised bibliographies

^۶ Library networking and interconnection of systems

^۷ Innovative library services

کتابخانه‌ها تولید شده‌اند (Creating a European library space telematics for libraries programmes 1990-1998, 2000).

نگاهی گذرا به چهار محور مذکور نشانگر این است که محور دوم با عنوان شبکه‌سازی کتابخانه‌ها و روابط درونی سیستم‌ها به صورت مستقیم و سایر محورهای مذکور به صورت غیرمستقیم با این پژوهش می‌توانند مرتبط باشند. با توجه به اهمیت پروژه تله‌متیکس برای کتابخانه‌ها و به‌ویژه ارتباط مستقیم برخی از زیر پروژه‌های آن در ادامه برخی از زیر پروژه‌های مرتبط با این پروژه معرفی می‌شوند.

یکی از پروژه‌های مرتبط با این پژوهش توسط بینگام^۲ (۱۹۹۵) با عنوان یکپارچه‌سازی کتابخانه و اسناد^۳ اجرا شد. محور اصلی این پروژه دستیابی یکپارچه به منابع الکترونیکی برای مبادله خدمات کتابخانه‌ای جدید بود. این پروژه اقدام به ایجاد، استفاده و ارزیابی خدمات مبادله اسناد چندرسانه‌ای در محیط توزیع‌یافته کرد. پروژه حاضر همانگونه که ذکر شد سرویس جستجو و بازیابی را برای مبادله منابع چندرسانه‌ای در سه کشور اروپایی ایجاد کرد. خدمات مذکور توسط ناشران و متخصصان اروپا در این پروژه به اجرا درآمد (Bingham, 1995). همچنین فیلیپس^۴ در پروژه دیگری با عنوان درگاه Z39.50 اروپا^۵ اقدام به شخصی‌سازی پروتکل مذکور در اروپا کرد. هدف از این پروژه ایجاد و استفاده از خدمات درگاه نمونه با استفاده از پروتکل انسی Z39.50 و ایزو اس.آر.^۶ برای دستیابی پیوسته به فهرست‌ها بود. پروژه حاضر منتج به دستیابی بین‌المللی به فهرست‌ها و اطلاعات کتابشناختی با حداقل تجهیزات لازم بود (Phillips, 1996). بستر ارتباطی استفاده شده برای انتقال داده‌ها استفاده از قابلیت‌های پست الکترونیکی بود. لونگو^۷ (۱۹۹۸) در پروژه جایگزین‌هایی برای دسترس‌پذیری اسناد بین‌المللی (آیدا)^۸ به ایجاد شبکه کتابخانه‌های پرتقالی و ایتالیایی در راستای تهیه خدمات بین کتابخانه‌ای بین‌المللی پرداخت. پروژه مذکور مجموعه‌ای از استانداردها را برای ایتالیا و پرتقال، شامل ایزو آی.ال.ال.^۹ تهیه کرد (LUNGO, 1998). در پروژه دیگری با عنوان دستیابی به فهرست‌های راه دور با استفاده از عملگرهای جستجو و بازیابی^{۱۰}، فاوارو^{۱۱} (۱۹۹۷) نرم‌افزاری عمومی تولید کرد که آپک‌های موجود را قادر می‌کرد تا مانند سیستم‌های جستجو و بازیابی عمل کنند. همچنین پروژه حاضر سرویس گیرنده جستجو و بازیابی با رابط کاربر گرافیکی تولید کرد که پذیرای دستیابی کاربران نهایی به آپک بود. نرم‌افزار مذکور مبتنی بر فیلدهای آپک، مانند نقاط دستیابی، ذخیره شده در واژه‌نامه با ساختار مناسب عمل می‌کند. واژه‌نامه مذکور شامل تمام داده‌های مرتبطی است که برای ترجمه پرس و جوا مورد نیاز هستند (Favaro, 1998). پروژه دیگری توسط مائورر^{۱۲} (۱۹۹۸) با عنوان کتابخانه‌ها: دستیابی الکترونیکی به اطلاعات در شبکه-ها^{۱۳} در راستای پروژه تله‌متیکس به اجرا درآمد. در این پروژه اقدام به گردآوری مجموعه‌ای از منابع چندرسانه‌ای و ارائه دسترسی به آنها از طریق شبکه شد. در بستر شبکه اینترنت هزینه دستیابی به منابع از طریق روال‌های سنتی مانند اشتراک و پرداخت هزینه

¹ Technology-based library products and tools

² Bingham

³ Document and Library Integration

⁴ Phillips

⁵ European SR-Z39.50 Gateway

⁶ ISO SR

⁷ Lungo

⁸ Alternatives for International Document Availability (AIDA)

⁹ ISO ILL

¹⁰ Access to Remote Catalogues by implementing SR target functions

¹¹ Favaro

¹² Maurer

¹³ Libraries: Electronic Remote Access to Information over Networks

برای استفاده‌های موردی و در برخی موارد منابع با دسترسی رایگان انجام می‌شد. در پروژه حاضر از استاندارد فراداده دابلین‌گر استفاده شد (Maurer, 1998).

نگاهی جامع به پروژه‌های ذکر شده در زیر مجموعه پروژه تله‌متیکس نشان می‌دهد که پروژه‌های مرتبط با یکپارچه‌سازی و دستیابی یکپارچه بیشتر به تعامل نحوی پرداخته‌اند و هیچ یک از پروژه‌ها به تعامل معنایی که هدف اصلی پژوهش حاضر می‌باشد نپرداخته‌اند.

چن، فنین، و جاشی^۱ (۲۰۰۳) در گزارشی با عنوان هستی‌نگاری برای محیط‌های رایانشی فراگیر آگاه از زمینه^۲، در زمینه به اشتراک گذاری اطلاعات مبتنی بر زمینه به مطالعه پرداختند. از نظر آنها هستی‌نگاری، مؤلفه کلیدی در ایجاد سیستم‌های رایانشی فراگیر دوردست توزیع یافته است. از این رو پژوهشگران در این پژوهش معماری شکست زمینه^۳ را با استفاده از هستی‌نگاری برای توصیف افراد، فضاها و مفاهیم پیشنهاد دادند. بستر مطالعاتی این پژوهش بر محور مؤلفه‌های جغرافیایی استوار بود. لذا معماری پیشنهادی آنها نیز محدود به این زمینه بود. به طور کلی پژوهشگران تأکید کمی بر علت ارائه خدمات و جنبه‌های مرتبط با آن مانند رابط کاربری و ابزارهای ارائه خدمات سیار داشتند. بنابراین، پژوهش مذکور با پژوهش حاضر از نظر اهداف و راهبرد متفاوت است ولی از نظر مؤلفه‌های مورد استفاده و نیز راهکار اجرایی تا حدودی در یک راستا قرار دارند (Chen, Finin, & Joshi, 2003).

پاتل، کوچ، دوئر، سیناراک^۴ (۲۰۰۵) در پژوهشی با عنوان مبادله معنایی در سیستم‌های کتابخانه دیجیتال^۵ به بررسی نحوه مبادله معنایی در سیستم‌های کتابخانه دیجیتال پرداختند. پژوهش حاضر بخشی از پروژه دلاس است که توسط اتحادیه اروپا از سال ۲۰۰۴ آغاز شده است. اهداف اصلی پروژه دلاس انجام پژوهش‌هایی است که نتایج آنها در حوزه عمومی، و انتقال فناوری از طریق توافقنامه‌های همکاری با حوزه‌های مورد علاقه است (DELOS, 2004). پروژه مبادله معنایی در سیستم‌های کتابخانه دیجیتال، زیر مجموعه پروژه دبلیو.پی.۵: استخراج دانش و مبادله معنایی^۶ است (Patel, Koch, Doerr, & Tsinaraki, 2005). پروژه مورد نظر، پروژه‌ای است که مقرر شده بود تا به مدت ۴۸ ماه به اهداف مورد نظر خود مبنی بر بیان اهمیت مبادله معنایی در سیستم‌های کتابخانه دیجیتال و نحوه استفاده از آن؛ شناسایی انواع مختلف تعامل معنایی؛ شناسایی روابط میان تعامل نحوی و معنایی؛ شناسایی روش‌های مرتبط، پیش‌نیازها، استانداردها و خدمات معنایی؛ و شناسایی روش‌های ارتقاء تعامل معنایی در کتابخانه‌های دیجیتال نائل آید. در پژوهش حاضر، سطوح تعامل معنایی در محیط کتابخانه دیجیتال مبتنی بر ساختارهای داده‌ای، داده‌های طبقه‌بندی شده، و داده‌های واقعی مشخص شده است. گزارش فاز اول پروژه مورد بررسی در سال ۲۰۰۵ منتشر شده است. با وجود این پژوهشگر با بررسی‌هایی که انجام داد موفق به دسترسی به گزارش نهایی این پروژه نشد. به نظر می‌رسد که این پژوهش هنوز به اتمام نرسیده باشد.

وارن و السمیر^۷ (۲۰۰۵) در مقاله پژوهشی خود با عنوان بکارگیری فناوری معنایی در کتابخانه دیجیتال: مطالعه موردی^۸ به توصیف نحوه استفاده از فناوری دانش معنایی در راستای ارتقاء کتابخانه دیجیتال پرداختند. پژوهشگران در این مقاله چالش‌های پژوهشی

¹ Chen, Finin, and Joshi

² An ontology for context-aware pervasive computing environments

³ Context Broker Architecture (CoBrA)

⁴ Patel, Koch, Doerr, Tsinaraki

⁵ Semantic Interoperability in Digital Library Systems

⁶ WP5: Knowledge Extraction and Semantic Interoperability (KESI)

⁷ Warren and Alsmeyer

⁸ Applying semantic technology to a digital library: a case study

حوزه کتابخانه دیجیتال را مورد بررسی قرار دادند و کاربرد فناوری دانش معنایی را در پاسخ‌گویی به این چالش‌ها شناسایی کردند. عملکرد توسعه‌یافته در پروژه سکت^۱ جهت ارتقاء یک کتابخانه دیجیتال مورد استفاده قرار گرفت. ایجاد هستی‌نگاری خوب در چنین نرم‌افزاری مرکزیت داشت و این مقاله رویکرد مهندسی هستی‌نگاری خاص مناسب را توصیف کرده است. در بررسی چالش‌های کتابخانه دیجیتال پژوهشگران در این مقاله با چهار چالش عمده زیر مواجه شدند: دستیابی به مبادله اطلاعات، توصیف اشیاء و مخازن اطلاعاتی دیجیتال، مدیریت مجموعه‌های چندرسانه‌ای، و ارتقاء رابط کاربری و تعامل انسان-رایانه. کتابخانه دیجیتال سکت از فناوری دانش معنایی به منظور پاسخ‌گویی به دو چالش اول و نیز آخرین چالش مذکور استفاده می‌کند. بنا بر این مقاله پژوهشی مورد نظر نحوه اجرای فناوری مذکور را توصیف می‌کند (Warren & Alsmeyer, 2005).

هان^۲ (۲۰۰۵) در مقاله پژوهشی خود با عنوان سیستم کتابخانه دیجیتال مبتنی بر آر.دی.اف.^۳ به بررسی کاربرد آر.دی.اف. در مدیریت دانش و مدیریت فراداده پرداخت. در پروژه حاضر نرم‌افزار مبادله نشریات در کتابخانه‌های دانشگاه آریزونا^۴ به صورت نمونه جهت نشان دادن کاربرد آر.دی.اف. در سیستم‌های کتابخانه دیجیتال مورد استفاده قرار گرفت. در پروژه حاضر سیستم کتابخانه دیجیتال مبتنی بر آر.دی.اف. روشی آسان برای مدیریت منابع دیجیتال با یکپارچه ساختن سایر نرم‌افزارها بدون توجه به قالب‌های فراداده‌ای و حضور وب را فراهم کرد. لازم به ذکر است که نرم‌افزار مبادله نشریات و پایان‌نامه‌های دانشگاه از سال ۲۰۰۵ مبتنی بر آر.دی.اف. عمل می‌کنند. در نهایت مقاله استفاده از آر.دی.اف. (فناوری وب معنایی) را به عنوان رویکردی جدید جهت تسهیل مدیریت دانش و فراداده پیشنهاد می‌کند. به این ترتیب استفاده از فناوری آر.دی.اف. روش‌های جدیدی را برای مدیریت و کشف اطلاعات برای کتابخانه‌ها فراهم می‌کند (Han, 2006).

هوانگ، ک و یانگ^۵ (۲۰۰۵) در پژوهشی با عنوان ارتقاء پرس‌وجوی معنایی کتابخانه دیجیتال با استفاده از مدل استنباط خدمات و محتوا (سی.اس.آی.ام.)^۶ به معرفی مدل سی.اس.آی.ام. در راستای ارتقاء سیستم پرس‌وجوی معنایی در کتابخانه دیجیتال پرداختند. پژوهشگران در پژوهش خود به اهمیت ارتباط بین دو جنبه مهم کتابخانه‌های دیجیتال، محتوا و خدمات، و تأثیر آنها بر تولید روابط معنایی تأکید کردند. پژوهش حاضر مدل سی.اس.آی.ام. را برای استنتاج ۱۵ رابطه میان دو جنبه محتوا و خدمات و تعریف عملکردهایی برای دستکاری این روابط پیشنهاد داد. پژوهشگران در گزارش پژوهش خود ذکر کرده‌اند که به کارگیری این مدل در کتابخانه‌های دیجیتال، پرس‌وجوهای معنایی را افزایش می‌دهد و هنگام ارائه خدمات جدید مانند رابط پرس‌وجوی کتابخانه دیجیتال، برنامه‌ریزی تهیه منابع و سیستم فهرست مشترک مجازی، حجم امور اجرایی را کاهش می‌دهد (Huang, Ke, & Yang, 2005).

گوها^۷ (۲۰۰۶) در مطالعه‌ای با عنوان خدمات معنایی کتابخانه دیجیتال^۸ به مطالعه مسائل مربوط به تعامل معنایی در محیط کتابخانه دیجیتال در حوزه کشاورزی پرداخته است. در این مطالعه چارچوب هستی‌نگاری مدل‌سازی خدمات وب (دبلیو.اس.ام.او.)^۹ برای مواجهه با مشکلات ابهام در سطح هستی‌نگاری، سطح شما و سطح خدمات پیشنهاد شده است. علاوه بر این گوها معماری

^۱ SEKT project (<http://sekt.semanticweb.org>)

^۲ Han

^۳ An RDF-based digital library system

^۴ University of Arizona Libraries

^۵ Huang, Ke, Yang

^۶ Enhancing semantic digital library query using a content and service model (CSIM)

^۷ Guha

^۸ Semantic digital library services

^۹ Web Service Modeling Ontology (WSMO)

پیشنهادی مدل خدمات بر مبنای محتوا و واسط هستی نگاری و سازماندهی محتوا برای بازیابی معنایی اسناد را نیز در گزارش خود ارائه داده و به توصیف آن پرداخته است (Guha, 2006).

می‌هالسا، کورلی، استراپاراوا^۱ (۲۰۰۶) در مقاله‌ای با عنوان معیارهای مشابهت معنایی دانش مبنا و مبتنی بر مجموعه واژگان به بررسی روشی برای سنجش شباهت معنایی^۲، با استفاده از معیارهای مشابهت مبتنی بر دانش و مبتنی بر مجموعه واژگان پرداختند. تمرکز اصلی این مقاله بر سنجش شباهت معنایی متون کوتاه بود. یافته‌های پژوهش نشان داد که روش شباهت معنایی بهتر از روش‌های مبتنی بر شباهت لغوی ساده عمل می‌کند. در روش شباهت معنایی میزان خطا نسبت به روش شباهت لغوی با استفاده از معیار شباهت بُرداری سنتی حدود ۱۳ درصد کمتر است. بنا براین یافته‌های پژوهش حاضر نشان داد که استفاده از روش شباهت معنایی بهتر از سایر روش‌های سنتی برای بازیابی محتوا در متون محدود و کوتاه مانند چکیده‌ها عمل می‌کند (Mihalcea, Corley, & Strapparava, 2006).

آیساک، شلوباخ، متزینگ، و زین^۳ (۲۰۰۸) در مقاله خود با عنوان دستیابی یکپارچه به منابع میراث فرهنگی از طریق بازنمون و تطبیق واژگان کنترل شده^۴، به بررسی چگونگی استفاده از وب معنایی در مرتفع ساختن مسائل مبادله معنایی در حوزه منابع فرهنگی پرداختند. در این مطالعه مسائل عدم تعانس معرفی شدند. جهت برطرف نمودن مسائل موجود در زمینه ایجاد دستیابی یکپارچه، راهکارهای وب معنایی با تمرکز بر بازنمون واژگان کنترل شده و تطابق معنایی آنها معرفی شدند. در این زمینه پروژه‌های نمونه‌ای نیز در متن مقاله مورد اشاره قرار گرفتند که از این راهکار استفاده کرده بودند. در این مقاله دو رویکرد استفاده از مدل سیستم سازماندهی ساده دانش (اس.کی.او.اس.)^۵ و ابزارها و روش‌های تطبیق واژگان خودکار برای فایق آمدن بر مسائل موجود پیشنهاد شد (Issac, Schlobach, Mattheizing, & Zinn, 2008).

پراساد و مدالی^۶ (۲۰۰۸) در مقاله پژوهشی خود با عنوان زیرساخت مناسب کتابخانه‌های دیجیتال معنایی^۷ به ارتقاء بازیابی اطلاعات در کتابخانه‌های دیجیتال با استفاده از زیرساخت معنایی پرداختند. پژوهش حاضر منتج به ارائه مدل عمومی برای ایجاد زیرساخت معنایی کتابخانه‌های دیجیتال شد. در گزارش این پژوهش ذکر شده که هستی نگاری‌های موضوعی کتابخانه‌های دیجیتالی تنها یکی از رویکردهای موجود در زمینه تعامل معنایی در این کتابخانه‌ها هستند. اما همین رویکرد ممکن است با گروه‌هایی که با رویکردهای مختلف برای عملی ساختن بازیابی مؤثر در کتابخانه‌های دیجیتالی کار می‌کنند سازگار باشد (Pasad & Madalli, 2008).

میر، موچکه، و پتراس^۸ (۲۰۰۸) در مقاله پژوهشی خود با عنوان کاهش پیچیدگی معنایی در کتابخانه‌های دیجیتال توزیع یافته^۹ به یکپارچه‌سازی با استفاده از سه روش ارزش افزوده ساختاری با هدف کاهش پیچیدگی‌های معنایی موجود در کتابخانه‌های دیجیتال توزیع یافته در چند مرحله در فرآیند بازیابی اطلاعات (تولید پرس و جو، جستجو و رتبه‌بندی، و رتبه‌بندی مجدد) پرداختند. یکی از رویکردهای موجود برای یکپارچه‌سازی استفاده از پورتال است. نمونه‌ای از پورتال‌ها که در این پژوهش مورد بررسی قرار

¹ Mihalcea, Corley, Strapparava

² Corpus-based and Knowledge-based Measures of Text Semantic Similarity

³ Isaac, Schlobach, Mattheizing, and Zinn

⁴ Integrated access to cultural heritage resources through representation and alignment of controlled vocabularies

⁵ Simple Knowledge Organization System (SKOS) Model

⁶ Prasad and Madalli

⁷ Faceted infrastructure for semantic digital libraries

⁸ Mayer, Mustchke, and Petras

⁹ Reducing semantic complexity in distributed digital libraries

گرفت پورتال علوم و سکودا^۱ بود که مجموعه‌های اطلاعاتی کیفی و ساخت یافته را از بیش از ۴۰ مجموعه اطلاعاتی با استفاده از فناوری موتور جستجو یکپارچه می‌کند. نخستین تجربه‌ها در حوزه پورتال نشانگر برخی از ضعف‌های این رویکرد بود. بنا بر این هدف اصلی این پژوهش پیشنهاد مدل‌هایی برای کاهش پیچیدگی معنایی در کتابخانه‌های دیجیتال نامتجانس بود. در این پژوهش از دو روش برگرفته از علم سنجی و تحلیل شبکه در راستای اهداف این پژوهش برای رتبه‌بندی مجدد یافته‌ها در ویژگی‌های زیر استفاده شد: رتبه‌بندی یافته‌ها توسط نشریات هسته (که با عنوان روش بردفورد شناخته می‌شود) و رتبه‌بندی با مرکزیت نویسندگان در شبکه‌های هم تألیفی. یافته‌های این پژوهش نشان داد که روش‌هایی که اجرا خواهند شد بر پرس‌وجو و نتایج پرس‌وجو تمرکز خواهند داشت و به منظور تأثیر مثبت بر یکدیگر طراحی خواهند شد. به لحاظ مفهومی آنها کیفیت پرس‌وجوها را افزایش خواهند داد و تضمین خواهند داد که مرتبط‌ترین منابع در رتبه‌های بالایی نمایش داده خواهند شد (Mayer, Mutschke, & Petras, 2008).

نیلسون^۲ (۲۰۱۰) در پایان‌نامه دکتری خود با عنوان از میان‌کنش‌پذیری تا هماهنگ‌سازی در استانداردسازی فراداده: طراحی چارچوبی قابل شمول برای هماهنگ‌سازی فراداده^۳ به ترکیب ویژگی‌های فراداده در هماهنگ‌سازی فراداده‌های پراکنده در سیستم‌های اطلاعاتی نامتجانس پرداختند. پایان‌نامه مذکور به ارائه تحلیلی مبتنی بر راهکار از مسائل جاری در هماهنگ‌سازی فراداده پرداخته است. تاکید ویژه این پژوهش بر یکپارچه‌سازی و میان‌کنش‌پذیری میان سیستم‌های اطلاعات کتابخانه دیجیتال، سیستم‌های اطلاعاتی مبتنی بر فناوری آموزش الکترونیکی با استفاده از چارچوب آر.دی.اف. بود. پژوهش حاضر به این نتیجه رسید که چارچوب آر.دی.اف. ابزار مناسبی برای هماهنگ‌سازی فراداده در سیستم‌های اطلاعاتی کتابخانه دیجیتال و آموزش الکترونیکی است (Nilsson, 2010).

۳-۱۰-۲. تحلیل پیشینه‌های پژوهش

بررسی پیشینه‌های داخلی پژوهش نشان از چهار پژوهش و چهار طرح اجرایی دارد. یکی از پژوهش‌ها اختصاص به بررسی وضعیت مبادله اطلاعات در محیط کتابخانه‌های سنتی دارد. دو پژوهش دیگر مرتبط با مبادله اطلاعات در بستر نحوی در کتابخانه‌های دیجیتال می‌باشند و پژوهش دیگر مرتبط با کارآمدی هستی‌نگاری‌ها نسبت به اصطلاحنامه‌ها در بازیابی اطلاعات در سیستم‌های اطلاعاتی نوین می‌باشد. از طرفی نیز از بین طرح‌های اجرا شده دو طرح امین و غدیر مرتبط با اشتراک منابع بین کتابخانه‌ها و تسهیل در دستیابی به کتابخانه‌ها در بستر کتابخانه‌های سنتی هستند. یک طرح مرتبط با گردآوری اطلاعات کتابشناختی در پایگاه اطلاعاتی واحد و ایجاد امکان دستیابی به آن برای کاربران است. متأسفانه با توجه به روزآمد نبودن محتواهای پایگاه اطلاعاتی مذکور و عدم جامعیت آن از کارآمدی مناسبی برخوردار نیست. طرح دیگر اجرا شده در یکپارچه‌سازی نیز مرتبط با کنسرسیوم محتوای ملی است که به علت عدم جامعیت و عدم پوشش محتواهای اطلاعات نرم‌افزارهای کتابخانه‌ای هنوز نتوانسته در بستر نحوی پوشش مناسبی بر محتواهای کتابخانه‌ای بدهد. تنها طرح متناسب با یکپارچه‌سازی نحوی آخرین طرح اشاره شده است و همانگونه که ذکر شد تا حدودی به یکپارچه‌سازی نحوی پرداخته است. علاوه بر موارد ذکر شده یکپارچه‌سازی‌های نحوی دیگری نیز توسط نرم‌افزارهای کتابخانه‌ای شکل گرفته است. در این موارد یکپارچه‌سازی در نرم‌افزارهای متجانس شکل گرفته

¹ Vascoda

² Nilsson

³ From interoperability to harmonization in metadata standardization: designing an evolvable framework for metadata harmonization

است. فعالیت‌های مذکور نیز از جمله فعالیت‌های مناسب در یکپارچه‌سازی نحوی هستند. مهم‌ترین مسئله‌ای که در این زمینه وجود دارد متجانس بودن^۱ سیستم‌های اطلاعاتی می‌باشد و راهکار اجرایی مناسبی برای یکپارچه‌سازی نحوی در سیستم‌های اطلاعاتی نامتجانس نمی‌باشد. لازم به ذکر است که یکپارچه‌سازی نحوی پیش‌درآمد یکپارچه‌سازی معنایی است و بدون اجرای یکپارچه‌سازی نحوی امکان کاربردی کردن یکپارچه‌سازی معنایی فراهم نخواهد بود. با وجود فعالیت‌های پژوهشی و اجرایی ذکر شده هنوز پژوهش و یا طرح اجرایی در زمینه یکپارچه‌سازی معنایی در داخل کشور شکل نگرفته است.

مطالعه پژوهش‌های خارجی در زمینه یکپارچه‌سازی سیستم‌های اطلاعاتی نیز نشانگر این است که پروژه‌ها و پژوهش‌های اجرا شده در این زمینه از سال ۱۹۸۵ آغاز شده‌اند و تا کنون ۱۲ پروژه در این زمینه به اجرا درآمده است. برخی از این پروژه‌ها به تولید ابزارهای مورد نیاز مانند وُردنت (Szymański, 2011) یا هستی نگاری‌ها (Chen, Finin, & Joshi, 2003) در یکپارچه‌سازی معنایی اختصاص یافته‌اند. برخی دیگر از این پروژه‌ها نیز زیرمجموعه‌هایی از پروژه‌های بزرگ‌تر بودند که مرتبط با یکپارچه‌سازی کتابخانه‌ها و یکپارچه‌سازی معنایی بودند، از جمله این پروژه‌ها می‌توان به پروژه تله‌متیکس و دلاس اشاره کرد. همانگونه که قبلاً نیز اشاره شد، پروژه تله‌متیکس شامل ۱۰۲ پروژه در چهار محور اصلی بود که یکی از محورهای اصلی آن مرتبط با شبکه‌سازی کتابخانه‌ها بود. از بین پروژه‌های مذکور تعداد پنج پروژه نسبتاً مرتبط در بالا معرفی شدند که مطالعه گزارش‌های مربوطه نشانگر توجه پروژه‌های مذکور به تعامل نحوی در یکپارچه‌سازی کتابخانه‌ها بود و پروژه مرتبطی با یکپارچه‌سازی معنایی در پروژه جامع مذکور به اجرا در نیامده است. همچنین در پروژه دلاس که از سال ۲۰۰۴ آغاز شده زیر پروژه‌های متعددی تعریف و اجرا شده‌اند که یکی از آنها مربوط به مبادله معنایی در سیستم‌های کتابخانه دیجیتال بود.

سایر پروژه‌ها و پژوهش‌های معرفی شده نیز مستقیماً با یکپارچه‌سازی سیستم‌های اطلاعاتی کتابخانه دیجیتال مرتبط بودند. برخی از این پژوهش‌ها اختصاص به مطالعه موردی (Warren & Alsmeyer, 2005) (Guha, 2006) (Issac, Schlobach, 2008) (Mattheizing, & Zinn, 2008) داشتند، برخی مبتنی بر فناوری خاصی مانند آر.دی.اف. (Han, 2006) بودند، برخی به معرفی مدل خاصی در حوزه بازایی معنایی اختصاص (Huang, Ke, & Yang, 2005) (Pasad & Madalli, 2008) یافته بودند، و برخی نیز به روش‌های یکپارچه‌سازی (Mayer, Mutschke, & Petras, 2008) پرداخته بودند.

به طور کلی با بررسی پژوهش‌ها و پروژه‌های اجرا شده در حوزه تعامل معنایی در کتابخانه‌های دیجیتال مشخص می‌شود که تا کنون پژوهش و یا پروژه خاصی به یکپارچه‌سازی معنایی در داخل کشور نپرداخته است. همچنین در پژوهش‌ها و پروژه‌های خارجی نیز تنها یک مورد، پژوهش مایر، موچکه و پتراس، تا حدودی به حوزه مورد نظر در این پژوهش پرداخته است و سایر موارد مرتبط دیگر به بخش‌هایی از موضوع مورد پژوهش در این حوزه پرداخته‌اند. لذا تحلیل پژوهش‌های پیشین نشان از فضای بکری در حوزه یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال توزیع یافته دارد که پژوهش حاضر سعی به پرداختن به این حوزه در کتابخانه‌های دیجیتالی توزیع یافته در ایران دارد.

^۱ سیستم‌های اطلاعاتی متجانس به سیستم‌های اطلاعاتی گفته می‌شود که دارای ساختار و معماری یکسانی هستند. به عنوان مثال تمامی کتابخانه‌هایی که از یک سیستم اطلاعاتی مشخص استفاده می‌کنند. از جمله نمونه‌های اجرا شده توسط نرم‌افزارهای کتابخانه‌ای می‌توان به نرم‌افزارهای کتابخانه‌اس سیمرغ، پارس آذرخش و نرم‌افزار کتابخانه دیجیتال پیام مشرق اشاره کرد.

فصل ۳

روش‌شناسی پژوهش

۳-۱. نوع و روش پژوهش

در پژوهش حاضر به لحاظ ساختار پژوهش از روش‌های مختلفی نیز بهره گرفته شده است. در بخش ادبیات پژوهش، شناسایی و تحلیل پیشینه‌های پژوهش، و نیز شناسایی و بهره‌گیری از مؤلفه‌های مورد نظر در پژوهش از روش سندی یا کتابخانه‌ای استفاده شده است. در مرحله دوم پژوهش برای شناسایی کیفیت و نوع محتوای کتابشناختی کتابخانه‌های دیجیتال ایران از روش پیمایش تحلیلی استفاده شده است. در این روش از ابزارهای معمول آن برای گردآوری و تحلیل یافته‌ها استفاده شده است. در مرحله سوم برای ارائه الگوی پیشنهادی پژوهش نیز از روش تحلیل سیستم استفاده شده است.

۳-۲. شیوه و ابزار گردآوری اطلاعات

در بخش اول پژوهش که از روش سندی استفاده شده است، برای گردآوری اطلاعات اقدام به جستجو در پایگاه‌های اطلاعاتی مهم و کاربردی مانند امرالد، الزویر، آی.ای.ای.، لیستا، ایستا، و غیره انجام گرفت. سپس با بررسی و مطالعه چکیده‌های منابع اطلاعاتی بازایی شده اقدام به انتخاب منابع مرتبط در گام اول گردید. سپس در مرحله بعدی منابع انتخاب شده به طور کامل مورد بررسی و مطالعه قرار گرفت که از میان آنها منابع اطلاعاتی کاملاً مرتبط انتخاب شدند و در طول فرآیند انجام پژوهش بویژه در بخش مبانی نظری و ادبیات پژوهش مورد استفاده قرار گرفتند.

در بخش دوم پژوهش که از روش پیمایش تحلیلی استفاده شده است، برای گردآوری اطلاعات از ابزار پرسشنامه و در بخشی نیز (برای پاسخ به پرسش اساسی ۵) از سیاهه واری و مشاهده استفاده شده است. با توجه به اینکه تا کنون پژوهشی در این سطح در سطح کشور انجام نگرفته است و نیز اینکه ابزار مشابهی برای این منظور موجود نبود، پژوهشگر با استفاده از منابع اطلاعاتی، پژوهش‌ها و نمونه‌های موجود اقدام به تهیه پرسشنامه پژوهشگر ساخته (پیوست ۱) و سیاهه واری (پیوست ۲) کرده است. مراحل پالایش پرسشنامه در بخش بعدی همین فصل توضیح داده شده است. پرسشنامه نهایی شده برای گردآوری اطلاعات در اختیار جامعه پژوهش حاضر قرار گرفته است و پس از تکمیل در مراحل بعدی پژوهش جهت توصیف و تحلیل مورد استفاده قرار گرفته است. برای کسب اطمینان به یافته‌های پژوهش، پژوهشگر سعی نموده است تا حد امکان با مراجعه به محل کتابخانه‌های دیجیتال مورد بررسی با استفاده از ابزار مشاهده مستقیم اقدام به تصحیح پاسخ‌ها نماید. بنابراین در این فرآیند برخی از داده‌های گردآوری شده با توجه به مشاهده صورت گرفته در این مرحله اصلاح شدند. این فرآیند باعث شده است تا یافته‌های پژوهش از اطمینان بیشتری برخوردار باشند.

در بخش سوم پژوهش که از روش تحلیل سیستم استفاده شده است، پس از تحلیل یافته‌های پژوهش و نیز با استفاده از یافته‌های پژوهش‌های پیشین اقدام به طراحی الگوی پیشنهادی یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال شده است. لذا با استفاده از یافته‌های این پژوهش و نیز با بهره گرفتن از تجربیات پژوهش‌های پیشین در این بخش سعی شده است تا الگوی کاربردی مناسبی با

وضعیت موجود کتابخانه‌های دیجیتال ایران تهیه و ارائه شود. از این رو، برای بهره‌گیری و آزمون کاربردی بودن الگوی پیشنهادی، الگوی مذکور در طرح سمات به جهت اعتبار سنجی مورد آزمون قرار گرفته است.

۳-۳. روایی و پایایی ابزار پژوهش

برای کسب روایی ابزار پژوهش اقدام به توزیع پرسشنامه تهیه شده به پنج نفر از متخصصان حوزه علوم کتابداری و اطلاع‌رسانی آشنا به حوزه فناوری اطلاعات شد. پس از دریافت نظرات متخصصان و اعمال اصلاحات پیشنهادی در پرسشنامه اولیه، ویرایش جدیدی از پرسشنامه حاصل شد. پرسشنامه اصلاح شده در مرحله دوم در پنج کتابخانه دیجیتالی که حائز تمام ویژگی‌های جامعه پژوهش بودند نیز توزیع گردید و از آنها خواسته شد تا ضمن پاسخ‌گویی به پرسش‌ها موارد مبهم و یا ناقص را نیز مشخص نمایند. با دریافت پرسشنامه‌ها و اعمال اصلاحات پیشنهادی، ویرایش جدید از پرسشنامه حاصل شد که دارای روایی مناسبی می‌باشد.

پس از دریافت نظرات متخصصان و نمونه‌های پاسخ داده شده اصلاحات نهایی روی پرسشنامه انجام گرفت و پرسشنامه نهایی تهیه شد. با توجه به اینکه داده‌های مورد نیاز در این پژوهش داده‌های اسمی بودند و امکان دریافت آزمون پایایی وجود نداشت آزمونی برای کسب پایایی ابزار پژوهش صورت نگرفت.

۳-۴. جامعه پژوهش

جامعه پژوهش در گام نخست منابع اطلاعاتی مرتبطی بودند که به جهت آشنایی با ادبیات پژوهش مورد استفاده پژوهشگر قرار گرفتند. در گام بعدی نیز به جهت بررسی وضعیت کتابخانه‌های دیجیتال فعال در کشور، کتابخانه‌های دیجیتال فعال جزو جامعه پژوهش حاضر قلمداد شدند. بررسی در این زمینه نشان داد که آشفتگی بسیار زیادی در این زمینه وجود دارد که علت اصلی آن دیدگاه‌های متنوعی است که در این زمینه وجود دارد. لذا سعی شد تا معیارهایی جهت شناسایی کتابخانه‌های دیجیتال فعال تهیه شود تا از طریق معیارهای مذکور اقدام به شناسایی کتابخانه‌های دیجیتال نمایم. بنا براین، در پژوهش حاضر، کتابخانه دیجیتال به کتابخانه‌ای گفته شده است که در آن: (۱) منابع اطلاعاتی دیجیتالی مبتنی بر اصول مجموعه‌سازی، گردآوری یا تولید شده باشند؛ (۲) کارکنان متخصصی به انتخاب و سازماندهی منابع اطلاعاتی بپردازند؛ (۳) اشیاء دیجیتال به جامعه کاربران آن کتابخانه ارائه شود؛ (۴) مدیریت منابع اطلاعاتی با استفاده از سازوکارهای مدیریت فایل، مانند مدیریت در نمایش و ارائه اطلاعات به کاربران، و مدیریت سطح دسترسی صورت پذیرد؛ و (۵) اطلاعات دیجیتالی به مدت زمان طولانی به مثابه منابع اطلاعاتی پایدار قابل دسترس باشند. کتابخانه‌هایی که همه این عناصر را دارا باشند، در این پژوهش کتابخانه دیجیتال نامیده شده‌اند.

جهت شناسایی کتابخانه‌های دیجیتال فعال ابتدا به دنبال فهرستی جامع از این کتابخانه‌ها بودیم. ولی متأسفانه هیچ مرجع مناسب و کاملی در این زمینه یافت نشد. از این رو، اقدام به شناسایی نرم‌افزارهای کتابخانه دیجیتال فعال در کشور کردیم. بررسی در این زمینه منتج به فهرست ۱۱ نرم‌افزار زیر شد: (فهرست زیر به ترتیب الفبا مرتب شده است)

- نرم‌افزار کتابخانه دیجیتال آذرخش^۱

^۱ <http://www.parslib.com/web/guest/551>

- نرم‌افزار کتابخانه دیجیتال ارم^۱
- نرم‌افزار کتابخانه دیجیتال پایپروس^۲
- نرم‌افزار کتابخانه دیجیتال پروان^۳
- نرم‌افزار کتابخانه دیجیتال پیام^۴
- نرم‌افزار کتابخانه دیجیتال تبیان^۵
- نرم‌افزار کتابخانه دیجیتال دید^۶
- نرم‌افزار کتابخانه دیجیتال سنا^۷
- نرم‌افزار کتابخانه دیجیتال سیمرخ^۸
- نرم‌افزار کتابخانه دیجیتال نورلیب^۹
- نرم‌افزار کتابخانه دیجیتال وستا^{۱۰}

بررسی اولیه وبگاه نرم‌افزارهای مذکور نشان داد که برخی از این نرم‌افزارها فقط در یک کتابخانه دیجیتال مورد استفاده قرار می‌گیرند. از جمله این نرم‌افزارها می‌توان نرم‌افزار کتابخانه دیجیتال دید، نورلیب، تبیان، سنا، ارم و وستا را نام برد. سایر نرم‌افزارها مانند سیمرخ، آذرخش، پیام، پروان و پایپروس با توجه به استفاده‌های متعددی که از آنها شده جزو نرم‌افزارهای کتابخانه دیجیتال تجاری هستند. بررسی در فهرست مشتریان نرم‌افزارهای تجاری نیز نشان داد که برخی از مشتریان نرم‌افزارهای مذکور هنوز در حال سرویس دهی نیستند. همچنین لازم به ذکر است که برخی از کتابخانه‌های دیجیتال نیز، مانند کتابخانه دیجیتال دانشگاه علوم پزشکی اردبیل، از نرم‌افزارهای منبع باز^{۱۱} استفاده کرده‌اند. به طور کلی بررسی در این زمینه منتج به تهیه فهرستی از کتابخانه‌های دیجیتال به تعداد ۴۶ کتابخانه دیجیتال شد.^{۱۲} در ادامه اقدام به بررسی اولیه ۴۶ کتابخانه دیجیتالی ذکر شده در فهرست ضمیمه شد. بررسی اولیه نشان داد که ۱۴ مورد از ۴۶ کتابخانه مورد بررسی در تعریف کتابخانه دیجیتال پژوهش حاضر قرار نمی‌گیرند، یا در حال حاضر کتابخانه دیجیتال نیستند و یا در حال نصب و راه‌اندازی کتابخانه دیجیتال هستند. از این رو، جامعه پژوهش حاضر به ۳۲ کتابخانه دیجیتال محدود شد. سپس پرسشنامه پژوهش به ۳۲ کتابخانه دیجیتال مورد نظر ارسال شود و نسبت به مورد برای هر مورد مشاوره‌هایی نسبت به محتوای پرسشنامه داده شد. در انتها ۲۶ پرسشنامه تکمیل شد. با پیگیری‌های متعدد به عمل آمده متأسفانه

^۱ <http://www.dlib.ir/fa/index.aspx>

^۲ فاقد سایت قابل بازدید

^۳ <http://www.parvanpajoooh.ir>

^۴ <http://www.payamnet.com>

^۵ <http://www.tebyan.net/Library.html>

^۶ www.did.ir/

^۷ <http://database.irandoc.ac.ir/DL/Search/>

^۸ <http://www.nosa.com/nosasoft/webUI/Home.aspx>

^۹ <http://www.noorlib.ir/dlview/default.aspx>

^{۱۰} <http://www.vestasoft.ir/CyberLib.aspx>

^{۱۱} Open source

^{۱۲} فهرست کتابخانه‌های دیجیتال در ضمیمه ۲ همین گزارش ارائه شده است.

موفق به گردآوری اطلاعات از تمام ۳۲ کتابخانه نشدیم. با توجه به اینکه موفق به گردآوری اطلاعات مربوط به ۸۱/۲۵ درصد از جامعه پژوهش به پرسشنامه‌ها شدیم، بر همین اساس توصیف و تحلیل‌ها را بر اساس ۲۶ پرسشنامه گردآوری شده انجام دادیم.

۳-۵. روش تجزیه و تحلیل یافته‌ها

با توجه به پرسشنامه تهیه شده توسط پژوهشگر، جامعه مطالعاتی مذکور از جهت یکپارچه‌سازی معنایی مورد بررسی قرار گرفتند. داده‌های گردآوری شده، با استفاده از نرم‌افزار اکسل^۱ توصیف، سپس برآوردهای آماری نظیر فراوانی، درصد فراوانی، فراوانی تجمعی و نما و طراحی نمودارهای مورد نیاز با استفاده از نرم‌افزار مذکور انجام گرفت. بنابراین با استفاده از داده‌های آماری گردآوری شده و توصیف و تحلیل آنها، که در بخش بعد به تفصیل بیان شده‌اند، قادر به پاسخ‌گویی به پرسش‌های اساسی طرح شده در ابتدای پژوهش حاضر باشیم.

^۱ Excel

فصل ۴

توصیف یافته‌های پژوهش

۱-۴. مقدمه

با توجه به استفاده از پرسشنامه برای گردآوری اطلاعات مورد نیاز بخشی از پژوهش حاضر (پرسش‌های اساسی شماره ۴ و ۵ ذکر شده در فصل اول) در این بخش اقدام به توصیف یافته‌های حاصل از پرسشنامه کرده‌ایم. لازم به ذکر است که در پرسشنامه اقدام به شناسایی وضعیت موجود محتواهای فراداده‌ای کتابخانه‌های دیجیتال ایران که مرتبط با برقراری تعامل معنایی می‌باشند پرداخته‌ایم. بنا براین یافته‌ها و توصیف آنها نیز مرتبط با این محور می‌باشند.

۲-۴. توصیف یافته‌های پژوهش

۱-۲-۴. توصیف یافته‌های مربوط به وضعیت موجود محتواهای پیشینه‌های کتابشناختی

در رابطه با انواع منابع اطلاعاتی موجود در کتابخانه‌های دیجیتال مورد بررسی نتایج زیر حاصل شد:

جدول ۴-۱. انواع منابع اطلاعاتی موجود در کتابخانه‌های دیجیتال ایران^۱

انواع منابع اطلاعاتی	تعداد	درصد
کتاب دیجیتال	۲۶	۱۰۰
نشریات دیجیتال	۱۳	۵۰
مقالات دیجیتال	۱۵	۵۸
عکس‌های دیجیتال	۵	۱۹
نقشه‌های دیجیتال	۲	۸
گزارش‌های فنی دیجیتال	۷	۲۷
پایان‌نامه‌های دیجیتال	۱۶	۶۲
فایل‌های صوتی دیجیتال	۱۳	۵۰
فایل‌های ویدئویی دیجیتال	۱۰	۳۸
منابع چندرسانه‌ای	۸	۳۱
منابع اینترنتی	۷	۲۷
سایر	۱	۴

^۱ با توجه به وجود پایگاه‌های اطلاعاتی مجزا برای منابع مرتبط مانند نشریات و مقالات به صورت مجزا یافته‌های این حوزه به صورت جدا از هم بیان شده‌اند. همچنین فایل‌های صوتی فقط محدود به قابلیت صدا می‌باشند و فایل‌های ویدئویی دارای قابلیت صدا و تصویر و فایل‌های چندرسانه‌ای منابعی را شامل می‌شوند که ترکیبی از چند رسانه (بیش از دو رسانه) در آنها مورد استفاده قرار گرفته‌اند.

یافته‌ها نشانگر این است که کتاب در قالب دیجیتال پرکاربردترین منبع اطلاعاتی با استفاده در ۱۰۰ درصد جامعه پژوهش مورد استفاده قرار می‌گیرد. علاوه بر این منبع اطلاعاتی از پایان‌نامه‌ها در قالب دیجیتال با ۶۲ درصد و مقالات دیجیتال با ۵۸ و نشریات دیجیتال نیز با ۵۰ درصد بیشترین استفاده را در جامعه مورد پژوهش دارند. در مقابل نقشه‌ها با ۸ درصد و عکس‌های دیجیتال با ۱۹ درصد نیز کمترین کاربرد را در جامعه مورد مطالعه دارند.

با توجه به اینکه برای برقراری تعامل معنایی میان محتواهای دیجیتال نیاز به شناسایی فیلدهای اصلی مورد نیاز در تعامل است، طی پرسشی به این موضوع پرداخته شد.

جدول ۴-۲. تکمیل فیلدهای کتابشناختی توسط کتابخانه‌های دیجیتال مورد بررسی

فیلدهای کتابشناختی	تعداد	درصد
عنوان	۲۶	۱۰۰
پدیدآور	۲۵	۹۶
پدیدآور همکار	۲۴	۹۲
توصیف (کلیدواژه نمایه)	۱۶	۶۲
موضوع	۲۴	۹۲
ناشر	۲۵	۹۶
شابک	۲۰	۷۷
شاپا	۱۲	۴۶
شابم	۱	۴
دی.او.آی.	۲	۸

بررسی در مورد تکمیل فیلدهای اطلاعاتی نشان داد که فیلد اطلاعاتی عنوان آثار با ۱۰۰ درصد، پدیدآور و ناشر با ۹۶ درصد، پدیدآور همکار و موضوع با ۹۲ درصد با بیشترین اقبال در تکمیل فیلدهای توسط کتابخانه‌های دیجیتال مورد بررسی مواجه هستند و در مقابل فیلدهای مربوط به کد شناسایی مانند شابم با ۴ درصد، دی.او.آی. با ۸ درصد کمترین استفاده را در کتابخانه‌های دیجیتال مورد بررسی دارا هستند.

موضوع مهم دیگر در امکان ایجاد روابط معنایی در یکپارچه‌سازی کتابخانه دیجیتال کسب اطلاعات در زمینه میزان و نحوه مستندسازی اسامی توسط سازمان‌های مورد مطالعه می‌باشد.

جدول ۴-۳. مستند سازی اسامی پدیدآوردندگان و ناشران

مستندسازی اسامی پدیدآوردندگان	تعداد	درصد	مستندسازی اسامی ناشران	تعداد	درصد
اسامی پدیدآوردندگان مستند نمی شوند	۳	۱۲	اسامی ناشران مستند نمی شوند	۹	۳۵
قسمت شناسنامه اثر در دست بررسی	۶	۲۳	قسمت شناسنامه اثر در دست بررسی	۷	۲۷
فهرست مستند اسامی مشاهیر و مؤلفان	۱۶	۶۲	فهرست مستند اسامی نهادها، سازمان ها و موسسات دولتی ایران	۵	۱۹
پایگاه اطلاعاتی کتابخانه ملی	۲۰	۷۷	مرجع نشر	۴	۱۵
پایگاه اطلاعاتی کتابخانه کنگره امریکا	۱۸	۶۹	وبگاه ناشر	۵	۱۹
فهرست مستندات موجود در نرم افزار مورد استفاده	۴	۱۵	سایر	۰	۰

همانگونه که از جدول ۴-۳ نیز برمی‌آید حدود ۸۸ درصد از کتابخانه‌های دیجیتال مورد بررسی اقدام به مستندسازی اسامی پدیدآوردندگان می‌نمایند. این در حالی است که تنها ۶۵ درصد کتابخانه‌های دیجیتال مورد بررسی اقدام به مستند سازی نام ناشران می‌نمایند. علاوه بر این استفاده از پایگاه اطلاعاتی مستند پدیدآوردندگان کتابخانه ملی با ۷۷ درصد، پایگاه اطلاعاتی کتابخانه کنگره امریکا با ۶۹ درصد و فهرست مستند اسامی مشاهیر و مولفات نیز با ۶۲ درصد پر استفاده‌ترین پایگاه‌های اطلاعاتی جهت مستند سازی نام پدیدآوردندگان هستند.

پراستفاده‌ترین منبع نیز برای مستندسازی نام ناشران خود منبع در دست بررسی با ۲۷ درصد و فهرست مستند اسامی نهادها، سازمان-ها و موسسات دولتی ایران است. همین مساله نشانگر عدم وجود منبع مناسب جهت استفاده در مستندسازی نام ناشران است.

از جمله عناصر مهم دیگر در شناسایی امکان یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال کسب اطلاعات در زمینه نحوه تحلیل موضوعی منابع اطلاعاتی در پایگاه‌های اطلاعاتی است.

جدول ۴-۴. روش‌های تحلیل موضوعی منابع اطلاعاتی توسط کتابخانه‌های دیجیتال مورد بررسی

روش های تحلیل موضوعی منابع اطلاعاتی	تعداد	درصد
موضوع دهی کنترل شده با استفاده از سرعنوان های موضوعی	۲۲	۸۵
نمایه سازی آزاد با استفاده از کلیدواژه های انتخابی از منبع مورد نظر	۱۰	۳۸
مایه سازی آزاد با استفاده از کلیدواژه های انتخابی نمایه ساز با توجه به تجربه تخصصی	۵	۱۹
نمایه سازی کنترل شده با استفاده از اصطلاحنامه ها	۱۴	۵۴
مایه سازی کنترل شده با استفاده از هستی نگاری ها	۰	۰

یافته‌ها در زمینه روش‌های تحلیل موضوعی منابع اطلاعاتی در کتابخانه‌های دیجیتال مورد بررسی نشانگر این است که حدود ۸۵ درصد جامعه مورد مطالعه از سرعنوان‌های موضوعی برای موضوع دادن به منابع اطلاعاتی به ویژه کتاب‌های دیجیتال استفاده می‌کنند. همچنین ۵۴ درصد مراکز مورد بررسی از اصطلاحنامه‌ها برای توصیف منابع اطلاعاتی استفاده می‌نمایند. در مقابل یافته‌ها نشان داد که هیچ یک از کتابخانه‌های دیجیتال مورد بررسی از هستی‌نگاری‌ها برای این موضوع استفاده نمی‌کنند. بر این اساس در صورتی که نیاز به برقراری یکپارچه‌سازی معنایی باشد نیاز به تبدیل اصطلاحنامه‌های موجود به ساختار هستی‌نگاری است. در پرسش دیگری که در رابطه با انتخاب موضوع برای منابع اطلاعاتی پرسیده شد یافته‌های ذکر شده از جدول ۴-۵ حاصل شد:

جدول ۴-۵. انتخاب موضوع برای تحلیل منابع اطلاعاتی توسط کتابخانه‌های دیجیتال مورد بررسی

انواع منابع اطلاعاتی	تعداد	درصد
کتاب ها	۲۴	۹۲
نشریات	۷	۲۷
مقالات	۸	۳۱
عکس ها	۲	۸
نقشه ها	۲	۸
گزارش های فنی	۵	۱۹
پایان نامه ها	۷	۲۷
شنیداری	۸	۳۱
ویدئوئی	۵	۱۹
چند رسانه ای	۶	۲۳

یافته‌ها نشان داد که کتاب‌ها با حدود ۹۲ درصد بیشترین منابع اطلاعاتی هستند که از موضوع‌دهی برای تحلیل محتوای آنها استفاده می‌شود. در سایر منابع اطلاعاتی از اختصاص موضوع استفاده بسیار کمی صورت می‌گیرد.

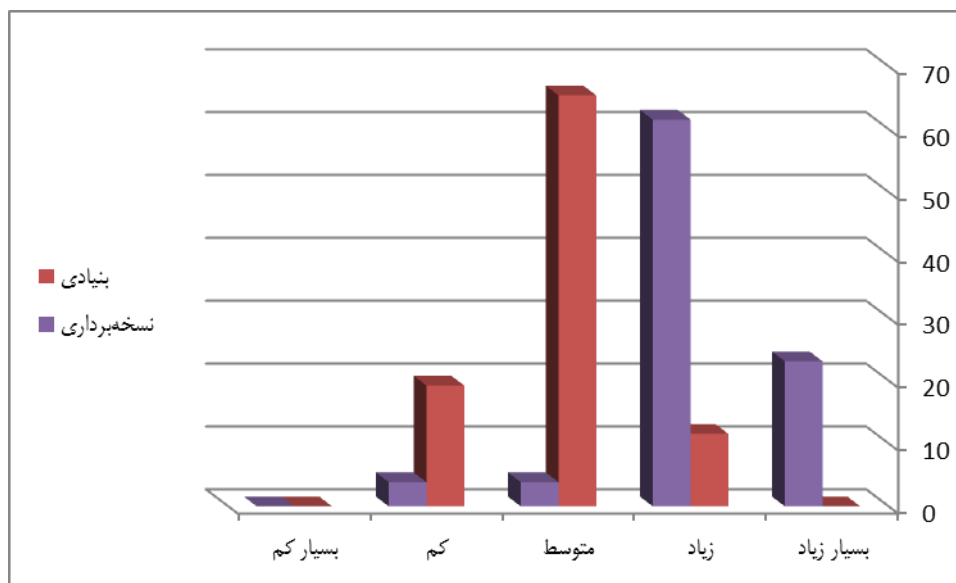
در مقابل پرسشی با همین مضمون برای استفاده از کلیدواژه در تحلیل محتوای منابع اطلاعاتی پرسیده شد. یافته‌ها در این زمینه در جدول ۴-۶ ذکر شده‌اند:

جدول ۴-۶. انتخاب کلیدواژه برای تحلیل منابع اطلاعاتی توسط کتابخانه‌های دیجیتال مورد بررسی

انواع منابع اطلاعاتی	تعداد	درصد
کتاب‌ها	۸	۳۱
نشریات	۴	۱۵
مقالات	۸	۳۱
عکس‌ها	۲	۸
نقشه‌ها	۲	۸
گزارش‌های فنی	۷	۲۷
پایان‌نامه‌ها	۱۳	۵۰
شنیداری	۶	۲۳
ویدئویی	۵	۱۹
چندرسانه‌ای	۴	۱۵

یافته‌ها حاکی از این است که پایان‌نامه‌ها با حدود ۵۰ درصد بیشترین منابع اطلاعاتی هستند که توسط کتابخانه‌های دیجیتال توسط کلیدواژه‌ها مورد تحلیل محتوا قرار می‌گیرند. پس از پایان‌نامه‌ها، کتاب‌ها و مقالات با حدود ۳۱ درصد و گزارش‌های فنی با حدود ۲۷ درصد توسط کلیدواژه‌ها مورد تحلیل محتوا قرار می‌گیرند. لازم به ذکر است که همانگونه که در جدول ۴-۱ ذکر شده است کتاب‌ها، پایان‌نامه‌ها و مقالات دیجیتال بیشترین منابعی هستند که توسط کتابخانه‌های دیجیتال در اختیار کاربران قرار می‌گیرند.

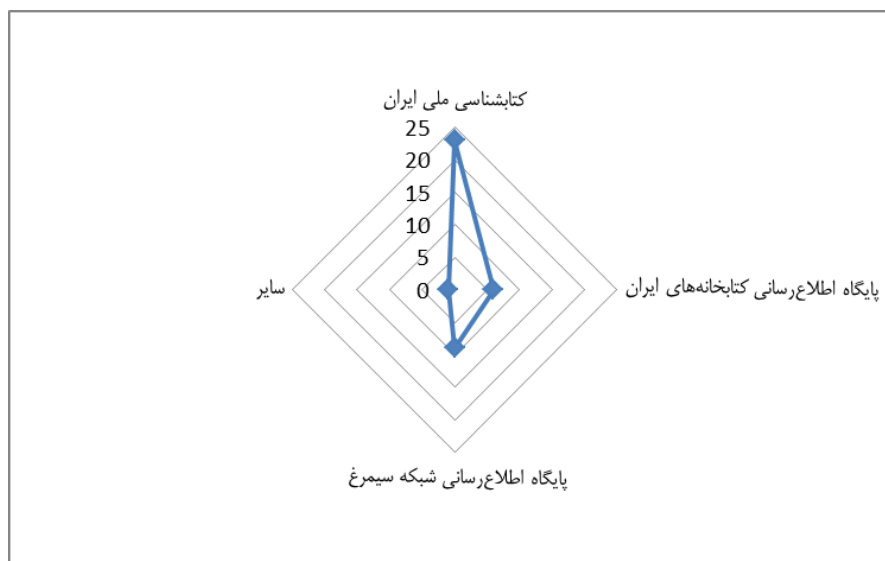
در پرسش دیگری از مخاطبان در مورد شیوه سازماندهی منابع اطلاعاتی (بنیادی یا نسخه‌برداری) پرسش شد. یافته‌ها نشان داد که ۵۴ درصد کتابخانه‌های دیجیتال مورد بررسی از شیوه نسخه‌برداری و ۴۶ درصد دیگر از شیوه سازماندهی بنیادی برای تحلیل و توصیف منابع اطلاعاتی خود استفاده می‌کنند. یافته‌ها در زمینه میزان بهره‌برداری از هر یک از روش‌های بالا در شکل ۴-۱ نشان داده شده است:



شکل ۴-۱. میزان بهره‌گیری از روش‌های فهرست‌نویسی بنیادی و نسخه‌برداری توسط کتابخانه‌های دیجیتال مورد بررسی

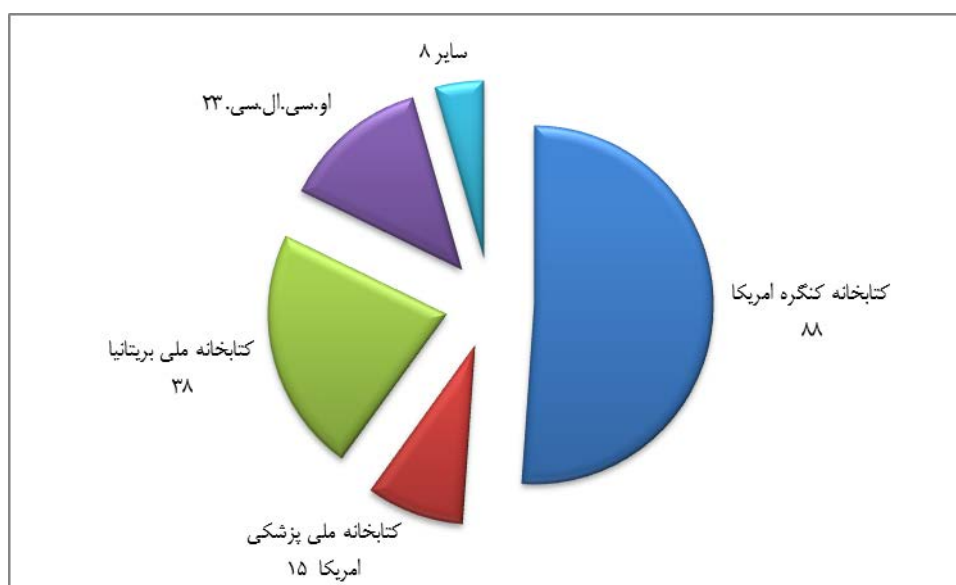
یافته‌ها همانگونه که از شکل ۴-۱ نیز برمی‌آید نشانگر این است که حدود ۷۷ درصد جامعه مورد مطالعه بیش از حد متوسط از فهرست‌نویسی بنیادی و حدود ۸۵ درصد جامعه مورد مطالعه نیز از فهرست‌نویسی به شیوه نسخه‌برداری استفاده می‌کنند. با توجه به تبدیل قالب منابع اطلاعاتی از قالب فیزیکی به دیجیتالی به ویژه در نسخه‌های خطی و منابع قدیمی در اکثر این کتابخانه‌ها استفاده از فهرست‌نویسی بنیادی قابل توجه است. علاوه بر این کتابخانه‌های دیجیتالی مورد بررسی برای برخی از منابع خود از فهرست‌نویسی بنیادی و برای برخی دیگر از منابع از روش فهرست‌نویسی نسخه‌برداری (شناسایی و استفاده از فراداده‌های موجود در کتابخانه‌های بزرگ کشور و بین‌المللی) استفاده می‌کنند. همین امر باعث شده تا جمع درصدهای دو مورد بالا (فهرست‌نویسی بنیادی و نسخه‌برداری) بیش از صد درصد باشد.

در ادامه پرسش حاضر از مخاطبان پرسیده شد که در صورت استفاده از نسخه‌برداری کدام پایگاه‌های اطلاعاتی بیشترین استفاده را برای این منظور دارند.



شکل ۴-۲. نمودار پایگاه‌های اطلاعاتی مورد استفاده برای نسخه‌برداری جهت توصیف و تحلیل منابع اطلاعاتی

یافته‌های ذکر شده در شکل ۴-۲ نشان می‌دهد که کتابشناسی ملی ایران با ۸۸ درصد بیشترین استفاده را برای نسخه‌برداری جهت توصیف و تحلیل منابع اطلاعاتی دارد. در مورد منابع مورد استفاده جهت نسخه‌برداری در رابطه با منابع اطلاعاتی غیرفارسی نیز یافته‌های شکل ۴-۳ در این پژوهش حاصل شده است:

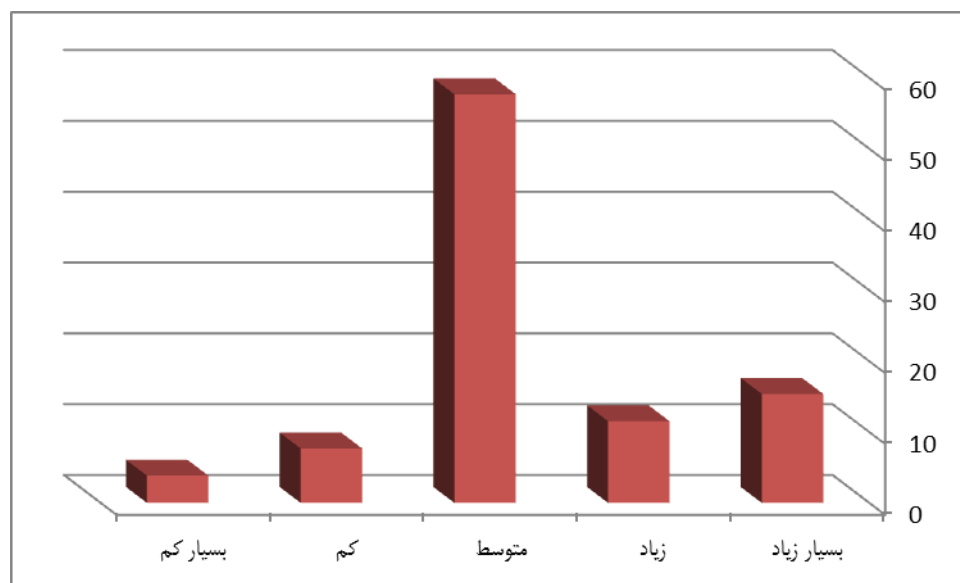


شکل ۴-۳. نمودار پایگاه‌های اطلاعاتی غیر فارسی مورد استفاده برای نسخه‌برداری جهت توصیف و تحلیل منابع اطلاعاتی

همانگونه که از شکل ۴-۳ نیز برمی‌آید پایگاه اطلاعاتی کتابخانه کنگره آمریکا با حدود ۸۸ درصد بیشترین استفاده را در نسخه-برداری جهت توصیف و تحلیل منابع اطلاعاتی توسط کتابخانه‌های دیجیتال مورد بررسی دارا است.

۷۳ □ بررسی یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال و ارائه الگوی پیشنهادی ...

در ادامه پرسشی در مورد اصلاح فیلدهای اطلاعاتی نسخه برداری شده از پایگاه‌های اطلاعاتی و میزان اصلاحات پرسیده شد. یافته‌ها در این زمینه در شکل ۴-۴ نشان داده شده‌اند:



شکل ۴-۴. نمودار میزان اصلاح فیلدهای اطلاعاتی توسط کتابخانه‌های دیجیتال مورد بررسی

همانگونه که از شکل نیز بر می‌آید حدود ۸۵ درصد جامعه مورد مطالعه بیش از حد متوسط نسبت به اصلاح فیلدهای اطلاعاتی اخذ شده از پایگاه‌های اطلاعاتی دیگر استفاده می‌کنند.

در ادامه پرسشی در مورد استفاده از منابع مختلف در فهرست‌نویسی بنیادی جهت شناسایی منابع مورد استفاده و میزان استفاده از آنها پرسیده شد.

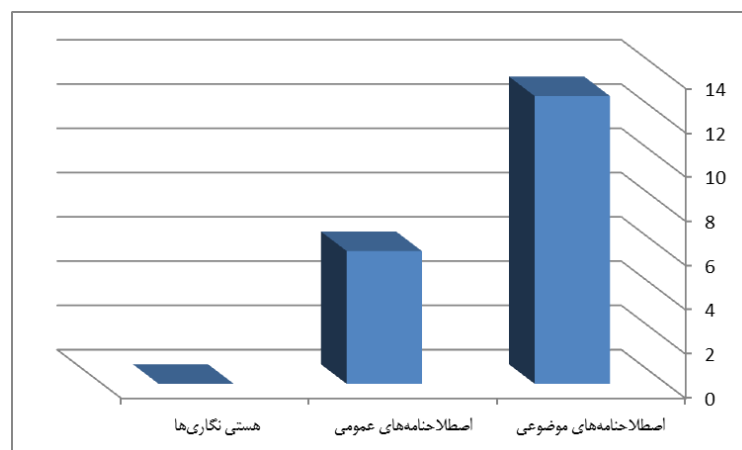
جدول ۴-۷. منابع مورد استفاده در فهرست‌نویسی بنیادی توسط کتابخانه‌های دیجیتال مورد بررسی

منبع مورد استفاده برای فهرست‌نویسی بنیادی	تعداد	درصد
استفاده از فهرست نویسی پیش از انتشار	۱۵	۵۸
سرعنوان‌های موضوعی فارسی	۱۸	۶۹
سرعنوان‌های موضوعی کنگره امریکا	۱۹	۷۳
سرعنوان‌های موضوعی سیرز	۰	۰

یافته‌ها همانگونه که از جدول ۴-۷ نیز بر می‌آید نشانگر این است که اکثر کتابخانه‌های دیجیتال مورد بررسی از سرعنوان‌های موضوعی فارسی و کنگره امریکا و نیز از فهرست‌نویسی پیش از انتشار منتشر شده در آثار به ویژه کتاب‌های فارسی استفاده می‌-

نمایند. لازم به ذکر است که در کنار فهرست نویسی به شیوه نسخه برداری کتابخانه‌های دیجیتالی برای برخی از منابع خود از فهرست‌نویسی بنیادی نیز استفاده می‌نمایند.

در ادامه پرسشی در زمینه شناسایی منابع مورد استفاده در نمایه‌سازی کنترل شده از مخاطبان پرسیده شد. یافته‌ها در این زمینه در شکل ۴-۵ نشان داده شده است:



شکل ۴-۵. ابزارهای مورد استفاده در نمایه‌سازی کنترل شده در کتابخانه‌های دیجیتال مورد بررسی

یافته‌های نشان داده شده در شکل ۴-۵ نشانگر این است که حدود ۵۰ درصد کتابخانه‌های دیجیتال مورد بررسی از اصطلاحنامه‌های موضوعی و ۲۳ درصد نیز از اصطلاحنامه‌های عمومی در نمایه‌سازی کنترل شده منابع اطلاعاتی استفاده می‌کنند. در مقابل هستی‌نگاری‌ها هنوز توسط جامعه مورد مطالعه مورد استفاده قرار نمی‌گیرند.

در پرسش دیگری در زمینه اصطلاحنامه‌های مورد استفاده جهت نمایه‌سازی کنترل شده یافته‌های زیر حاصل شده است:

جدول ۴-۸. اصطلاحنامه‌های مورد استفاده جهت نمایه‌سازی کنترل شده

بررسی یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال و ارائه الگوی پیشنهادی ... □ ۷۵

اصطلاحنامه های موضوعی	تعداد	درصد	اصطلاحنامه های موضوعی	تعداد	درصد
اصطلاحنامه کشاورزی	۴	۱۵	اصطلاحنامه پزشکی فارسی	۴	۱۵
اصطلاحنامه ارتقاء بهداشت	۳	۱۲	اصطلاحنامه جامعه شناسی	۴	۱۵
اصطلاحنامه ریاضیات	۶	۲۳	اصطلاحنامه شیمی	۶	۲۳
اصطلاحنامه علوم زمین	۳	۱۲	اصطلاحنامه علوم زیستی	۳	۱۲
اصطلاحنامه فنی و مهندسی	۸	۳۱	اصطلاحنامه فیزیک	۶	۲۳
اصطلاحنامه مدیریت بحران	۳	۱۲	اصطلاحنامه علوم قرآن	۵	۱۹
اصطلاحنامه فلسفه	۵	۱۹	اصطلاحنامه منطق	۴	۱۵
اصطلاحنامه اصول فقه	۵	۱۹	اصطلاحنامه اخلاق	۴	۱۵
اصطلاحنامه کلام اسلامی	۵	۱۹	سایر	۲	۸

یافته‌های ذکر شده در جدول بالا نشان از استفاده عام از اصطلاحنامه‌های موجود در نمایه‌سازی منابع اطلاعاتی دارد.

پرسش ذکر شده در بالا در مورد منابع اطلاعاتی لاتین نیز پرسیده شد و یافته‌های حاصل از آن در جدول ۴-۹ نشان داده شده است:

جدول ۴-۹. اصطلاحنامه‌های لاتین مورد استفاده جهت نمایه‌سازی کنترل شده

اصطلاحنامه های موضوعی لاتین	تعداد	درصد
Agrovoc	۳	۱۲
Chemistry	۳	۱۲
Engineering	۵	۱۹
Health promotion	۱	۴
Physics	۳	۱۲
Inspec	۳	۱۲
Biological Science	۳	۱۲
Crisis management	۲	۸
Geosciences	۱	۴
Mathematics	۴	۱۵
Sociology	۲	۸

یافته‌ها در زمینه استفاده از اصطلاحنامه‌های لاتین در نمایه‌سازی منابع اطلاعاتی نیز همانگونه که در جدول بالا نشان داده شده است نشان می‌دهد که به طور عمومی اکثر اصطلاحنامه‌های مطرح لاتین در کتابخانه‌های دیجیتال مورد بررسی مورد استفاده قرار می‌گیرند.

در مورد استفاده از اصطلاحنامه‌های عمومی نیز یافته‌ها حاکی از این است که اصطلاحنامه سه زبانه اصفا با ۳۱ درصد و اصطلاحنامه نما با ۱۹ درصد پر استفاده‌ترین اصطلاحنامه‌ها توسط کتابخانه‌های دیجیتال هستند. در مورد استفاده از هستی‌نگاری‌ها نیز یافته‌ها نشان داد که هیچ هستی‌نگاری توسط کتابخانه‌های دیجیتال مورد مطالعه ایجاد و یا مورد استفاده قرار نمی‌گیرد. همین مساله نشان از توجه ویژه به این منبع بسیار مهم در نمایه‌سازی و تحلیل منابع اطلاعاتی دارد که باید بسیار مورد توجه قرار گیرد.

۲-۲-۴. توصیف یافته‌های مربوط به وضعیت موجود خروجی‌های فراداده نرم‌افزارهای کتابخانه‌های دیجیتال مورد بررسی

در این راستا پرسش‌هایی با استفاده از سیاهه واری طرح گردید و بر اساس آن از مسئولان شرکت‌های نرم‌افزار کتابخانه دیجیتال موجود در سطح کشور اقدام به گردآوری اطلاعات در زمینه وضعیت ارائه خروجی‌های فراداده نرم‌افزارهای کتابخانه دیجیتال شد. با توجه به تعداد نرم‌افزارهای مرتبط فعال، ۱۱ سیاهه واری تکمیل شد. همچنین به علت کم اهمیت بودن اسامی نرم‌افزارهای مورد بررسی در تحلیل یافته‌ها، معادل‌های حرفی برای هر یک از نرم‌افزارها در جداول توصیفی در نظر گرفته شد. معادل‌های حرفی و اسامی نرم‌افزارهای مورد بررسی به ترتیب حروف الفبای نام نرم‌افزار مورد بررسی در زیر ارائه شده‌اند:

- | | |
|---|--------------------------------------|
| A | ■ نرم‌افزار کتابخانه دیجیتال آذرخش |
| B | ■ نرم‌افزار کتابخانه دیجیتال ارم |
| C | ■ نرم‌افزار کتابخانه دیجیتال پایپروس |
| D | ■ نرم‌افزار کتابخانه دیجیتال پروان |
| E | ■ نرم‌افزار کتابخانه دیجیتال پیام |
| F | ■ نرم‌افزار کتابخانه دیجیتال تبیان |
| G | ■ نرم‌افزار کتابخانه دیجیتال دید |
| H | ■ نرم‌افزار کتابخانه دیجیتال سیمرخ |
| I | ■ نرم‌افزار کتابخانه دیجیتال نور |
| J | ■ نرم‌افزار کتابخانه دیجیتال وستا |

در این رابطه ۸ پرسش اصلی طرح شده است که در ادامه به توصیف یافته‌های مرتبط پرداخته شده است. پرسش ۱ به بررسی خروجی‌های فراداده‌ای پرداخته است. داده‌های برگرفته در این رابطه در جدول ۴-۱۰ ارائه شده است.

جدول ۴-۱۰. خروجی‌های فراداده‌ای نرم‌افزارهای مورد بررسی

	استاندارد در لایه زمینه					استاندارد در لایه بالاتر				
	ایزو	متن ساختیافته	ایکس.ام.ال.	رابط او.دی.بی.سی.	اچ.تی.ام.ال.	مارک	مارک ایکس.ام.ال.	دابلین گز	مدز	متس
A	✓					✓	✓	✓	✓	
B			✓			✓				
C			✓					✓		✓
D			✓					✓		
E	✓		✓		✓	✓	✓	✓		✓
F					✓				✓	
G			✓					✓		
H	✓		✓			✓		✓		
I										
J	✓		✓			✓				
تعداد	۴	۰	۷	۰	۲	۵	۲	۶	۲	۲
درصد	۴۰	۰	۷۰	۰	۲۰	۵۰	۲۰	۶۰	۲۰	۲۰

یافته‌ها در زمینه خروجی‌های فراداده نشانگر این است که در زمینه استانداردها در لایه زمینه در حال حاضر حدود ۷۰ درصد نرم‌افزارهای مورد بررسی از ایکس.ام.ال. استفاده می‌کنند. همچنین ۴۰ درصد از کتابخانه‌ها نیز از محیط متنی ایزو ۲۷۰۹ برای این منظور استفاده می‌کنند. از نظر استانداردها در لایه بالاتر برای فراداده نیز مارک و دابلین گز هر کدام با ۵۰ درصد استفاده توسط نرم‌افزارهای مورد بررسی پر استفاده‌ترین استانداردها می‌باشند. همچنین مارک ایکس.ام.ال.، مدز و متس نیز هر کدام با ۲ مورد استفاده (معادل یا ۲۰ درصد جامعه مورد مطالعه) دارای کمترین استفاده در نرم‌افزارهای مورد بررسی هستند.

در مورد ارائه شمای خروجی مورد استفاده توسط نرم‌افزارهای مورد بررسی نیز مشخص شد که هیچکدام از نرم‌افزارها اقدام به ارائه شمای مورد استفاده نمی‌کنند. بنا براین این مسئله می‌تواند چالش‌هایی را در ایجاد سیستم‌های اطلاعاتی یکپارچه فراهم نماید.

در زمینه نحوه ایجاد ارتباط بین خروجی فراداده و فایل منبع که موضوع پرسش سوم این بخش از پژوهش حاضر می‌باشد نیز یافته‌های زیر که در جدول ۴-۱۱ ارائه شده، حاصل شده است.

جدول ۴-۱۱. نحوه ارتباط فراداده با فایل منبع

عدم وجود ارتباط	قابلیت فرامتن	به صورت دروانه‌ای	سایر
A		✓	
B	✓		
C		✓	
D		✓	
E		✓	
F	✓		
G	✓		
H		✓	
I	✓		خروجی فراداده وجود ندارد.
J		✓	
تعداد	۱	۴	۵
درصد	۱۰	۴۰	۵۰

یافته‌های جدول ۴-۱۱ نشان می‌دهد که در یک مورد از نرم‌افزارهای مورد بررسی به خاطر عدم توجه مناسب به بحث فراداده خروجی فراداده‌ای در اختیار کاربر قرار نمی‌گیرد. در این نرم‌افزار اطلاعات مقدماتی مانند عنوان، نویسنده، ناشر و تاریخ نشر در مورد منابع مورد استفاده قرار می‌گیرد و بازیابی اطلاعات شامل بریده‌ای از متن و سپس خود متن خواهد بود. ولی قابلیت فرامتن و ارتباط درونه‌ای در فیلدهای داده‌ای هر کدام در حدود ۴۰ درصد نرم‌افزارهای مورد بررسی برای برقراری ارتباط بین فراداده و شیء دیجیتال مورد استفاده قرار می‌گیرند.

در زمینه وجود امکان نمایه‌سازی داده‌های نرم‌افزارهای مورد بررسی توسط موتورهای جستجوی عمومی که پرسش چهارم سیاهه واریسی را به خود اختصاص داده نیز جدول ۴-۱۲ گویای یافته‌های این پژوهش در این زمینه می‌باشد.

جدول ۴-۱۲. نمایه‌سازی و دسترسی به فراداده توسط موتورهای جستجوی عمومی

	بله (با برنامه ریزی)	بله (به صورت عادی)	خیر
A		✓	
B	✓		
C		✓	
D			✓

✓			E
	✓		F
		✓	G
		✓	H
	✓		I
✓			J
۳	۴	۳	تعداد
۳۰	۴۰	۳۰	درصد

یافته‌های جدول بالا نشان می‌دهد که امکان نمایه‌سازی محتوای فراداده‌ای توسط موتورهای جستجوی عمومی وب در حدود ۷۰ درصد نرم‌افزارهای مورد بررسی یا وجود دارد و یا امکان ایجاد آن در نرم‌افزارها وجود دارد و تنها در ۳۰ درصد نرم‌افزارهای مورد بررسی این قابلیت وجود ندارد و یا برنامه‌ای برای ایجاد این قابلیت وجود ندارد. پرسش پنجم این بخش به امکان مبادله اطلاعات موجود در سیستم‌های اطلاعاتی مورد بررسی پرداخته است که در جدول ۴-۱۳ به آن پرداخته شده است.

جدول ۴-۱۳. وجود امکان مبادله اطلاعات

	بله	خیر	تا حدودی	توضیحات
A	✓			WSDL
B		✓		
C		✓		
D		✓		
E	✓			Z39.50
F	✓			
G		✓		
H	✓			Z39.50
I	✓			WSDL
J		✓		
تعداد	۵	۵	۰	
درصد	۵۰	۵۰	۰	

یافته‌ها در این زمینه همان‌گونه که از جدول بالا نیز بر می‌آید نشانگر این است که قابلیت مبادله اطلاعات در حدود نیمی از نرم‌افزارهای مورد مبادله وجود دارد. لازم به ذکر است که در بیش از ۵۰ درصد نرم‌افزارهایی که از این قابلیت برخوردار هستند نیز این قابلیت تنها جهت دریافت اطلاعات وجود دارد نه برای ارسال اطلاعات. از طرفی نیز این قابلیت در حدود ۵۰ درصد نرم‌افزارهایی که از این قابلیت برخوردار هستند از طریق تفاهم‌نامه Z39.50 انجام می‌پذیرد و در ۵۰ درصد بقیه نیز این قابلیت از طریق وب سرویس انجام می‌گیرد.

پرسش ششم در زمینه امکان اجرای قابلیت مبادله اطلاعات در نرم‌افزارهایی که از این قابلیت برخوردار نیستند نیز نشان داد که در هیچ‌یک از نرم‌افزارهای مورد بحث امکان اجرای آن در آینده نزدیک وجود ندارد.

یافته‌ها در زمینه سطح مبادله اطلاعات نیز که محتوای پرسش هفتم پژوهش حاضر را به خود اختصاص داده بود نشان داد که به غیر از یک نرم‌افزار این قابلیت در سایر موارد تنها در سطح فراداده می‌باشد و تنها در یک نرم‌افزار است که این قابلیت در سطح فراداده و اشیاء دیجیتال وجود دارد.

یافته‌ها در پاسخ به پرسش هشتم پژوهش حاضر که مربوط به نوع تفاهم‌نامه مورد استفاده بود در جدول ۴-۱۴ نشان داده شده است.

جدول ۴-۱۴. استفاده از تفاهم‌نامه جهت مبادله اطلاعات

	Z39.50	اس.آر.یو./ دبلیو.	او.ای.آی.	سایر	توضیحات
A	✓				
B					وجود ندارد
C					وجود ندارد
D					وجود ندارد
E	✓	✓			
F	✓				
G					وجود ندارد
H	✓				
I				✓	WSDL
J					وجود ندارد
تعداد	۴	۱	۰	۱	

یافته‌های جدول ۴-۱۴ نشان می‌دهد که در ۴ نرم‌افزار از ۶ نرم‌افزاری که از قابلیت مبادله اطلاعات برخوردار هستند، تفاهم‌نامه Z39.50 مورد استفاده قرار می‌گیرد و تنها در یک نرم‌افزار از قابلیت وب سرویس‌ها برای این منظور استفاده می‌شود. همان‌گونه که در بالا نیز اشاره شد در بیش از ۵۰ درصد این نرم‌افزارها نیز این تفاهم‌نامه فقط در حد سرویس گیرنده مورد استفاده قرار می‌گیرد.

پرسش نهم سیاهه واریسی اختصاص به فرمت‌های مورد استفاده در لایه زمینه، لایه بالاتر و استاندارد برای فراداده پوششی یافته بود که یافته‌های حاصل از آن در جدول ۴-۱۵ ارائه شده است.

جدول ۴-۱۵. فرمت‌های مورد استفاده برای مبادله اطلاعات

توضیحات	استاندارد در لایه بالاتر برای ابر داده ساختاری و مدیریتی		استاندارد در لایه بالاتر				استاندارد در لایه زمینه					
	متن	ای.ای.دی.	مدز	دابلین کر	مارک ایکس.ام.ال	مارک	ای.جی.ام.ال	رابط اودی.بی.سی.	ایکس.ام.ال	متن ساخته	ایزو	
					✓	✓			✓		✓	A
وجود ندارد												B
وجود ندارد												C
وجود ندارد												D
				✓	✓	✓		✓	✓		✓	E
						✓	✓					F
وجود ندارد												G
				✓	✓	✓			✓		✓	H
استاندارد مشخصی استفاده نمی‌شود.										✓		I
وجود ندارد												J
تعداد				۲	۳	۴	۱	۱	۳	۱	۳	
درصد				۲۰	۳۰	۴۰	۱۰	۱۰	۳۰	۱۰	۳۰	

یافته‌های جدول ۴-۱۵ در زمینه فرمت‌های مورد استفاده برای مبادله اطلاعات نشان می‌دهد که ایکس.ام.ال. و ایزو ۲۷۰۹ با حدود ۶۰ درصد بیشترین استفاده را در لایه زیرین به عنوان محمل مبادله اطلاعات دارند. در زمینه فرمت‌های فراداده‌ای مورد استفاده برای مبادله اطلاعات نیز مارک، مارک ایکس.ام.ال. و دابلین کر بیشترین استفاده را در این زمینه دارند. با وجود این هیچ‌یک از نرم‌افزارهای مورد بررسی از قابلیت فرمت پوششی مانند متس برای مبادله اطلاعات استفاده نمی‌کنند.

فصل ۵

تحلیل یافته‌های پژوهش

۱-۵. مقدمه

در این فصل اقدام به تحلیل یافته‌های پژوهش شده است که با استفاده از پرسشنامه گردآوری شده و در فصل پیشین توصیف شده‌اند. برای وصول به اهداف پژوهش حاضر لازم بود تا بخشی از یافته‌ها از طریق پرسشنامه محقق ساخته گردآوری شود، بخشی با استفاده از سیاهه واری و بخشی نیز لازم است از طریق دریافت خروجی‌های استاندارد از نرم‌افزارهای مورد مطالعه گردآوری شوند. بنابراین با توجه به یافته‌های گردآوری شده از طریق پرسشنامه و سیاهه واری و مشاهده که در فصل پیشین توصیف شدند، در این فصل اقدام به تحلیل این یافته‌ها می‌شود. ساختار تحلیل یافته‌ها نیز مبتنی بر پرسش‌های اساسی است که در فصل اول متن حاضر به آنها پرداخته شده است. بر این اساس به ترتیب پرسش‌های اساسی اقدام به تحلیل یافته‌ها در این فصل می‌نمایم. لازم به ذکر است که پرسش‌های اساسی پژوهش حاضر به شرح زیر تعیین شده‌اند:

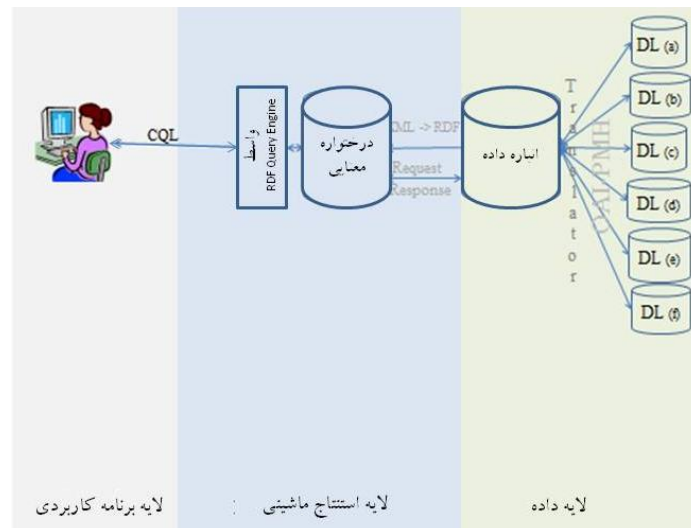
۱. عناصر دخیل در یکپارچه‌سازی معنایی کتابخانه‌های دیجیتال کدام‌ها هستند؟
۲. از چه روش‌هایی برای یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال استفاده می‌شود؟
۳. از چه ابزارهایی در یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال استفاده می‌شود؟
۴. وضعیت موجود محتواهای پیشنهادی کتابشناختی کتابخانه‌های دیجیتال مورد بررسی چگونه است؟
۵. وضعیت موجود خروجی‌های فراداده نرم‌افزارهای کتابخانه دیجیتال مورد بررسی چگونه است؟
۶. الگوی مناسب برای یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال ایران چیست؟

۲-۵. تحلیل یافته‌ها

۱. عناصر دخیل در یکپارچه‌سازی معنایی کتابخانه‌های دیجیتال کدام‌ها هستند؟
برای پاسخ به این پرسش اقدام به گردآوری و بیان داده‌ها در فصل ۲ شد. داده‌های ارائه شده در این زمینه برگرفته از پروژه‌های یکپارچه‌سازی معنایی بود که پیش از این برای ایجاد یکپارچه‌سازی معنایی در سیستم‌های اطلاعاتی متجانس و نامتجانس ایجاد و راه‌اندازی و مورد استفاده قرار گرفته بودند. با توجه به یافته‌های ذکر شده عناصر دخیل در یکپارچه‌سازی معنایی بستگی به رویکرد (Shen R. , 2006) و متجانس یا نامتجانس بودن سیستم‌های اطلاعاتی حاضر در شبکه یکپارچه‌سازی معنایی دارد. در سیستم‌های اطلاعاتی متجانس همه سیستم‌های اطلاعاتی حاضر در شبکه یکپارچه‌سازی از یک نرم‌افزار استفاده می‌کنند. در این صورت با توجه به اینکه روال‌ها، خروجی‌ها و استانداردهای مورد استفاده یکی است فقط کافی است مجموعه واژگان و روابط معنایی بین آنها با استفاده از یک اصطلاحنامه یا هستی‌نگاری و یا مجموعه‌ای از آنها (برای پوشش کامل موضوعی) برقرار گردد. البته لازم به ذکر است که این رویکرد مورد نظر پژوهش حاضر نیست و به همین علت به صورت دقیق و کامل به آن نمی‌پردازیم. با توجه به اینکه موضوع اصلی این پژوهش مربوط به یکپارچه‌سازی معنایی در سیستم‌های اطلاعاتی نامتجانس است در اینجا به عناصر دخیل در یکپارچه‌سازی چنین سیستم‌هایی پرداخته شده است. برای ایجاد یکپارچه‌سازی معنایی نیاز به زیرساخت فناوری‌ها مانند پروتکل‌های مبادله اطلاعات (Alipour-Hafezi, Horri, Shiri, Ghaebi, 2010)، نظام‌های کدگذاری و نظام‌های شناساگر، استانداردهای فراداده، مسائل سازمانی و قانونی و پایگاه‌های دانشی معنایی مانند هستی‌نگاری‌ها و سیستم‌های سازماندهی دانش (DELOS, 2004) است (Patel, Koch, Doerr, & Tsinaraki, 2005).

۲. از چه روش‌هایی برای یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال استفاده می‌شود؟

با توجه به مطالب ذکر شده در فصل ۲ در مورد روش‌های متعدد ایجاد یکپارچه‌سازی معنایی و با تحلیل محتواهای بیان شده در این زمینه می‌توان اظهار داشت که یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتالی با توجه به پژوهش‌های پیشین به دو روش اصلی انجام می‌گیرد. روش اول استفاده از مترجم‌ها برای ترجمه استانداردهای فراداده به استاندارد مورد نظر در سیستم و ماشین استنتاج^۱ است (Vdovjak & Houben, 2001) (Patel, Koch, Doerr, & Tsinaraki, 2005) (Guha, 2006) (Shen R., 2006). در این روش فراداده‌های موجود در سیستم‌های اطلاعاتی عضو شبکه یکپارچه‌سازی با استفاده از تفاهم‌نامه‌های مبتنی بر برداشت اطلاعات در پایگاه اطلاعاتی واسط گردآوری می‌شوند. در این مرحله نیاز به استفاده از مترجم‌هایی است تا فراداده‌ها با استفاده از استانداردهای مختلف به استاندارد مورد نظر در سیستم تبدیل شوند. پس از ایجاد پایگاه اطلاعاتی واسط و ترسیم فرآیند گردآوری فراداده‌ها نیاز به افزودن لایه دیگری برای استنتاج است. در این لایه با استفاده از شبکه‌های معنایی اقدام به برقراری روابط معنایی میان فراداده‌ها و ارائه داده‌های معنایی برای پرس و جوهای کاربران می‌شود. در شکل ۵-۱ روش انجام چنین یکپارچه‌سازی نشان داده شده است.



شکل ۵-۱. روش استفاده از مترجم‌ها و ماشین استنتاج

شکل ۵-۱ نشانگر این است که لایه داده^۲ مربوط به یکپارچه‌سازی نحوی می‌شود. در این لایه فراداده‌ها با استفاده از تفاهم‌نامه او.ای.آی. گردآوری شده و در انبار داده^۳ نگهداری می‌شود. در این لایه کتابخانه‌های دیجیتال عضو شبکه یکپارچه کتابخانه‌ای امکان مرور پایگاه فراداده خود را به خزانه‌گر مورد استفاده در انبار داده می‌دهد (علیپور حافظی م. و ۱۳۹۰). در مسیر انتقال فراداده‌ها مترجمی وجود دارد که فراداده‌های مورد استفاده در کتابخانه‌های دیجیتال را به استاندارد مورد نیاز در انبار داده ترجمه می‌کند.

در گام بعدی داده‌های معنایی مستخرج از فیلدهای فراداده با ساختار سلسله مراتبی موجود در پایگاه داده پیش‌همارا^۴ مطابقت داده شده و به انضمام روابط سلسله مراتبی معنایی در پایگاه اطلاعاتی نگهداری می‌شود. جستجوی کاربر در این پایگاه اطلاعاتی شامل

^۱ Inference Machine

^۲ Data Layer

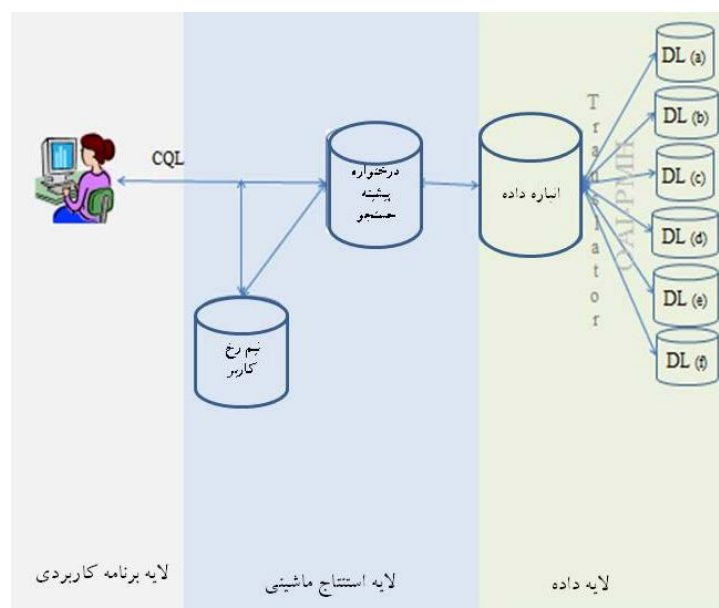
^۳ Data Warehouse

^۴ Semantic Store

فرا داده‌های کتابخانه‌های دیجیتالی عضو شبکه یکپارچه کتابخانه‌ای و روابط معنایی محتواهای فرا داده‌ها برای کاربر بازیابی و نمایش داده می‌شود.

از جمله پروژه‌هایی که بر این مبنا استوار شده‌اند می‌توان به پروژه دلاس (DELOS, 2004)، پروژه تی.اس.آی.ام.آی.اس. [Chawathe, et al., 1996] و پروژه یورویانا (Introduction, 2011) اشاره کرد. پژوهش صنعت جو و فتحیان در زمینه مقایسه کارآمدی اصطلاحنامه و هستی‌شناسی در بازنمون دانش نیز در نهایت به بیان استفاده از هستی‌نگاری‌ها در بازنمون دانش برای استنتاج داده‌ها رسیده است. از این رو روش پیشنهادی مشابه هم می‌تواند باشد. همچنین هوانگ، کِ و یانگ (Huang, Ke, & Yang, 2005)، آیساک، شلوباچ، متزینگ و زین (Issac, Schlobach, Mattheizing, & Zinn, 2008) و پراساد و مدالی (Pasad & Madalli, 2008) نیز به استفاده از روش مشابهی رهنمون شده‌اند.

در روش دوم که معماری آن در شکل ۵-۲ نشان داده شده است، برقراری روابط معنایی برعکس روش پیشین مبتنی بر معماری پس‌همارایی است. تاکید اصلی در این روش بر جستجوی کاربر است. لایه مربوط به یکپارچه‌سازی نحوی همانند روش پیشین مبتنی بر معماری باز و مدل برداشت اطلاعات است. در لایه استنتاج پایگاه اطلاعاتی مبتنی بر ثبت پیشینه جستجوی کاربر وجود دارد و روابط معنایی میان کلیدواژه‌های جستجو و منابع بازیابی شده در آن برقرار و ثبت می‌شود. هنگام جستجوی کاربر در این سیستم جستجوی او با روابط ثبت شده در پایگاه اطلاعاتی مذکور برقرار شده و بر اساس وزن حاصل از حجم رابطه بین کلیدواژه‌ها و فرا داده‌ها یافته‌های معنایی مرتبط بازیابی شده و به کاربر نمایش داده می‌شود.



شکل ۵-۲. معماری مبتنی بر ثبت پیشینه جستجو

پروژه ان.اس.دی.ال. (Lagoze, et al., 2002) از جمله پروژه‌هایی است که مبتنی بر پروفایل کاربر امکان دستیابی معنایی به اطلاعات را فراهم می‌کند. در این زمینه پروژه‌هایی در فضای عام وب برای دستیابی به اطلاعات اجرا شده است ولی با توجه به اینکه آنها در فضای کتابخانه‌های دیجیتال نبودند در اینجا به این نمونه‌ها اشاره‌ای نمی‌شود.

۳. از چه ابزارهایی در یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال استفاده می‌شود؟

یافته‌های مربوط به این پرسش نیز در فصل ۲ و با تحلیل متون مرتبط حاصل شده و بیان شده‌اند. در این زمینه بیان شد که پیش‌نیازهای زیادی از جمله زیر ساخت فناوانه (مانند ساختارها، تفاهم‌نامه‌ها، روش‌های ترکیبی، طرح‌های کدگذاری، و نظام‌های شناساگر)، استانداردها، مسائل سازمانی و قانونی، خدمات حمایتی مانند ثبت‌های مختلف و پایگاه‌های دانشی معنایی و سیستم سازماندهی دانش (از جمله هستی‌نگاری‌های بنیادی و هسته‌ای) (Patel, Koch, Doerr, & Tsinaraki, 2005)، برای یکپارچه‌سازی معنایی و محقق ساختن یکپارچه‌سازی معنایی مورد نیاز هستند. از جمله ابزارهای مهم در ایجاد یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتالی می‌توان به موارد زیر اشاره کرد:

- هستی‌نگاری‌ها. هستی‌نگاری‌ها برای برقراری روابط معنایی مورد استفاده قرار می‌گیرند. این ابزارها در اغلب موارد از جمله ابزارهای پیش‌همارا در نظام‌های سازماندهی دانش هستند. هستی‌نگاری‌ها بر اساس دانش موجود و روابط بین آنها ایجاد می‌شوند. هستی‌نگاری‌ها یا از طریق تبدیل اصطلاحنامه‌ها و یا به تدریج و با رشد دانش در نظام‌های اطلاعاتی شکل می‌یابند. از جمله هستی‌نگاری‌های موجود می‌توان به وردنت اشاره کرد.
- سیستم‌های سازماندهی دانش. همانگونه که پیشتر ذکر شده، سیستم‌های سازماندهی دانش به واژگان کنترل شده و نیز سیستم‌ها، ابزارها یا خدماتی اشاره دارد که به منظور سازماندهی دانش ایجاد شده‌اند (ابزارهایی که تفسیر سازماندهی شده‌ای از ساختار دانش ارائه می‌نمایند) (Patel, Koch, Doerr, & Tsinaraki, 2005). از جمله این سیستم‌ها می‌توان به اسکاس اشاره کرد.
- پایگاه دانشی. این پایگاه شامل واژگان و روابط معنایی بین آنها است. در حقیقت این پایگاه در برگیرنده هر دو ابزار ذکر شده در بالا است. این پایگاه همواره در نظام‌های یکپارچه اطلاعاتی در لایه استنتاج قرار می‌گیرند و مغز متفکر آن را تشکیل می‌دهد. این پایگاه مبتنی بر نظام‌های یادگیرنده عمل می‌کند و هوشمندی شبکه معنایی کتابخانه دیجیتالی نیز بر اساس آن حاصل می‌شود.
- مترجم‌ها. مترجم‌ها از جمله ابزارهایی هستند که وظیفه یکدست‌سازی فراداده‌ها و تبدیل آنها به یک استاندارد مشخص و مورد نیاز در نظام بر عهده دارند. تمامی مترجم‌ها مبتنی بر نقشه‌هایی است که از قبل برای استاندارد‌ها و یا حتی فراداده‌های غیر استاندارد تهیه و در سیستم تعبیه می‌شوند. نوع فعالیت این مترجم‌ها بر اساس استانداردسازی منفعل (Shen R., 2006) می‌باشد. همانگونه که ذکر شد مترجم‌ها فراداده‌ها را دریافت و آنها را بر اساس نقشه‌های تبدیلی که دارند به استاندارد مورد استفاده در انبار داده تبدیل و منتقل می‌کنند.
- زبان‌های توصیفی و تحلیلی. برای انتقال داده‌ها در شبکه‌های یکپارچه نیاز به استفاده از زبان توصیفی مانند ایکس.ام.ال. است. بدون آن امکان انتقال داده‌ها امروزه امکان پذیر نیست. بنابراین می‌توان اظهار داشت که زبان توصیفی ایکس.ام.ال. جزء لازم و ضروری انتقال داده‌ها است. علاوه بر این برای ایجاد شبکه‌های معنایی و هستی‌نگاری‌ها نیاز به استفاده از زبان‌های تحلیلی مانند آر.دی.اف.، دمل+اوایل و یا او.دبلیو.ال. است. این زبان‌ها قابلیت برقراری روابط معنایی را در سیستم‌های دانشی بر عهده دارند.

۴. وضعیت موجود محتواهای پیشینه‌های کتابشناختی کتابخانه‌های دیجیتال مورد بررسی چگونه است؟

جهت پاسخ به این پرسش از پرسشنامه محقق ساخته جهت گردآوری اطلاعات استفاده شد. بر این اساس یافته‌های مربوط به این پرسش در فصل ۴ متن حاضر توصیف شده‌اند. در اینجا پس از ارائه خلاصه‌ای از یافته‌های مذکور اقدام به تحلیل و پاسخ‌گویی به این پرسش شده است. به طور خلاصه یافته‌ها در این رابطه به شرح زیر هستند:

- یافته‌ها در این زمینه حاکی از تاکید ویژه کتابخانه‌های مورد بررسی بر منابع متنی (کتاب، پایان‌نامه، و مقالات) است.
 - برای توصیف منابع اطلاعاتی از فراداده‌هایی که شامل فیلدهای اطلاعاتی متعددی هستند بهره‌برداری می‌شود و از بین فیلدهای اطلاعاتی مورد استفاده محتوای فیلدهای عنوان، پدیدآور، پدیدآور همکار، موضوع، ناشر، و شابک (برای کتاب-ها) بیشتر تکمیل می‌گردند.
 - در زمینه مستندسازی اسامی پدیدآورندگان، یافته‌ها حاکی از این است که از فهرست مستند اسامی مشاهیر و مولفان و پایگاه اطلاعاتی کتابخانه ملی و برای مستند سازی نام ناشران بیشتر از قسمت شناسنامه اثر در دست بررسی استفاده می‌شود.
 - در تحلیل موضوعی منابع اطلاعاتی بیشتر از موضوع‌دهی کنترل شده با استفاده از سرعنوان‌های موضوعی و نمایه‌سازی کنترل شده با استفاده از اصطلاحنامه‌ها بهره‌برداری می‌شود. همچنین موضوع‌دهی در بیشتر موارد در مورد کتاب‌ها انجام می‌گیرد.
 - در توصیف موضوعی با استفاده از کلیدواژه‌ها، پایان‌نامه‌ها، کتاب‌ها و مقالات بیشترین بهره را برده‌اند.
 - در بیشتر موارد از فهرست‌نویسی به شیوه نسخه‌برداری استفاده می‌شود. جهت نسخه‌برداری نیز کتابشناسی ملی بیشترین کاربرد را دارد. برای منابع انگلیسی زبان نیز از پایگاه اطلاعاتی کتابخانه کنگره امریکا استفاده می‌شود. در اکثر موارد نیز نسبت به اصلاح محتواهای بازیابی شده از پایگاه‌های اطلاعاتی مذکور اقدام می‌شود.
 - در فهرست‌نویسی بنیادی نیز از سرعنوان‌های موضوعی کنگره امریکا، فهرست‌نویسی پیش از انتشار و سرعنوان‌های موضوعی فارسی استفاده می‌شود.
 - در نمایه‌سازی منابع اطلاعاتی نیز از اصطلاحنامه‌های موضوعی استفاده می‌شود. در این زمینه اکثر اصطلاحنامه‌های موضوعی موجود مورد استفاده قرار می‌گیرند. اصطلاحنامه سه زبانه اصفا نیز از جمله اصطلاحنامه‌های عمومی پر استفاده است.
- با توجه به یافته‌های ذکر شده در مورد وضعیت محتواهای پیشینه‌های کتابشناختی کتابخانه‌های دیجیتال مورد بررسی باید گفت که با توجه به اینکه در حال حاضر اکثر منابع موجود منابع متنی هستند تدابیر لازم جهت دستیابی یکپارچه به منابع اطلاعاتی در وضعیت موجود می‌تواند مبتنی بر این منابع باشد ولی باید امکان یکپارچه‌سازی برای سایر منابع اطلاعاتی را نیز در آینده نزدیک در نظر داشت. این یافته همسو با یافته پژوهش علیپور حافظی (۱۳۸۸) و علیپور حافظی، حری، اسفندیاری مقدم و علیپور حافظی (۱۳۹۰) که به منظور شناسایی وضعیت سازماندهی منابع اطلاعاتی کتابخانه‌های دیجیتال انجام گرفته می‌باشد. همچنین متأسفانه اغلب فیلدهای اطلاعاتی مربوط به منابع اطلاعاتی در کتابخانه‌های دیجیتال تکمیل نمی‌شوند و تنها محدود فیلدهای ضروری و اصلی برای توصیف و تحلیل منابع اطلاعاتی حاوی محتواهای اطلاعاتی هستند. این یافته نیز همسو با یافته علیپور حافظی، حری، اسفندیاری مقدم و علیپور حافظی (۱۳۹۰) است. در زمینه مستندسازی نیز که در یکپارچه‌سازی محتواهای اطلاعاتی بسیار مهم است برای نام پدیدآورندگان در برخی از موارد از فهرست مستند اسامی مشاهیر و مولفان استفاده می‌شود. در مقابل در برخی موارد دیگر نیز مستندسازی در زمینه اسامی افراد صورت نمی‌گیرد. در مورد مستندسازی اسامی ناشران وضعیت بدتر نیز می‌باشد و در اکثر موارد نسبت به مستندسازی اسامی ناشران اقدام نمی‌شود. همین مساله باعث ایجاد مشکلاتی در یکپارچه‌سازی معنایی و محتوایی منابع پراکنده در کتابخانه‌های دیجیتال می‌شود.
- در زمینه سازماندهی نیز با توجه به اینکه از دو شیوه بنیادی و نسخه‌برداری به طور هم‌زمان استفاده می‌شود، وضعیت موجود تا حدودی مناسب است. زیرا سازماندهی کتاب‌های دیجیتال از روش موضوع‌دهی و با استفاده از سرعنوان‌های موضوعی فارسی و کنگره امریکا انجام می‌گیرد و سایر منابع اطلاعاتی با استفاده از اصطلاحنامه‌های متعدد موجود نمایه‌سازی می‌شوند. در صورتی که سازماندهی به صورت نسخه‌برداری نیز باشد در این شرایط نیز کتابشناسی ملی و کتابخانه کنگره امریکا که هر دو از سرعنوان‌های موضوعی استفاده می‌کنند، باز شرایط تا حدودی می‌تواند مطلوب باشد. استفاده از منابع مرجع مذکور مزیتی بر یکپارچه‌سازی پیشینه‌های کتابشناختی محسوب می‌شود. در چنین شرایطی بهره‌گیری از منابع مرجع مذکور می‌تواند تاثیر گذاری مناسبی در برقراری تعامل معنایی میان منابع اطلاعاتی داشته باشد. با وجود این هنوز ابزارهای استاندارد تعامل معنایی در محیط دیجیتال مانند آنتولوژی‌های موضوعی و زبان‌های توصیفی مانند او.دبلیو.ال. یا آر.دی.اف. توسط کتابخانه-

های دیجیتال مورد بررسی مورد بهره‌برداری قرار نگرفته‌اند. بر این اساس همین مساله می‌تواند محدودیتی در راستای ایجاد یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال مورد بررسی اعمال کند.

مساله دیگر توجه محدود به منابع چندرسانه‌ای در ارائه خدمات کتابخانه‌های دیجیتال مورد بررسی است. این منابع روز به روز جایگاه بهینه‌ای در ارائه خدمات اطلاعاتی به کاربران می‌یابند. همین مساله توجه جدی کتابخانه‌های دیجیتال به منابع غیر متنی را می‌طلبد. بر این اساس خواه ناخواه کتابخانه‌های دیجیتال باید به این سمت حرکت کنند و قابلیت‌های مورد نیاز را باید ایفا نمایند. بنا براین در ارائه راهکار این پژوهش باید به این مهم توجه جدی نمود و الزامات اولیه مورد نیاز را برای بهره‌گیری از این منابع در یکپارچه‌سازی معنایی میان منابع اطلاعاتی مذکور فراهم نمود.

۵. وضعیت موجود خروجی‌های فراداده نرم‌افزارهای کتابخانه دیجیتال مورد بررسی چگونه است؟

جهت پاسخ به پرسش اساسی فوق نه پرسش در سیاهه واری برای بررسی جنبه‌های مختلف آن در نظر گرفته شد. در این زمینه ۳ متغیر اصلی مد نظر قرار گرفت که عبارت بودند از: خروجی‌های فراداده‌ای، ارتباط فراداده با شیئی دیجیتال، و قابلیت نمایه‌سازی توسط موتورهای جستجوی عمومی.

خروجی فراداده در سیستم‌های اطلاعاتی همواره با دو مفهوم متفاوت مرتبط است. نخست اینکه فراداده باید در قالب یک فرمت استاندارد ارائه شود. دوم اینکه فراداده مذکور در یک قالب استاندارد باید با استفاده از یک قالب یا محیط یا ظرفی ارائه شود. به بیانی دیگر فراداده به عنوان مطرووف نیازمند ظرفی برای انتقال و یا نمایش است. به عنوان مثال فراداده می‌تواند در قالب یک فایل متنی مانند ایزو ۲۷۰۹ و یا یک فایل سلسله مراتبی مانند ایکس.ام.ال. ارائه شود. استفاده از فرمت مُدز یا دابلین‌گُر به عنوان استاندارد فراداده و ایکس.ام.ال. به عنوان بستر سلسله مراتبی جهت انتقال محتوای فراداده در بسیاری از پروژه‌های مبادله اطلاعات و یکپارچه‌سازی اطلاعات در سطح کشورهای توسعه یافته مورد استفاده قرار می‌گیرد (Arms, Dushay, Fulker, & Lagoze, 2003; Hagedorn, 2003; Jayakanth, Maly, Zubair, & Aswath, 2006; Nelson, Rocker, & Harrison, 2003; Jin, 2004; Lam & Chan, 2007). بررسی حاضر در نرم‌افزارهای کتابخانه دیجیتال در ایران نشان داد که این نرم‌افزارها از فرمت مارک و دابلین‌گُر هر کدام با حدود ۵۰ درصد، و از محیط سلسله مراتبی ایکس.ام.ال. با حدود ۷۰ درصد تبعیت می‌کنند. بنابراین، یافته‌های پژوهش حاضر در این زمینه با یافته‌ها در همین مورد در سیستم‌های اطلاعاتی سایر کشورها تقریباً در یک راستا قرار دارند. همچنین می‌توان اظهار داشت که نرم‌افزارهای مورد مطالعه از نظر فنی تقریباً دارای قابلیت ارائه خروجی مورد نیاز برای یکپارچه‌سازی اطلاعات هستند.

خروجی‌های فراداده در قالب‌های ذکر شده باید ارتباط مناسبی به صورت درونه‌ای با شیئی دیجیتال داشته باشند. در این صورت است که یافته کاربر در سطح فراداده می‌تواند او را به شیئی دیجیتال رهنمون شود. یافته‌های پژوهش حاضر نیز در این راستا نشان می‌دهد که استفاده از قابلیت ارتباط درونه‌ای و فرامتن هر کدام با حدود ۴۶ درصد بیشترین استفاده را در این زمینه دارند.

فراداده‌های ذخیره شده و نیز ارائه شده در سیستم‌های اطلاعاتی برای اینکه بتوانند در دسترس بیشتر کاربران قرار گیرند، باید قابلیت نمایه شدن توسط موتورهای جستجوی عمومی را داشته باشند. همچنین این قابلیت به سیستم‌های اطلاعاتی اجازه می‌دهد تا به راحتی زمینه لازم را برای یکپارچه‌سازی و مبادله اطلاعات داشته باشند. یافته‌های پژوهش حاضر در این زمینه نشان داد که در

شرایط عادی این قابلیت در حدود ۳۰ درصد نرم‌افزارهای مورد مطالعه وجود دارد و نیز در ۳۰ درصد دیگر نرم‌افزارها این قابلیت می‌تواند به راحتی فراهم شود و تنها در حدود ۳۰ درصد نرم‌افزارها این قابلیت در دسترس نمی‌باشد. بنابراین می‌توان اظهار داشت که این قابلیت در اکثر نرم‌افزارهای مورد مطالعه وجود دارد.

به طور کلی در پاسخ به پرسش اساسی ۵ این پژوهش یافته‌ها نشان دادند که قابلیت ارائه خروجی‌های مناسب، ارتباط خروجی‌ها با شیء دیجیتالی و نیز قابلیت نمایه‌سازی توسط سایر سیستم‌های اطلاعاتی در سیستم‌های اطلاعاتی مورد بررسی در حد نسبتاً مطلوبی قرار دارد.

۶. الگوی مناسب برای یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال ایران چیست؟

با توجه به یافته‌های پژوهش حاضر در پرسش‌های اساسی ۱ تا ۳ متوجه می‌شویم که عناصر، ابزارها و روش‌های متعددی در یکپارچه‌سازی معنایی در سیستم‌های اطلاعاتی کتابخانه‌های دیجیتالی مورد استفاده قرار می‌گیرند. به طور کلی تمامی این عناصر در قالب الگوهای متعددی می‌توانند مطرح شوند. روش‌های موجود در این زمینه در پاسخ به پرسش ۲ این پژوهش معرفی شدند. همچنین این روش‌ها به طور مبسوط در فصل ۲ پژوهش حاضر نیز معرفی شده‌اند. علاوه بر این در این پژوهش با تلفیق یافته‌های پژوهش‌ها و پروژه‌های اجرا شده در زمینه تعامل معنایی با یافته‌های مرتبط با وضعیت موجود کتابخانه‌های دیجیتالی ایران از منظر قابلیت‌های مبادله اطلاعات (نحوی و معنایی) الگویی متناسب با کتابخانه‌های دیجیتالی ایران در پاسخ به این پرسش طراحی شد.

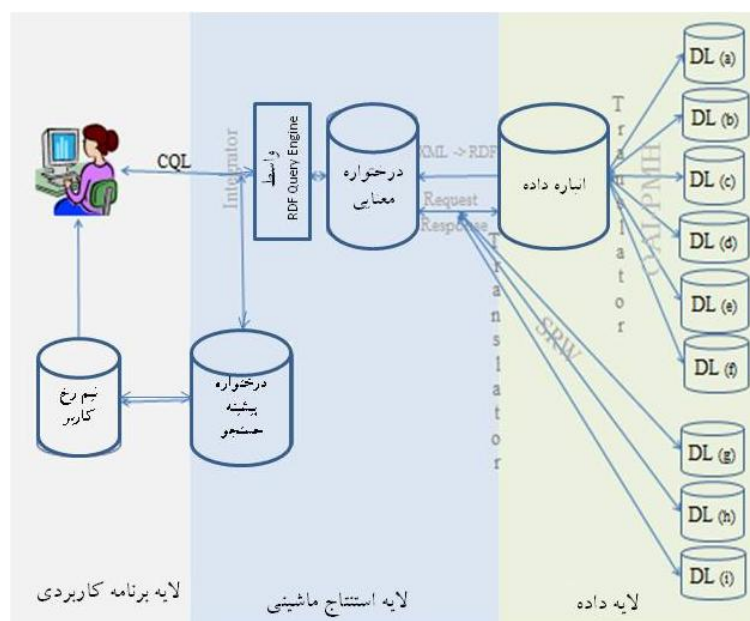
نتایج پژوهش حاضر نشان داد که وضعیت موجود کتابخانه‌های دیجیتالی به لحاظ تعامل معنایی چگونه است. همچنین روش‌ها و ابزارهای مورد نیاز با مرور و تحلیل متون استخراج و معرفی شد. در اینجا برای پاسخ به این پرسش، الگوی مناسب برای یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال ایران با تحلیل تمام موارد پیش گفته ارائه می‌شود.

الگوی پیشنهادی برای تعامل معنایی لاجرم نیازمند اشاره به تعامل نحوی نیز می‌باشد. بنابراین الگوی پیشنهادی باید پوشش دهنده معماری یکپارچه‌سازی به طور کلی باشد. علت اصلی طرح این مسأله این است که در پژوهشی که در این زمینه اجرا شده مشخص شده که معماری اتصال کتابخانه‌های دیجیتال به این شبکه باید معماری دو وجهی باشد (علیپور حافظی، ۱۳۸۸). یعنی مدل ایجاد شبکه یکپارچه شامل اتصال برخی از اعضا به شیوه جستجوی هم‌زمان و برخی دیگر به شیوه برداشت اطلاعات است. این به این معنی است که برخی از نرم‌افزارهای کتابخانه دیجیتال قادر به ارائه دسترسی به اطلاعات پایگاه اطلاعاتی خود به صورت باز به مرورگرهای شناخته شده هستند. برخی دیگر نیز این قابلیت را در اختیار قرار نمی‌دهند و تنها از طریق پروتکل Z39.50 یا اس.آر.یو. یا اس.آر.دبلیو. این اجازه را به سایر سیستم‌های اطلاعاتی مشابه می‌دهند تا به محتوای اطلاعاتی دسترسی پیدا کنند. بنابراین در چنین شرایطی نیاز به سیستمی دو وجهی است که هر دو شیوه دسترسی به اطلاعات را بتواند پوشش دهد.

همچنین سیستم‌های اطلاعاتی عضو شبکه یکپارچه، سیستم‌های اطلاعاتی نامتجانسی هستند که دارای معماری و ساختار متفاوتی از هم هستند. بنابراین مسأله اصلی دیگر در این زمینه ترجمه و تبدیل زبان‌ها و ساختار ارتباطی آنها به یکدیگر است. از طرفی نیز استانداردهای متعددی توسط سیستم‌های اطلاعاتی عضو در شبکه در سطح فراداده پشتیبانی می‌شوند. این در حالی است که برای

برقراری ارتباط نیاز به استفاده از زبانی مشترک برای تعامل وجود دارد. این مسأله زمانی اهمیت خود را دو چندان نشان می‌دهد که هدف ایجاد یکپارچه‌سازی معنایی در ارائه خدمات به کاربران باشد.

بر این اساس الگوی پیشنهادی پژوهش حاضر با توجه به ویژگی‌های مذکور و نتایج حاصل از این پژوهش به شکل ۳-۵ در زیر ارائه می‌شود:



۳-۵. الگوی پیشنهادی یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال ایران

الگوی پیشنهادی پژوهش حاضر از سه لایه اصلی داده، استنتاج ماشینی و کاربرد تشکیل شده است. هر یک از سه لایه مذکور وظایف متعددی را برعهده دارند. در لایه داده، فراداده‌های متعدد موجود در کتابخانه‌های دیجیتال نامتجانس و نرم‌افزارهای کتابخانه دیجیتال با معماری مختلف مد نظر قرار می‌گیرند. در این لایه فراداده‌های متعدد با استانداردهای فراداده مختلفی حضور دارند که همانگونه که در بالا نیز اشاره شد و در شکل ۳-۵ نیز نشان داده شده است، نیاز به مترجم‌هایی برای تبدیل این استانداردها و فراداده‌ها به یکدیگر وجود دارد. در پیوست ۳ نمونه‌ای از یک فراداده در کتابخانه دیجیتال مجلس شورای اسلامی که از نرم‌افزار آذرخش استفاده می‌کند ارائه شده است.

در شکل مشخص است که کتابخانه‌های دیجیتال عضو شبکه یکپارچه در دو دسته قرار گرفته‌اند. یک دسته کتابخانه‌هایی هستند که امکان دستیابی به محتوای فراداده خود را به مرورگرها می‌دهند. این نوع کتابخانه‌ها با مدل برداشت اطلاعات و با استفاده از پروتکل او.ای.آی. به سیستم واسطه تعبیه شده در لایه استنتاج متصل می‌شوند و فراداده خود را در اختیار سیستم واسطه مذکور قرار می‌دهند. در مسیر انتقال فراداده‌ها مترجمی قرار می‌گیرد که وظیفه اصلی آن ترجمه استانداردها و یا فراداده‌های مختلف به استاندارد مورد نیاز سیستم واسطه است. با توجه به تنوع استانداردها و فرمت‌های مورد استفاده توسط کتابخانه‌های دیجیتال عضو در شبکه (ذکر شده در جدول ۴-۱۰) نیاز به یکدست سازی آنها و استفاده از استاندارد فراداده ساده‌ای بود تا پاسخگوی نیازهای

تعامل معنایی مورد نیاز این شبکه باشد. با توجه به تنوع منابع اطلاعاتی و فراداده‌های مورد استفاده برای آنها و نیز با توجه به سادگی فراداده مورد نیاز در یکپارچه‌سازی معنایی در این سیستم پیش بینی شده که تمامی استانداردها و فرمت‌های فراداده به استاندارد دابلین‌گر ساده با ۱۴ فیلد زیر تبدیل شوند. شمای تبدیل تمامی استانداردها و فیلدهای ابرداده‌ای در مترجم مورد استفاده در این فرآیند تعبیه می‌شود. فیلدهای فراداده مورد نظر عبارتند از:

- عنوان
- پدیدآور
- موضوع
- توصیف
- ناشر
- همکار
- تاریخ
- نوع
- قالب
- شناسگر
- منبع
- زبان
- رابطه
- حقوق

همین مترجم در فرآیند تبدیل فراداده‌ها در یکپارچه‌سازی سیستم‌های اطلاعاتی توزیع یافته نیز مورد استفاده قرار می‌گیرد. در بخش پایین لایه داده الگوی پیشنهادی، کتابخانه‌های دیجیتال توزیع یافته‌ای قرار دارند که اجازه دسترسی به محتوای فراداده را به سیستم واسط نمی‌دهند. همانگونه که در الگوی پیشنهادی نیز به این موضوع پرداخته شده است نحوه تعامل با این سیستم‌ها با استفاده از مدل جستجوی هم زمان خواهد بود. در این مدل از پروتکل اس.آر.دبلیو. استفاده خواهد شد. همانگونه که پیشتر نیز ذکر شده این پروتکل نسل جدید پروتکل Z39.50 برای برقراری ارتباط هم زمان در کتابخانه‌های دیجیتال ارائه شده است. در فرآیند دستیابی به محتواهای این کتابخانه‌های دیجیتالی در قالب فراداده استاندارد مورد نیاز در این سیستم نیاز به مترجم برای تبدیل استانداردها و قالب‌های فراداده به استاندارد مورد نیاز در این سیستم است. دستور دریافت فراداده از لایه استنتاج و هم زمان با ارسال دستور دریافت اطلاعات کاربر به سیستم اطلاعاتی واسط ارسال می‌شود و خروجی و پاسخ سیستم‌های اطلاعاتی توزیع یافته نیز هم زمان با دریافت پاسخ از سیستم واسط توسط پایگاه معنایی دریافت می‌شود.

در لایه میانی، لایه استنتاج، دو پایگاه اطلاعاتی مربوط به پایگاه اطلاعاتی شبکه معنایی فراداده‌ها و پایگاه اطلاعاتی شبکه ارتباطی پیشینه‌های جستجوی کاربران قرار دارند. منطق فعالیت این پایگاه‌های اطلاعاتی برقراری و ایجاد شبکه‌های معنایی مورد نیاز برای بازیابی معنایی است.

در پایگاه اطلاعاتی معنایی ساختار موجود مبتنی بر روابط سرعنوان‌های موضوعی، اصطلاحنامه‌های مورد استفاده در نمایه سازی، و مستند اسامی مورد استفاده در مستندسازی اسامی خواهد بود. برای این منظور نیاز به تبدیل فراداده‌ها از فرمت ایکس.ام.ال. به آر.دی.اف. برای استفاده در شبکه معنایی مذکور وجود دارد. در این مرحله تنها فیلدهای ضروری مورد استفاده در شبکه معنایی وارد این پایگاه اطلاعاتی می‌شوند. فیلدهای مذکور عبارتند از: پدیدآور، موضوع، توصیف، همکار و شناسگر. همچنین همانگونه که پیشتر نیز ذکر شد پاسخ‌های دریافتی از کتابخانه‌های دیجیتالی توزیع یافته نیز پس از دریافت وارد این پایگاه اطلاعاتی می‌شوند. بنابراین پس از اختصاص یافتن به گره‌های شبکه معنایی موجود در پایگاه اطلاعاتی در اختیار کاربر قرار می‌گیرند.

پرس و جوی کاربر با زبان سی.کی.ال. دریافت می‌شود و در میان‌افزاری تبدیل به پرس و جوی آر.دی.اف. می‌شود. پایگاه اطلاعاتی مربوط به ثبت پیشینه جستجوی کاربران نیز مبتنی بر سیستم‌های یادگیرنده^۱ است. سیستم‌های یادگیرنده در سیستم‌های هوشمند برای افزایش دانش و تغییر آن، افزایش کارایی و تصحیح اتوماتیک خطا مورد استفاده قرار می‌گیرند. اخذ دانش یکی از مهم‌ترین کاربردهای سیستم یادگیرنده است که امروزه بسیار مورد توجه قرار می‌گیرد. به این معنی که عمل یادگیری اطلاعات پایه را از محیط استخراج کرده و برای تحلیل اطلاعات آینده از آن بهره می‌گیرند. در این سیستم‌ها:

۱. امکان یادگیری و کسب دانش معتبر از مشاهدات تحت تأثیر عوامل موجود در محیط فراهم می‌شود.
۲. قابلیت کشف مفاهیم و ارتباط‌های بین آنها که قبلاً پیش‌بینی نشده‌اند و یا غیر قابل تصور هستند، در سیستم‌های یادگیرنده به وجود می‌آید. به این ترتیب، سیستم‌های یادگیرنده انسان را در کشفیاتش کمک کرده و خود را نیز گسترش می‌دهند.

بنابراین سیستم یادگیرنده برای اخذ ورودی و انتقال دانش به کاربر به شکل یک پایگاه دانش با محیط ارتباط برقرار می‌کند. به همین منظور باید داده‌های ورودی و دانش خروجی به یک زبان مشترک فرموله شوند، این زبان «زبان توصیفی» نامیده می‌شود. گویایی زبان توصیف برای توفیق و کارایی سیستم‌های یادگیرنده نقش حیاتی دارد. قوانین شبکه‌های معنایی و قاب‌ها نمونه‌هایی از زبان توصیفی هستند. این زبان باید بتواند توصیف جامعی از دامنه مورد یادگیری را نمایش دهد چرا که نقش عمده‌ای در یادگیری دارد. خروجی سیستم که به این زبان فرموله شده است باید قابل فهم باشد.

در الگوی پیشنهادی این پژوهش در این مرحله اطلاعات ورودی به سیستم پروفایل دریافتی از کاربران و ثبت پیشینه جستجوی کاربران است. از مهم‌ترین اجزاء یادگیری که مورد توجه بسیاری از محققین در این شاخه قرار گرفته است استراتژی‌های یادگیری می‌باشد که شامل دریافت، پردازش و پرداخت اطلاعات می‌شود.

در هنگامی که کاربر در حال جستجوی منابع مورد نظر خود در این سیستم است، پس از دیدن نتایج حاصله، کلمات کلیدی جدیدی در ذهن او شکل می‌گیرد. لذا سلسله کلماتی که او به سیستم وارد می‌کند دارای نظامی منطقی می‌تواند باشد و ناشی از برقرار شدن ارتباطات معنایی بین آنها از سوی فرد جستجو کننده است. به این مفهوم که احتمالاً مجموعه این کلمات از نظر معنایی با یکدیگر ارتباطی دارند که جستجوگر بعد از جستجو در مورد هر یک از آنها و دیدن بازخورد سیستم، از بین

¹ Learning systems

کلمات کلیدی ثانویه‌ای که سیستم به او بازخورد می‌دهد کلمات خاص دیگری را انتخاب می‌کند و در مورد آنها جستجو می‌کند.

اما مسئله این است که ممکن است در این فرآیند، جستجوگر کلمات نامرتبیطی را هم که به نظر او مرتبط می‌رسند در جستجوی خود وارد کند. بدون این که این کلمات اخیر ارتباطی با موضوع اصلی داشته باشند. اگر بنا داشته باشیم که سیستمی یادگیرنده بسازیم که با استفاده از کلماتی که جستجوگران مختلف در یک بازه زمانی در مورد آنها جستجو می‌کنند یک بانک کلمات کلیدی در مورد موضوع خاصی تشکیل دهیم، باید به هر دو مورد فوق توجه کنیم. یعنی اولاً از ارتباط معنایی بین کلماتی که کاربران مختلف به صورت پیاپی وارد سیستم می‌کنند استفاده کرده و کلماتی که توسط کاربران مختلف به کرات و به توالی در موردشان جستجو می‌شود را «دارای ارتباط معنایی با یکدیگر» تلقی کنیم، و ثانیاً بتوانیم نامرتبیط بودن برخی کلمات که در این توالی‌ها دیده می‌شوند را تشخیص داده و آنها را وارد این رده از کلمات مرتبط ننماییم.

یادگیرنده بودن سیستم، دارا بودن ویژگی‌های سومی را نیز برای آن تکلیف می‌کند، و آن این است که انتظار می‌رود که با توجه به پویا بودن مجموعه‌های علمی، کلمه‌ای که ممکن است قبلاً در رده‌ای از کلمات مرتبط قرار داشته، در طول زمان دیگر جزو آن رده محسوب نگردد و به رده معنایی دیگری تعلق گیرد. سیستم باید این توانایی را داشته باشد که این تغییر وضعیت را شناسایی کرده و رده‌بندی را به صورتی متناسب با آن تغییر دهد.

با این تفاسیر، روش کار در این مرحله برای ایجاد روابط معنایی با استفاده از کلیدواژه‌ها و نتایج حاصل از جستجوی کاربران و با استفاده از پایگاه پروفایل کاربران به این صورت باشد:

۱. دوره‌ای خاص به عنوان دوره یادگیری تعیین می‌شود؛ مثلاً یک سال.
۲. در طی این یک سال دنباله کلمات کلیدی که توسط کاربران مختلف در موردشان به توالی جستجو شده است مورد بررسی قرار می‌گیرند.
۳. از روی این بررسی، گروه‌هایی که کاربران مختلف به اجماع قائل به مرتبط بودن آنها با یکدیگر هستند در یک خانواده قرار داده می‌شوند، و به عبارتی دارای ارتباط معنایی با یکدیگر در نظر گرفته می‌شوند. به این ترتیب اثر جستجوهای نامرتبیط که طی این دوره توسط کاربران محدودی به اشتباه مرتبط با کلمه اصلی قلمداد شده‌اند خنثی شده، و این کلمه اخیر در خانواده قرار داده نمی‌شود.
۴. وقتی کاربری در مورد کلمه‌ای خاص به جستجو می‌پردازد، با توجه به دسته خانواده‌های ساخته شده در فوق، پیشنهادهایی در خصوص کلمات مرتبط دیگری که این کاربر می‌تواند در مورد آنها نیز به جستجو بپردازد به او ارائه می‌شود.
۵. وقتی دوره بکارگیری سیستم از یک سال بیشتر شد، هر روز اطلاعات جمع آوری شده از ۳۶۶ روز به قبل کنارگذاشته می‌شود و اطلاعات جمع آوری شده از روز آخر وارد فرآیند خانواده بندی مندرج در بند ۳ فوق می‌گردد. به این ترتیب پویایی سیستم در شبکه سازی همواره حفظ می‌شود.

به تدریج و با گذشت یک سال این پایگاه در اختیار قرار می‌گیرد و نتایج آن در نتایج بازیابی حاصل از شبکه معنایی تلفیق می‌شود و باعث ارتقاء نتایج معنایی در بازیابی اطلاعات از این سیستم می‌شود.

در لایه کاربردی که در سمت چپ الگوی پیشنهادی ذکر شده در شکل ۵-۳ نشان داده شده، ارتباط کاربران با پرتال جامع شبکه یکپارچه برقرار می‌شود. پرتال مذکور باید قابلیت شناسایی کاربران عضو در تمامی کتابخانه‌های دیجیتال عضو را داشته باشد. برای

این منظور پایگاه اطلاعاتی کاربران، اطلاعات خود را از پایگاه اطلاعاتی کاربران کتابخانه‌های دیجیتال عضو دریافت و به روزرسانی می‌کند. با مراجعه کاربر به این پرتال به تدریج شبکه معنایی پیشینه جستجوی کاربران شکل می‌گیرد و با گذشت یک سال از شروع می‌تواند نتایج خود را در بازایی معنایی به کاربران نشان دهد.

۳-۵. اعتبارسنجی مدل پیشنهادی

ارائه مدل بدون آزمون و اعتبارسنجی آن نمی‌تواند کارآمدی خود را به اثبات برساند. برای این منظور لازم است تا ارائه هر مدل^۱ الگو یا راهکاری پیش از انتشار از نظر اعتبار و کارآمدی مورد آزمون قرار گیرد. لازم به ذکر است که هر الگو^۲ مدل یا راهکاری برای حل یا تسهیل مسأله‌ای ارائه می‌شود. مهم‌ترین عنصر در آزمون مدل^۳ راهکار یا الگویی توجه به مسأله و هدف از ارایه آن است. برای آزمون مدل‌ها هفت روش تا کنون مطرح شده است. روش‌های مذکور عبارتند از: ارائه^۴ آزمون تجربی^۵ شبیه‌سازی^۶ استفاده از معیارها^۷ معیار سازی^۸ استدلال منطقی^۹ اثبات ریاضی^{۱۰} (Vaishnavi & Kuechler Jr., 2008).

با توجه به مسأله مورد نظر و راهکار ارائه شده لازم است تا یکی از روش‌های مذکور انتخاب و برای سنجش اعتبار راهکار ارائه شده مورد استفاده قرار گیرد. برخی از راهکارها مانند اثبات ریاضی و آزمون تجربی دارای اعتبار و ارزش بالایی هستند ولی هزینه و زمان بالایی نیاز دارند تا بتوانند مورد استفاده قرار گیرند. از طرفی نیز لازم است تا راهکار پیشنهادی در عمل قابلیت ارایه داشته باشند. برخی دیگر مانند شبیه‌سازی و معیار سازی می‌توانند بسیار پرکاربرد باشند. علت اصلی آن می‌تواند آزمون پیش از اجرای راهکار پیشنهادی باشد.

برای اعتبارسنجی مدل پیشنهادی این پژوهش از روش ارائه استفاده شد. در این روش لازم است تا ابتدا مدل پیشنهادی تهیه شود. این مدل نشانگر این است که در راستای نیاز طرح شده در ابتدای پژوهش طراحی شده و پاسخگوی آن می‌باشد. در ادامه باید نشان داد که این مدل می‌تواند مساله طرح شده را مرتفع نماید. به این منظور ابتدا در پژوهش حاضر با تحلیل محتواها و شناسایی وضعیت موجود کتابخانه‌های دیجیتالی ایران از منظر تعامل معنایی اقدام به طراحی مدل مفهومی پیشنهادی شد. سپس در گام بعدی اقدام به مطابقت آن با مساله طرح شده در ابتدای پژوهش شد. در این مرحله از نظر متخصصان حوزه علم اطلاعات و دانش شناسی آشنا به حوزه فناوری اطلاعات و متخصصان علوم کامپیوتر آشنا به تعامل سیستم‌های اطلاعاتی و وب معنایی استفاده شد. پس از جلسات متعدد با متخصصان مذکور در نهایت به قالب نهایی مدل پیشنهادی دست یافتیم که متناسب با شرایط موجود کتابخانه‌های دیجیتالی ایران قادر به ایجاد تعامل معنایی بوده و امکان توسعه متناسب با شرایط موجود را دارا است. در نهایت بر همین مبنا مدل پیشنهادی این پژوهش به لحاظ اعتبار مورد سنجش قرار گرفته و کارآمدی آن برای حل مساله این پژوهش مورد تایید قرار گرفت.

¹ Demonstration

² Experimentation

³ Simulation

⁴ Using metrics

⁵ Benchmarking

⁶ Logical reasoning

⁷ Mathematical proofs

۴-۵. پیشنهادهای پژوهش

با توجه به نتایج پژوهش و در ادامه پاسخ به پرسش‌های اساسی پژوهش حاضر در ادامه برخی پیشنهادهاى اجرایی جهت بهبود وضعیت کتابخانه‌های دیجیتالی به لحاظ برقراری روابط معنایی در یکپارچه‌سازی کتابخانه‌های دیجیتالی ایران ارائه می‌شوند:

- نتایج پژوهش حاضر در زمینه شناسایی عناصر دخیل در یکپارچه‌سازی معنایی کتابخانه‌های دیجیتالی نشان داد که برای این منظور نیاز به پروتکل‌ها، نظام‌های کدگذاری و نظام‌های شناساگر، استانداردهای فراداده، مسائل سازمانی و قانونی و پایگاه‌های دانشی معنایی مانند هستی‌نگاری‌ها^۱ و سیستم‌های سازماندهی دانش مانند اسکاس است. در این زمینه پیشنهاد می‌شود برنامه‌هایی برای پیاده‌سازی و ایجاد این عناصر توسط شرکت‌های نرم‌افزاری و سازمان‌ها و متولیان و تصمیم‌سازان حوزه مدیریت کتابخانه‌های دیجیتالی اجرا شود. در این زمینه نیاز به ایجاد قابلیت‌های برای تعامل مانند پروتکل‌های یکپارچه‌سازی کتابخانه دیجیتال و تبعیت از استانداردهای فراداده و تولید پایگاه‌های دانشی است.
- نتایج این پژوهش در زمینه روش‌های مورد استفاده برای یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتالی نشان داد که در این زمینه نیاز به استفاده از مترجم‌ها و ماشین‌های استنتاج است. در این زمینه پیشنهاد می‌شود تا شمای تبدیل استانداردهای فراداده به یکدیگر شناسایی و نرم‌افزارهایی برای این منظور تولید شود. در این راستا نرم‌افزارهای منبع بازی مانند مارک ادیت وجود دارند که از این قابلیت برخوردار هستند. شناسایی و شخصی‌سازی آنها می‌تواند سریع‌ترین و مطمئن‌ترین راه باشد.
- نتایج پژوهش در زمینه ابزارهای یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتالی نشان داد که در این زمینه نیاز به پایگاه‌های دانشی، هستی‌نگاری‌ها، سیستم‌های سازماندهی دانش و مترجم‌ها است. در این راستا پیشنهاد می‌شود اقدام‌های اولیه جهت ایجاد پایگاه‌های دانشی با توجه به ساختارهای هستی‌نگارانه مد نظر قرار گیرند. ایجاد هستی‌نگاری‌های موضوعی و تبدیل اصطلاحنامه‌ها به هستی‌نگاری‌ها می‌تواند در اولویت قرار گیرد.
- نتایج پژوهش در زمینه وضعیت موجود محتواهای پیشینه‌های کتابشناختی کتابخانه‌های دیجیتال مورد بررسی نشان داد که تاکید بر منابع متنی در کتابخانه‌های دیجیتال مورد بررسی وجود دارد. در این زمینه توجه به سایر منابع بویژه منابع چندرسانه‌ای و غیر متنی با توجه به تنوع اطلاعات موجود در دسترس می‌تواند مهم باشد. علاوه بر این توصیف داده‌ها در فراداده‌های تولیدی کتابخانه‌های دیجیتالی با مشکلات زیادی همراه است و در اکثر مواقع بسیاری از فیلدهای ضروری تکمیل نمی‌شوند. لذا پیشنهاد می‌شود در تکمیل فیلدهای اطلاعاتی توسط مسئولان کتابخانه‌ها اقدامات مقتضی به عمل آید.
- نتایج پژوهش در زمینه وضعیت موجود خروجی‌های فراداده نشان داد که اکثر نرم‌افزارهای مورد بررسی محدودیت‌هایی را در ایجاد دسترسی به محتواهای فراداده اعمال می‌نمایند. همچنین دسترسی به آنها را از طریق پروتکل‌های استاندارد

^۱ با توجه به اینکه در این پژوهش لازم بود تا الگویی برای یکپارچه‌سازی معنایی تهیه شود و یکی از ابزارهای مهم مورد نیاز هستی‌نگاری‌های مناسب برای این منظور بود لازم است تا پژوهش‌هایی در زمینه ایجاد هستی‌نگاری‌ها و کاربرد آنها انجام پذیرد. از آنجایی که نیاز به پژوهش مجزایی در این زمینه است و این پژوهش فقط نیاز به استفاده از هستی‌نگاری‌ها در بخشی از مدل پیشنهادی دارد^۲ به شیوه‌ها و نحوه ایجاد آنها نپرداخته و فقط به جایگاه و ضرورت و کاربرد آنها پرداخته است. توضیحات تکمیلی در این زمینه در فصل ۲ این گزارش نیز ذکر شده است.

می‌سازند. لذا پیشنهاد می‌شود شرکت‌های تولید کننده نرم‌افزارها امکان دستیابی به فراداده‌ها را به صورت باز برای نرم‌افزارهای دیگر مهیا سازند. در غیر اینصورت لازم است امکان ارائه محتوای فراداده را از طریق دریافت پرس و جوهای استاندارد فراهم نمایند. در مقابل مسئولان کتابخانه‌ها نیز باید به این امر اهتمام ورزیده و امکان دسترسی به محتواها را فراهم نمایند.

- نتایج حاصل از این پژوهش در زمینه الگوی مناسب برای یکپارچه‌سازی نیز نشان داد که با توجه به اینکه این الگو برگرفته از الگوهای موجود و شخصی سازی آن با توانمندی‌ها و ویژگی‌های کشورمان است، پیشنهاد می‌شود این الگو به اجرا درآمده و مورد استفاده قرار گیرد. در این زمینه سازمان‌های متولی مانند پژوهشگاه علوم و فناوری اطلاعات ایران می‌توانند اقدامات اولیه را به اجرا درآورند. این مسئله به ویژه می‌تواند در پروژه سهام (سامانه هوشمند اشتراک منابع) در پژوهشگاه علوم و فناوری اطلاعات ایران به اجرا درآید.

فهرست منابع

- اشباکر، چ. (۱۳۸۰). آموزش برنامه نویسی XML در ۲۴ ساعت. (ع. زارع پور، مترجم) تهران: مؤسسه علمی و فرهنگی نص.
- پائو، م. (۱۳۷۹). مفاهیم بازیابی اطلاعات. (ا. آزاد، & ر. فتاحی، مترجم) مشهد: دانشگاه فردوسی.
- پژوهشگاه علوم و فناوری اطلاعات ایران. (۱۳۸۹). گزارش عملکرد پنج دوره اجرایی طرح امین - امانت بین کتابخانه ها. بازیابی در آبان ۱۵، ۱۳۹۰، از پژوهشگاه علوم و فناوری اطلاعات ایران:
- <http://www.irandoc.ac.ir/files/sharing/amin-report.pdf>
- پناهی، س. (۱۳۸۴). ارزیابی طرح غدیر از دیدگاه کاربران. علوم و فناوری اطلاعات، ۲۱(۱)، ۱-۲۰.
- حسن زاده، م. و & حسینی، م. (۱۳۹۱). بررسی نظری و عملی کاربرد نظام ساده سازماندهی دانش در زبان فارسی. اصطلاحنامه و کاربردهای آن در محیط الکترونیکی (ص. ۶۱-۷۵). قم: نشر کتابدار.
- زارع زاده، ف. (۱۳۸۳). ایکس.ام.ال. چیست و چه کاربردهایی در کتابخانه دارد؟. فصلنامه کتاب، ۱۵(۲)، ۳۷-۴۴.
- ساربان قلی، م. (۱۳۸۴). بررسی مشکلات جستجو و بازیابی اطلاعات به زبان فارسی از اینترنت با مطالعه موردی بر روی کاربران مرکز اینترنت دانشگاه آزاد اسلامی واحد شبستر. دانشگاه آزاد اسلامی، واحد تهران شمال. تهران: دانشگاه آزاد.
- شمس فرد، م. (۱۳۸۹). پروژه ها. بازیابی در ۱۲، ۲، ۱۳۹۰، از مؤسسه تحقیقات، ارتباطات و فناوری اطلاعات:
- <http://www.itrc.ac.ir/p-portal/viewproject.php>
- صنعت جو، ا. و & فتحیان، ا. (۱۳۹۰). مقایسه کارآمدی اصطلاحنامه و هستی شناسی در بازنمون دانش (طراحی و ساخت نمونه هستی شناسی اصفاء). پژوهشنامه کتابداری و اطلاع رسانی، ۱۱(۱)، ۲۱۹-۲۴۰.
- علیپور حافظی، ح. (۱۳۹۰). بررسی وضعیت سازماندهی منابع اطلاعاتی در کتابخانه های دیجیتال ایران و ارائه راهکار پیشنهادی. پایان نامه کارشناسی ارشد. همدان، همدان، ایران: دانشگاه آزاد اسلامی همدان.
- علیپور حافظی، م. (۱۳۹۰). کتابخانه های دیجیتال: مبادله اطلاعات. تهران: سمت.
- علیپور حافظی، م. (۱۳۸۸). بررسی نحوه مبادله اطلاعات بین سیستمهای اطلاعاتی کتابخانه های دیجیتال در ایران و ارائه الگوی پیشنهادی. دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران. تهران: دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران.
- علیپور حافظی، م. (۱۳۹۰). بررسی هزینه-سازگاری سیستمهای اطلاعاتی کتابخانه های دیجیتال در ایران و ارائه الگوی پیشنهادی. پژوهشگاه علوم و فناوری اطلاعات ایران. تهران: پژوهشگاه علوم و فناوری اطلاعات ایران.

- محقق، ن. (۱۳۷۴). بررسی وضعیت مبادله اطلاعات و منابع میان کتابخانه‌های دانشگاهی شهر تهران. تهران: دانشگاه تهران.
- مزینانی، ع. (۱۳۸۲). کتابخانه و کتابداری. تهران: سمت.
- مؤسسه علمی فرهنگی دارالحديث. (۱۳۸۷). تاریخچه. بازیابی در آبان ۸، ۱۳۹۰، از پایگاه اطلاع رسانی کتابخانه های ایران:
- <http://www.lib.ir/LibView/default.aspx>
- نشاط، ن. (۱۳۹۰). در جستجوی معنا. تهران: نشر کتابدار.
- یوسفی راد، ا. (۱۳۸۸). آر.دی.اف.: الگویی برای توصیف منابع در وب معنایی. فصلنامه کتاب (۷۹)، ۹-۲۲.
- Alipour-Hafezi, M., Horri, A., Shiri, A & ,Ghaebi, A. (2010). Interoperability models in digital libraries: an overview .*The Electronic Library*, 28 (3), 438-452.
- Arms, W. Y., Dushay, N., Fulker, D & ,Lagoze, C. (2003). A case study in metadata harvesting: the NSDL .*Library Hi Tech*, 21(2), 228-237.
- Arms, W. Y., Hillmann, D., Lagoze, C., Krafft, D., Marisa, R., Saylor, J . . . ,Sompel, H. V. (2002). January .(A Spectrum of Interoperability: the site for science prototype for the NSDL .*D-Lib Magazine*.
- Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern information retrieval*. New York: ACM Press.
- Baker, C. J & ,Cheung, K.-H., Baker, C. J.; Cheung, K.-H. (2007). *Semantic Web:revolutionizing knowledge discovery in the life Sciences*, xix .(New York: Springer).
- Baker, T. (2012). Libraries, Languages of description, and linked data: a dublin Core perspective. *Library Hi Tech.*, 30(1), 116-133.
- Bergamaschi, S., Castano, S., De Capitani di Vimercati, S., Montanari, S & ,Vincini, M. (1998). An intelligent Approach to information integration .N. Guarino, In *Formal ontology in information systems*. Amsterdam, Netherlands: IOS press, 253-268.
- Bingham, G. (1995) Document and Library Integration, Retrieved 11 December 2012: <http://cordis.europa.eu/libraries/en/projects/dali.html>
- Borgman, C. L. (2000). *From Gutenberg to the global information infrastructure: access to information in the networked world* .Cambridge, MA: MIT press.
- Boteram, F. (2010). Content architecture: Semantic interoperability in an international comprehensive knowledge organisation system .*Aslib Proceedings: New Information*, 62(4/5), 406-414.
- Bouquet, P., Stoermer, H., Barczyński, W & ,Bocconi, S. (2010). Entity-Centric Semantic Interoperability .Y. Kalfoglou, در *Cases on semantic interoperability for information systems integration : practices and applications*. New York: InformatIon science reference, 1-21.

- Brown, D., Byard, e & ,Hart, C. (2001). *The impact of XML on information content organization in libraries*. Retrieved 2 June, 2013:
www.tennessee.edu/digitallibrary/xml/howisxmlgoingaffectlibraries.htm
- Büttcher, S., Clarke, C. L & ,Cormack, G. V. (2010). *Information retrieval: implementing and evaluating search engines* .Massachuset: MIT press.
- Chawathe, S., Garcia-Molina, H., Hammer, J., Ireland, K., Papakonstantinou, Y., Ullman, J., & Widom, J. ([1996]). *The TSIMMIS Project: Integration of Heterogeneous Information Sources*. Retrieved July 10, 2013, from Stanford University:
db.ucsd.edu/publications/tsimmis-overview.ps
- Chen, H., Finin, T & ,Joshi, A. (2003). An Ontology for Context-Aware Pervasive Computing Environments .*Knowledge engineering review*, 18(3), 197-207.
- Cheung, e.-H., Smith, A. K., Yip, e. Y., Baker, C. J & ,Gerstein, M. B. (2007). Semantic Web approach to database integration in the life sciences .e.-H. Cheung & ,C. J. Baker, In *Semantic Web: Revolutionizing Knowledge Discovery in the Life Sciences*. New York: Springer, 11-30.
- Creating a European library space telematics for libraries programmes 1990-1998 (2000)*. Retrieved 16 September 2012. Available: <http://cordis.europa.eu/libraries/en/intro.html>
- DAML. (2006, January 13). Retrieved January 28, 2013, from DAML: <http://www.daml.org/>
- Deegan, M & ,Tanner, S. (2000). *Digital futures: strategies for the information age* .London: library association publishing.
- DELOS (2004). *Welcome to the DELOS network of excellence*. Retrieved November 18, 2011:
http://www.delos.info/index.php?option=com_frontpage&Itemid=1
- Desmarais, N. (2000). *The ABC's of xml* .Houston: Houston Technology press.
- Ding, H & ,Sølvberg, I. T. (2004). TOWARDS THE SCHEMA HETEROGENEITY IN DISTRIBUTED DIGITAL LIBRARIES .*ICEIS 2004 - HUMAN-COMPUTER INTERACTION* .Porto: Universidade Portucalense and INSTICC, 307-312.
- Favaro, J. (1998). Access to Remote Catalogues by implementing SR target functions. Retrieved Dec. 18, 2011: <http://cordis.europa.eu/libraries/en/projects/arca.html>
- Fox, E. A., Suleman, H., Gaur, R. C., & Devika, M. P. (2005). Design Architecture: An Introduction and Overview. In Y.-L. Theng, & S. Foo, *Design and usability of digital libraries : case studies in the Asia-Pacific* (pp. 22-38). London: Information Science Publishing.
- Fulker, D & ,Janée, G. (2002). Components of an NSDL Architecture: Technical Scope and Functional Model .*JCDL 2002* .Ohio: ACM. Retrieved July 15, 2013:
<http://arxiv.org/ftp/cs/papers/0201/0201027.pdf>
- Garcia-Molina, H. (1994). *Summary94* Stanford University. Retrieved August July 10, 2013:
<http://infolab.stanford.edu/tsimmis/arpa/summary94.html>
- Gottschalk, P & ,Solli-Sæther, H. (2009). *E-government interoperability and information resource integration : frameworks for aligned development* .New York: Information Science reference.

- Guha, N. (2006). *Semantic Digital Library Services*. Retrieved July 15, 2011: http://www.l3s.de/kweb/kwepsy2006/FinalSubmissions/kwepsy2006_guha.pdf
- Gunn, S. (1999). The web is only the beginning. *Texas Library Journal* (Winter), 172-173.
- Hagedorn, K. (2003). OAIster: a "no dead ends" OAI service provider. *Library Hi Tech*, 21(2), 170-181.
- Han, Y. (2006). ARDF-based digital library system. *Library Hi Tech*, 24(2), 234-240. doi:10.1108/07378830610669600
- Hart, Peter E. (2010). *Cases on semantic interoperability for information system integration practices and applications*. New York: Information science reference.
- Horrocks, I. (2002). DAML+OIL: a description logic for the semantic web. *Bulletin of the technical committee on data engineering*, 25(1), 4-9.
- Huang, S.-H., Ke, H.-R & Yang, W.-P. (2005). Enhancing semantic digital library query using a content and service inference model (CSIM). *Information processing and management*, 41, 891-908.
- Introduction Europeana (2011). Retrieved July 12, 2013: <http://www.europeana.eu/portal/api-introduction.html>
- Issac, A., Schlobach, S., Mattheizing, H & Zinn, C. (2008). Integrated access to cultural heritage resources through representation and alignment of controlled vocabularies. *Library Review*, 57(3), 187-199, doi:10.1108/00242530810865475
- Jayakanth, F., Maly, K., Zubair, M & Aswath, L. (2006). A dynamic approach to make CDS/ISIS databases interoperable over the internet using the OAI protocol. *Program: electronic library and information systems*, 40(3), 277-285.
- Jin, Y. (2004). The development of the China Networked Digital Library of Thesis and Dissertations. *Online Information Review*, 28(5), 367-370.
- Kalfoglou, Y.; Bo, Hu (2010). *Streamlining semantic integration systems*. In Hart, P., *Cases on semantic interoperability for information systems integration: practices and applications* New York: Information science reference.
- Lagoze, C., Hoehn, W., Millman, D., Arms, W., Gan, S., Hillmann, D . . . , Kalt, T. (2002). Core Services in the Architecture of the National Digital Library for Science Education (NSDL). (*JCDL 2002* .Portland: ACM. Retrieved July 15, 2013: <http://arxiv.org/ftp/cs/papers/0201/0201025.pdf>
- Lakshminarayan, S. (1998). Semantic interoperability in digital libraries using inter-ontological relation. *A Thesis Submitted to the Graduate Faculty Of The University of Georgia in Partial Fulfillment of the Requirements for the Degree Master of science*.
- Lam, K.-T & Chan, D. L. (2007). Building an institutional repository: sharing experiences at the HKUST library. *OCLC systems and services: international Digital Library Perspectives*, 23(3), 310-323.

- Linckels, S & ,Meinel, C. (2011). *E-Librarian Service:User-Friendly Semantic Search in digital libraries*. New York: Springer.
- LUNGO, F. D. (1998). Alternatives for International Document Availability. Available July 02, 2012: <http://cordis.europa.eu/libraries/en/projects/aida.html>
- Manola, F & ,Miller, E. (2013). *Resource Description Framework*, Retrieved August 4, 2013: http://en.wikipedia.org/wiki/Resource_Description_Framework
- Maurer, h. H. (1998). *Cordis, Libraries: Electronic Remote Access to Information over Networks*. Available Dec. 18, 2012: <http://cordis.europa.eu/libraries/en/projects/liberati.html>
- Mayer, P., Mutschke, P & ,Petras, V. (2008). Reducing semantic complexity in distributed digital libraries: treatment of term vagueness and document re-ranking .*Library Review*, 57(3), 213-224. doi:10.1108/00242530810865484
- McDaniel, B & ,Kruk, S. R. (2009). introduction, In S. B. McDaniel, *Semantic Digital Libraries*, Berlin: springer, 1-6.
- McGuinness , D., & Harmelen, F. (2009, November 12). *OWL Web Ontology Language Overview* . Retrieved February 10, 2013, from W3C: <http://www.w3.org/TR/owl-features/>
- Mihalcea, R., Corley, C & ,Strapparava, C. (2006). Corpus-based and Knowledge-based Measures of Text Semantic Similarity .*AAAI'06*. American Association for Artificial Intelligence, 775-780.
- Miles, A., Matthews, B., Wilson, M., & Brickley, D. (2005). SKOS Core: Simple Knowledge Organisation for the Web. *DC-2005: Proc. Int. Conf. on Dublin Core and Metadata Applications*, (pp. 3-10). Madrid. Retrieved September 12, 2013
- Nelson, M. L., Rucker, J & ,Harrison, T. L. (2003). OAI and NASA's Scientific and Technical Information .*Library Hi Tech*, 21(2), 140150.
- Nilsson, M. (2010). From interoperability to harmonization in metadata standardization: designing an evolvable framework for metadata harmonization .*KTH school of computer science and communication* .Stockholm, Sweden: KTH school of computer science and communication.
- NSDL technical infrastructure white paper version 2.0*(2004). NSDL. Available August 17, 2012: http://nsdl.comm.nsdl.org/meeting/archives/smete/workgroups/technical/nsdl_tech_arch20.doc
- Paepcke, A., Chang, C.-C. K., Winograd, T & ,García-Molina, H. (1998). Interoperability for digital libraries worldwide .*Communications of the ACM*, 41-4, 33-42.
- Park, J & ,Ram, S. (2004). Information systems interoperability: what lies beneath ?*ACM Trans. Inf. Syst.* 22(4), 595-632.
- Partners*[(2008)]. DELOS. Retrieved July 10, 2013: http://www.delos.info/index.php?option=com_content&task=blogcategory&id=18&Itemid=72
- Pasad, A & ,Madalli, D. P. (2008). Faceted infrastructure for semantic digital libraries. *Library Review*, 57(3), 225-234. doi:10.1108/00242530810865493

- Patel, M., Koch, T., Doerr, M & ,Tsinarakis, C. (2005). *DELOS: A Network of Excellence on Digital Libraries* .
- Patel, M., Koch, T., Doerr, M & ,Tsinarakis, C. (2005). *Semantic Interoperability in Digital Library Systems* .UKOLN: DELOS, University of Bath. Retrieved November 18, 2011 : <http://opus.bath.ac.uk/23606/>
- Phillips, S. (1996). Cordis, European SR-Z39.50 Gateway. Retrieved Dec. 18, 2011: <http://cordis.europa.eu/libraries/en/projects/europaga.html>
- Rys, M. (1998). *TIMMIS*, Stanford University, Retrieved July 10, 2013: <http://infolab.stanford.edu/tsimmis/>
- Shen, R. (2006). *Applying the 5S Framework To Integrating Digital Libraries* .Virginia Polytechnic Institute .Virginia: Virginia Polytechnic Institute.
- Shen, X., Tan, B & ,Zhai, C. (2005). ContextSensitive Information Retrieval Using Implicit Feedback .*Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*. New York: ACM., 43-50.
- Shi, R. (2005). *Lightweight federation of non-cooperating digital libraries* .Computer Science. Old Dominion University.
- Soergel, D. (2009). Digital Libraries and Knowledge Organization .S. R. Kruk & ,B. McDaniel. In *Semantic Digital Libraries*. Berlin: Springer, 9-40.
- Synak, M., Dabrowski, M & ,Kruk, S. R. (2009). Semantic Web and Ontologies .S. R. Kruk & ,B. McDaniel ,در *Semantic Digital Libraries*, Berlin: Springer, 41-54.
- Szymański, J. (2011). Cooperative WordNet Editor for Lexical Semantic Acquisition .*IC3K 2009*. Berlin: Springer-Verlag, 187-196.
- Vaishnavi, Vijay K.; Kuechler Jr., William (2008). Design science research methods and patterns: innovating information and communication technology. Boca Raton: Auerbach publications.
- Vdovjak, R & ,Houben, G. J. (2001). *RDF Based Architecture for Semantic Integration of Heterogeneous Information Sources*. Retrieved February 10, 2013: http://iwayan.info/Research/Interoperability/Papers_Research/SemanticMediation/wiiw01.pdf
- Veltman, K. H. (2009). Syntactic and Semantic Interoperability: New Approaches to Knowledge and the Semantic Web .*The New Review of Information Networking*, 7, 159-183.
- Waldt, D. (2005). *Using XML and databases*. Retrieved February 18, 2013: www.gilbane.com
- Warren, P & ,Alsmeyer, D. (2005). Applying semantic technology to a digital library: a case study . *Library management*, 26(4-5), 196-205. doi:10.1108/01435120510596053

بررسی یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال و ارائه الگوی پیشنهادی ... □ ۱۰۳

پیوست‌ها

پیوست ۱. پرسشنامه بررسی وضعیت موجود محتواهای پیشینه‌های کتابشناختی کتابخانه‌های دیجیتال ایران

بنام خدا

مسئول محترم کتابخانه دیجیتال

با عرض سلام و ادب. پرسشنامه حاضر در راستای پروژه پژوهشی با عنوان بررسی یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال و ارائه الگوی پیشنهادی برای کتابخانه‌های دیجیتال ایران برای پاسخ‌گویی به بخشی از پژوهش مذکور با عنوان بررسی وضعیت موجود محتواهای پیشینه‌های کتابشناختی^۱ کتابخانه‌های دیجیتال ایران تهیه شده است. خواهشمند است پس از تکمیل پرسش‌ها، آن را به آدرس ایمیل اینجانب حداکثر تا تاریخ ۲۰ آبان ماه ۱۳۹۰ ارسال نمایید.

با تقدیم احترام

مهدی علیپور حافظی

استادیار پژوهشگاه علوم و فناوری اطلاعات ایران

alipour@irandoc.ac.ir

نرم‌افزار مورد استفاده:

مسئول مربوطه^۲:

تاریخ پاسخ‌گویی:

۱. لطفاً انواع منابع اطلاعاتی موجود در کتابخانه دیجیتال خود را انتخاب کنید (بیش از یک مورد را می‌توان انتخاب کرد):

- | | | | |
|----------------------------------|---|---|---|
| ☐ کتاب‌ها | ☐ نشریات | ☐ مقالات | ☐ عکس‌ها |
| ☐ نقشه‌ها | ☐ گزارش‌های فنی | ☐ پایان‌نامه‌ها | ☐ شنیداری‌ها |
| ☐ ویدئو | ☐ چندرسانه‌ای‌ها | ☐ منابع اینترنتی | ☐ سایر (لطفاً ذکر شود) |

۲. کدام یک از فیلدهای پیشینه‌های کتابشناختی را در مورد منابع اطلاعاتی موجود تکمیل می‌کنید؟ (بیش از یک مورد را می‌توان انتخاب کرد)

- | | |
|---|--|
| ☐ عنوان | ☐ پدیدآور |
| ☐ توصیف (کلیدواژه نمایه) | ☐ پدیدآور همکار |
| ☐ موضوع | ☐ ناشر |
| ☐ شابک (ISBN) | ☐ شاپا (ISSN) |
| ☐ شابم (ISMN) | ☐ دی.او.آی. (DOI) |

۳. اسامی پدیدآورندگان با استفاده از چه منبعی مستند می‌شوند؟ (بیش از یک مورد را می‌توان انتخاب کرد)

- ☐ اسامی پدیدآورندگان مستند نمی‌شوند
- ☐ قسمت شناسنامه اثر در دست بررسی
- ☐ فهرست مستند اسامی مشاهیر و مؤلفان

منظور اطلاعات کتابشناختی منابع اطلاعاتی است که در نرم‌افزار کتابخانه دیجیتال ثبت می‌شوند.^۱

مدیریت کتابخانه دیجیتال، مانند رئیس کتابخانه یا رئیس بخش سازماندهی اطلاعات^۲

- ☐ پایگاه اطلاعاتی کتابخانه ملی
- ☐ پایگاه اطلاعاتی کتابخانه کنگره امریکا
- ☐ فهرست مستندات موجود در نرم‌افزار کتابخانه دیجیتال مورد استفاده^۱ (در صورت وجود)
- سایر (لطفاً ذکر کنید)

۴. اسامی ناشران^۲ با استفاده از چه منبعی مستند می‌شوند؟ (بیش از یک مورد را می‌توان انتخاب کرد)

- ☐ اسامی ناشران مستند نمی‌شوند
- ☐ قسمت شناسنامه اثر در دست بررسی
- ☐ فهرست مستند / سامی نهادها، سازمان ها و موسسات دولتی ایران
- ☐ مرجع نشر
- ☐ وبگاه ناشر
- سایر (لطفاً ذکر کنید)

۵. تحلیل موضوعی منابع اطلاعاتی با استفاده از کدام یک از روش‌های زیر انجام می‌گیرد؟ (بیش از یک مورد را می‌توان انتخاب کرد)

- ☐ موضوع‌دهی کنترل شده با استفاده از سرعنوان‌های موضوعی
- ☐ نمایه‌سازی آزاد با استفاده از کلیدواژه‌های انتخابی از منبع مورد نظر
- ☐ نمایه‌سازی آزاد با استفاده از کلیدواژه‌های انتخابی نمایه‌ساز با توجه به تجربه تخصصی
- ☐ نمایه‌سازی کنترل شده با استفاده از اصطلاحنامه‌ها
- ☐ نمایه‌سازی کنترل شده با استفاده از هستی‌نگاری‌ها
- سایر (لطفاً ذکر کنید)

۶. از انتخاب موضوع برای تحلیل کدام یک از منابع اطلاعاتی زیر استفاده می‌کنید (بیش از یک مورد را می‌توان انتخاب کرد)؟

- | | | | |
|----------------------------------|---|--|-------------------------------------|
| ☐ کتاب‌ها | ☐ نشریات | ☐ مقالات | ☐ عکس‌ها |
| ☐ نقشه‌ها | ☐ گزارش‌های فنی | ☐ پایان‌نامه‌ها | ☐ شنیداری‌ها |
| ☐ ویدئو | ☐ چندرسانه‌ای‌ها | سایر (لطفاً ذکر شود) | |

۷. از انتخاب کلیدواژه برای تحلیل کدام یک از منابع اطلاعاتی زیر استفاده می‌کنید (بیش از یک مورد را می‌توان انتخاب کرد)؟

- | | | | |
|----------------------------------|---|--|-------------------------------------|
| ☐ کتاب‌ها | ☐ نشریات | ☐ مقالات | ☐ عکس‌ها |
| ☐ نقشه‌ها | ☐ گزارش‌های فنی | ☐ پایان‌نامه‌ها | ☐ شنیداری‌ها |
| ☐ ویدئو | ☐ چندرسانه‌ای‌ها | سایر (لطفاً ذکر شود) | |

۸. شیوه شما در سازماندهی منابع اطلاعاتی کدام یک از روش‌های زیر است (به ترتیب اولویت):

^۱ منظور فهرست‌هایی هستند که توسط نرم‌افزار کتابخانه دیجیتال برای یک‌دست ساختن اسامی (مؤلفان، مترجمان، ناشران و غیره) تولید می‌شوند.

^۲ منظور از ناشران تمامی سازمان‌هایی هستند که اقدام به تولید، انتشار و اشاعه اثر می‌کنند. به عنوان مثال ناشر مورد نظر برای پایان‌نامه‌ها دانشگاه محل تحصیل می‌باشد.

□ - بنیادی □ - نسخه برداری

۹. در صورت بهره‌گیری از شیوه نسخه‌برداری برای سازماندهی منابع فارسی از کدام یک از پایگاه‌های زیر استفاده می‌نمایید (به ترتیب اولویت شماره گذاری نمایید)

□ کتابشناسی ملی ایران

□ پایگاه اطلاع‌رسانی کتابخانه‌های ایران (lib.ir)

□ پایگاه اطلاع‌رسانی شبکه سیمرغ

(<http://www.nosa.com/simwebclt/WebAccess/SimWebPortal.dll/SrchDlg>)

سایر (لطفا ذکر کنید):

۱۰. در صورتی که از شیوه نسخه برداری در سازماندهی منابع لاتین بهره می‌گیرید از کدام پایگاه‌های اطلاعاتی بیشتر استفاده می‌نمایید (به ترتیب اولویت شماره گذاری نمایید).

□ کتابخانه کنگره امریکا (www.loc.org)

□ کتابخانه ملی پزشکی امریکا ([http://locatorplus.gov/cgi-](http://locatorplus.gov/cgi-bin/Pwebrecon.cgi?DB=local&PAGE=First)

[bin/Pwebrecon.cgi?DB=local&PAGE=First](http://locatorplus.gov/cgi-bin/Pwebrecon.cgi?DB=local&PAGE=First))

□ کتابخانه ملی بریتانیا (http://catalogue.bl.uk/F/?func=file&file_name=login-bl-list)

□ اوس.ال.سی. (<http://www.oclc.org>)

سایر (لطفا ذکر کنید):

۱۱. در صورتی که از شیوه نسخه برداری استفاده می‌کنید، لطفا میزان اصلاحی که روی فیلدها انجام می‌دهید را مشخص نمایید:

□ بسیار زیاد □ زیاد □ متوسط □ کم □ بسیار کم

۱۲. در صورتی که از شیوه سازماندهی بنیادی استفاده می‌کنید، لطفا منبع مورد استفاده برای ارائه موضوعات را ذکر کنید (بیش از یک مورد را می‌توان انتخاب کرد):

□ استفاده از فهرست‌نویسی پیش از انتشار (منتشر شده در صفحه حقوقی کتاب)

□ سرعنوان‌های موضوعی فارسی

□ سرعنوان‌های موضوعی کنگره امریکا

□ سرعنوان‌های موضوعی سیرز

سایر (لطفا ذکر کنید)

۱۳. برای سازماندهی منابع اطلاعاتی خود به چه میزان از شیوه سازماندهی بنیادی یا نسخه برداری استفاده می‌کنید؟

سازماندهی به شیوه بنیادی

□ بسیار زیاد □ زیاد □ متوسط □ کم □ بسیار کم

سازماندهی به شیوه نسخه برداری

۱۰۷ □ بررسی یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال و ارائه الگوی پیشنهادی ... □

□ بسیار زیاد □ زیاد □ متوسط □ کم □ بسیار کم

۱۴. در صورت استفاده از نمایه‌سازی کنترل شده، لطفاً منبع مورد استفاده را ذکر کنید (بیش از یک مورد را می‌توان انتخاب کرد):

□ اصطلاحنامه‌های موضوعی □ اصطلاحنامه‌های عمومی
□ هستی‌نگاری‌ها (Ontologies) □ سایر (لطفاً ذکر کنید):

۱۵. در صورت استفاده از اصطلاحنامه‌های موضوعی، لطفاً اصطلاحنامه‌های مورد استفاده را مشخص کنید (بیش از یک مورد را می‌توان انتخاب کرد):

□ اصطلاحنامه کشاورزی	□ اصطلاحنامه پزشکی فارسی
□ اصطلاحنامه ارتقاء بهداشت	□ اصطلاحنامه جامعه‌شناسی
□ اصطلاحنامه ریاضیات	□ اصطلاحنامه شیمی
□ اصطلاحنامه علوم زمین	□ اصطلاحنامه علوم زیستی
□ اصطلاحنامه فنی و مهندسی	□ اصطلاحنامه فیزیک
□ اصطلاحنامه مدیریت بحران	□ اصطلاحنامه علوم قرآن
□ اصطلاحنامه فلسفه	□ اصطلاحنامه منطق
□ اصطلاحنامه اصول فقه	□ اصطلاحنامه اخلاق
□ اصطلاحنامه کلام اسلامی	□ سایر (لطفاً ذکر کنید)

۱۶. در صورت استفاده از اصطلاحنامه‌های موضوعی لاتین برای منابع اطلاعاتی لاتین، از کدام یک از اصطلاحنامه‌های زیر استفاده می‌کنید (بیش از یک مورد را می‌توان انتخاب کرد):

□ Agrovoc	□ Biological Science
□ Chemistry	□ Crisis management
□ Engineering	□ Geosciences
□ Health promotion	□ Mathematics
□ Physics	□ Sociology
□ Inspec	Others (Please specify)

۱۷. در صورت استفاده از اصطلاحنامه‌های عمومی، لطفاً اصطلاحنامه‌های مورد استفاده را مشخص کنید (بیش از یک مورد را می‌توان انتخاب کرد):

□ اصطلاحنامه سه زبانه اصفا
□ اصطلاحنامه نما
□ سایر (لطفاً ذکر کنید)

۱۸. در صورت استفاده از هستی‌نگاری‌ها، منابع مورد استفاده را ذکر کنید (بیش از یک مورد را می‌توان انتخاب کرد):

□ وردنت فارسی

□ Wordnet

۱۰۸ □ بررسی یکپارچه‌سازی معنایی اطلاعات در کتابخانه‌های دیجیتال و ارائه الگوی پیشنهادی ... □

□ Corposes

سایر (لطفا ذکر کنید):

پیوست ۲. سیاهه واری ارزیابی قابلیت‌های خروجی نرم‌افزارهای کتابخانه دیجیتال

۱. چه خروجی‌های ابر داده‌ای به صورت پیش تعریف در سیستم وجود دارد؟ (لطفاً نسخه‌های مربوطه نیز ذکر شوند) (بیش از یک مورد را می‌توان انتخاب کرد)
استاندارد در لایه زمینه (Context)

- ISO text – □ Structured text – □ XML –
□ ODBC Interface – □ HTML – سایر (لطفاً ذکر کنید) –

استاندارد در لایه بالاتر

- (MARC) – مارک ایکس.ام.ال. □ (MARCXML) –
□ (Dublin Core) – دابلین کُر □ (MODS) – سایر (لطفاً ذکر شود) –

۲. آیا شمای خروجی‌های ابر داده مورد استفاده در دسترس کاربران هستند؟ (آیا شمای مورد نظر می‌تواند در اختیار پژوهشگر قرار گیرد؟)
- بلی □ - خیر □

۳. در صورت ارائه خروجی ابر داده از پایگاه اطلاعاتی، ارتباط آن با فایل منبع چگونه برقرار می‌شود؟
- ارتباط بین ابر داده و شیء دیجیتال در فرمت‌های خروجی وجود ندارد □
- با استفاده از قابلیت فرامتن موجود در ابر داده □
- به صورت درونه‌ای (Embedded) □
- سایر (لطفاً ذکر شود)

۴. آیا امکان نمایه‌سازی و دسترسی به ابر داده توسط موتورهای جستجوی عمومی یا تخصصی وب وجود دارد؟
- بلی (با برنامه‌ریزی اختصاصی در این خصوص) □ - بلی (به صورت عادی) □
- خیر (محدود شده است) □

۵. آیا امکان مبادله اطلاعات با سیستم‌های اطلاعاتی سایر مراکز وجود دارد؟
- بلی (لطفاً استاندارد ارتباطی مورد استفاده را ذکر کنید) □
- خیر □ - تا حدودی □

۶. در صورت عدم اجرای مبادله در حال حاضر، آیا چنین قابلیتی در سیستم وجود دارد که بدون نیاز به حضور متخصصان نرم‌افزار قابل اجرا باشد؟ یا به سادگی بتواند در اختیار کتابداران قرار گیرد؟
- بلی □ - خیر □ - تا حدودی □

۷. در صورت وجود امکان مبادله اطلاعات در سیستم، این قابلیت در چه سطحی قابل اجرا است؟
- اطلاعات کتابشناختی (ابر داده) □ - اشیاء دیجیتال □

– اطلاعات کتابشناختی و اشیاء دیجیتال □ – سایر (لطفاً ذکر شود)

۸. برای مبادله اطلاعات از چه تفاهم نامه‌ای استفاده می‌شود؟ (بیش از یک مورد را می‌توان انتخاب کرد)

– 339.50 (سرویس دهنده/ سرویس گیرنده) □ – اس.آر.یو. (SRU) □

– اس.آر.دبلیو. (SRW) □ – او ای آی (OAI) □

– سایر (لطفاً ذکر شود)

۹. فرمت مورد استفاده در مبادله اطلاعات کتابشناختی چیست؟

استاندارد در لایه زمینه (Context)

– ISO text □ – Structured text □ – XML □

– ODBC Interface □ – HTML □ – سایر (لطفاً ذکر کنید)

استاندارد در لایه بالاتر برای ابر داده توصیفی

– مارک (MARC) □ – مارک ایکس ام ال (MARCXML) □

– دابلین کُر (Dublin Core) □ – مُدز □

– سایر (لطفاً ذکر شود)

استاندارد در لایه بالاتر برای ابر داده ساختاری و مدیریتی

– متر (METS) □ – ای.ای.دی. (EAD) □

– سایر (لطفاً ذکر شود)

پیوست ۳. نمونه فایل ایکس.ام.ال. دریافتی از کتابخانه دیجیتال مجلس

```

xml version="1.0" encoding="UTF-8"?><collection>
xmlns="http://www.loc.gov/MARC21/slim"><?xml version = '1.0' encoding = 'UTF-8'?>
<record><leader>00987nam0022002771n04500</leader><controlfield
tag="001">848868</controlfield><datafield tag="020" ind1=" " ind2=" "><subfield
code="a">IR</subfield><subfield code="b">/>١٢-٣٦٧</subfield></datafield><datafield
tag="100" ind1=" " ind2=" "><subfield code="a">20101002 per 50
</subfield></datafield><datafield tag="101" ind1="0" ind2=" "><subfield
code="a">ara</subfield></datafield><datafield tag="102" ind1=" " ind2="
"><subfield code="a">ET</subfield></datafield><datafield tag="200" ind1="1" ind2="
"><subfield code="a">والادب</subfield><subfield code="f" />"
/>؛ صححه و ضبط الفاظه و علق</subfield><subfield code="g"> [تأليف] الشريف ابى القاسم على بن الطاهر ابى احمد الحسين</
/>؛ على نفقه احمد ناجى الجمالى و محمد امين الخانجى و</subfield><subfield code="g"> حواشيه احمد بن الشنقيطى</
/>اخيه</subfield></datafield><datafield tag="210" ind1=" " ind2=" "><subfield
code="e">مصر</subfield><subfield code="g"> مطبعة السعادة</subfield><subfield
code="h"> ١٢٨٥ ، = ١٩٠٧ = ١٣٢٥ ق.</subfield></datafield><datafield tag="215" ind1="
" ind2=" "><subfield code="a">/> ج. ٢ مجلد</subfield></datafield><datafield
tag="606" ind1="2" ind2=" "><subfield code="9">قرن ٥</subfield></datafield><datafield
tag="606" ind1="2" ind2=" "><subfield code="9">قرن ٥</subfield></datafield><datafield
tag="606" ind1="1" ind2=" "><subfield code="9">قرن ٥</subfield></datafield><datafield
tag="680" ind1=" " ind2=" "><subfield code="a">BP٩٤</subfield><subfield
code="b">/ الف ٨ع ١٢٨٥</subfield></datafield><datafield tag="700" ind1=" "
ind2="1"><subfield code="9">٣٥٥</subfield></datafield><datafield tag="702" ind1=" " ind2="1"><subfield
code="9">٣٤٦</subfield></datafield><datafield tag="801" ind1=" "
ind2="0"><subfield code="a">IR</subfield><subfield code="b">كتابخانه مجلس شوراى
اسلامى</subfield><subfield code="c">20121015</subfield></datafield><datafield
tag="856" ind1=" " ind2=" "><subfield code="f">/><subfield code="e">26-
02646.pdf</subfield></datafield><datafield tag="910" ind1=" " ind2=" "><subfield
code="b">arghash</subfield><subfield
code="c">20121015</subfield></datafield><datafield tag="911" ind1=" " ind2="
"><subfield code="a">emami</subfield><subfield
code="b">20121015082331.0</subfield></datafield><datafield tag="915" ind1=" "
ind2=" "><subfield code="a">فهرست نویسی قبلى</subfield></datafield><datafield tag="930"
ind1=" " ind2=" "><subfield code="a">SF</subfield><subfield
code="b">215332</subfield><subfield code="c">1</subfield></datafield><datafield
tag="932" ind1=" " ind2=" "><subfield
code="b">Y</subfield></datafield></record><?xml version = '1.0' encoding = 'UTF-
8'?>
<record><leader>00624naa 2200193 450 </leader><controlfield
tag="001">896028</controlfield><datafield tag="020" ind1=" " ind2=" "><subfield
code="b">/>١٢-٤٩٤</subfield></datafield><datafield tag="100" ind1=" " ind2="
"><subfield code="a">20110206 per 50
</subfield></datafield><datafield tag="101" ind1="0" ind2=" "><subfield
code="a">ara</subfield></datafield><datafield tag="200" ind1="0" ind2="
"><subfield code="a">دليل السالك لمذهب الامام مالك فى جميع العبادات و المعاملات و</
/>تأليف محمد محمد سعد</subfield></datafield><datafield
tag="210" ind1=" " ind2=" "><subfield code="e">مصر</subfield><subfield
code="g"> مكتبة و مطبعة محمد على صبيح و اولاده</subfield><subfield
code="h"> ١٣٣٦ ق. =</subfield><subfield code="h"> ١٢٩٦</subfield></datafield><datafield tag="215" ind1=" " ind2=" "><subfield
code="a">/> ١٩٦٦</subfield></datafield><datafield tag="305" ind1=" " ind2="

```

```

"><subfield code="a/>چاپ سربی.</subfield></datafield><datafield tag="801" ind1=" "
ind2=" "><subfield code="a/>ایران</subfield><subfield code="b">کتابخانه، موزه و مرکز اسناد مجلس</subfield></datafield><datafield
tag="856" ind1=" " ind2=" "><subfield code="f"/><subfield code="e">26-
02594.pdf</subfield></datafield><datafield tag="910" ind1=" " ind2=" "><subfield
code="b">f-bagheri</subfield><subfield
code="c">20120929</subfield></datafield><datafield tag="911" ind1=" " ind2="
"><subfield code="a">sh-aslani</subfield><subfield
code="b">20120929133256.0</subfield></datafield><datafield tag="930" ind1=" "
ind2=" "><subfield code="a">SF</subfield><subfield
code="b">215332</subfield><subfield code="c">1</subfield></datafield><datafield
tag="932" ind1=" " ind2=" "><subfield code="b">Y
/><subfield></datafield></record><?xml version = '1.0' encoding = 'UTF-8?'
<record><leader>00939nam0022003251n04500</leader><controlfield
tag="001">953720</controlfield><datafield tag="020" ind1=" " ind2=" "><subfield
code="a">IR</subfield><subfield code="b">/>۱۲-۲۵۱۵</subfield></datafield><datafield
tag="100" ind1=" " ind2=" "><subfield code="a">20110629 per 50
</subfield></datafield><datafield tag="101" ind1="0" ind2=" "><subfield
code="a">ara</subfield></datafield><datafield tag="102" ind1=" " ind2="
"><subfield code="a">IR</subfield></datafield><datafield tag="200" ind1="1" ind2="
"><subfield code="a/>لمولفه ابوعلی الفضل</subfield><subfield code="f">مجمع البيان في تفسير القرآن</subfield></datafield><datafield tag="210" ind1=" " ind2=" "><subfield
code="a/>صیدا</subfield><subfield code="c/>مطبعة العرفان</subfield></datafield><datafield tag="215" ind1=" "
ind2=" "><subfield code="a">ج۵</subfield></datafield><datafield tag="305" ind1=" "
ind2=" "><subfield code="a/>چاپ سربی</subfield></datafield><datafield tag="606"
ind1="1" ind2=" "><subfield code="9">قرن ۶</subfield></datafield><datafield tag="680" ind1=" " ind2=" "><subfield
code="a">BP۵/۹۴</subfield><subfield code="b">/>ط ۲
۳</subfield></datafield><datafield tag="700" ind1=" " ind2="1"><subfield
code="9">۴۶۸</subfield></datafield><datafield tag="801" ind1=" "
ind2="0"><subfield code="a">IR</subfield><subfield code="b">کتابخانه مجلس شورای
اسلامی</subfield><subfield code="c">20121104</subfield></datafield><datafield
tag="856" ind1=" " ind2=" "><subfield code="f"/><subfield code="e">18-
01086.pdf</subfield></datafield><datafield tag="856" ind1=" " ind2=" "><subfield
code="f"/><subfield code="e">18-01166.pdf</subfield></datafield><datafield
tag="856" ind1=" " ind2=" "><subfield code="f"/><subfield code="e">19-
03949.pdf</subfield></datafield><datafield tag="856" ind1=" " ind2=" "><subfield
code="f"/><subfield code="e">24-05057.pdf</subfield></datafield><datafield
tag="910" ind1=" " ind2=" "><subfield code="b">az-eskandarlo</subfield><subfield
code="c">20121104</subfield></datafield><datafield tag="911" ind1=" " ind2="
"><subfield code="a">emami</subfield><subfield
code="b">20121029114148.0</subfield></datafield><datafield tag="911" ind1=" "
ind2=" "><subfield code="a">emami</subfield><subfield
code="b">20121030080658.0</subfield></datafield><datafield tag="911" ind1=" "
ind2=" "><subfield code="a">emami</subfield><subfield
code="b">20121030111717.0</subfield></datafield><datafield tag="911" ind1=" "
ind2=" "><subfield code="a">emami</subfield><subfield
code="b">20121104094153.0</subfield></datafield><datafield tag="915" ind1=" "
ind2=" "><subfield code="a">فهرست نویسی قبلی</subfield></datafield><datafield tag="930"
ind1=" " ind2=" "><subfield code="a">SF</subfield><subfield
code="b">215332</subfield><subfield code="c">1</subfield></datafield><datafield
tag="932" ind1=" " ind2=" "><subfield
code="b">Y</subfield></datafield></record><?xml version = '1.0' encoding = 'UTF-
8'?>
<record><leader>00507naa 2200181 450 </leader><controlfield
tag="001">885154</controlfield><datafield tag="020" ind1=" " ind2=" "><subfield
code="b">/>۱۲-۴۷۹</subfield></datafield><datafield tag="100" ind1=" " ind2="

```

```

"><subfield code="a">20110112 per 50
</subfield></datafield><datafield tag="101" ind1="0" ind2=" "><subfield
code="a">ara</subfield></datafield><datafield tag="200" ind1="0" ind2="
"><subfield code="a/>حاشیه میرزا زاهد علی شرح التهذيب</subfield></datafield><datafield
tag="210" ind1=" " ind2=" "><subfield code="e/>قزان</subfield><subfield
code="g/>مطبع الملکی</subfield><subfield code="h"> = ۱۳۰۵ ق. </subfield></datafield><datafield tag="215" ind1=" " ind2=" "><subfield
code="a">/>۱۲۶۶ ص. ۱۳۳</subfield></datafield><datafield tag="801" ind1=" " ind2="
"><subfield code="a/>ایران</subfield><subfield code="b">موزه و مرکز اسناد مجلس شورای</subfield></datafield><datafield
tag="856" ind1=" " ind2=" "><subfield code="f/><subfield code="e">18-
01211.pdf</subfield></datafield><datafield tag="910" ind1=" " ind2=" "><subfield
code="b">admin</subfield><subfield
code="c">20121030</subfield></datafield><datafield tag="911" ind1=" " ind2="
"><subfield code="a">emami</subfield><subfield
code="b">20121030083914.0</subfield></datafield><datafield tag="930" ind1=" "
ind2=" "><subfield code="a">SF</subfield><subfield
code="b">215332</subfield><subfield code="c">1</subfield></datafield><datafield
tag="932" ind1=" " ind2=" "><subfield code="b">Y
/><subfield></datafield></record><?xml version = '1.0' encoding = 'UTF-8?>
<record><leader>01074nam0022003371n04500</leader><controlfield
tag="001">821299</controlfield><controlfield
tag="005">20110702072254.0</controlfield><datafield tag="020" ind1=" " ind2="
"><subfield code="a">IR</subfield><subfield code="b">۱۲-
/>۲۶۰<subfield></datafield><datafield tag="100" ind1=" " ind2=" "><subfield
code="a">20100706 per 50 </subfield></datafield><datafield
tag="101" ind1="0" ind2=" "><subfield
code="a">ara</subfield></datafield><datafield tag="102" ind1=" " ind2="
"><subfield code="a">IR</subfield></datafield><datafield tag="200" ind1="1" ind2="
"><subfield code="a/>زبدہ الصحائف فی اصول المعارف</subfield><subfield code="f
/>تالیف نوفل</subfield></datafield><datafield tag="210" ind1=" " ind2=" "><subfield
code="a/>بمینی</subfield><subfield code="c/>مطبعه ناصری</subfield><subfield
code="d"> ۱۲۸۳ ق. = ۱۳۲۳ ق. </subfield></datafield><datafield tag="215" ind1=" "
ind2=" "><subfield code="a/>ج</subfield></datafield><datafield tag="305" ind1=" "
ind2=" "><subfield code="a/>چاپ سربی</subfield></datafield><datafield tag="324"
ind1=" " ind2=" "><subfield code="a">۱. اسکن شده</subfield></datafield><datafield tag="517" ind1="1" ind2=" "><subfield
code="a/>زبدہ الصحائف فی سیاحه المعارف</subfield></datafield><datafield tag="606" ind1="1"
ind2=" "><subfield code="9">۱۴ متون قدیمی تا قرن</subfield></datafield><datafield tag="680" ind1=" " ind2=" "><subfield
code="a">AG۵۵/۱۱۰</subfield><subfield code="b">۴ ز ۲</subfield></datafield><datafield tag="700" ind1=" " ind2="1"><subfield
code="9"></subfield></datafield><datafield tag="801" ind1=" " ind2="0"><subfield code="a">IR</subfield><subfield code="b">کتابخانه مجلس شورای</subfield></datafield><datafield
tag="856" ind1=" " ind2=" "><subfield code="f">13-
00150.pdf</subfield></datafield><datafield tag="856" ind1=" " ind2=" "><subfield
code="f/><subfield code="e">18-01047.pdf</subfield></datafield><datafield
tag="910" ind1=" " ind2=" "><subfield code="b">admin</subfield><subfield
code="c">20121029</subfield></datafield><datafield tag="911" ind1=" " ind2="
"><subfield code="a">mo-asadolahi</subfield><subfield
code="b">20080807120658.0</subfield></datafield><datafield tag="911" ind1=" "
ind2=" "><subfield code="a">e-pouya</subfield><subfield
code="b">20110702072254.0</subfield></datafield><datafield tag="911" ind1=" "
ind2=" "><subfield code="a">e-pouya</subfield><subfield
code="b">20111017080817.0</subfield></datafield><datafield tag="911" ind1=" "
ind2=" "><subfield code="a">emami</subfield><subfield

```

```
code="b">20121029102842.0</subfield></datafield><datafield tag="915" ind1=" "
ind2=" "><subfield code="a">فهرست نویسی قبلی</subfield></datafield><datafield tag="930"
ind1=" " ind2=" "><subfield code="a">SF</subfield><subfield
code="c">1</subfield><subfield code="b">215332</subfield></datafield><datafield
tag="932" ind1=" " ind2=" "><subfield
code="b">Y</subfield></datafield></record><?xml version = '1.0' encoding = 'UTF-
8'?>
<record><leader>00727nam 22002171i 450 </leader><controlfield
tag="001">852420</controlfield><datafield tag="020" ind1=" " ind2=" "><subfield
code="b">۱۲-۳۲۷</subfield></datafield><datafield tag="100" ind1=" " ind2="
"><subfield code="a">20101012 per 50
</subfield></datafield><datafield tag="101" ind1="0" ind2=" "><subfield
code="a">per</subfield></datafield><datafield tag="200" ind1="1" ind2="
"><subfield code="a">وزارت ۱۳۰۷ آخرین پروگرام جدید ۱۳۰۷ وزارت
تالیف محمدحسین /<subfield code="f"></subfield></datafield><datafield tag="210" ind1=" " ind2=" "><subfield
code="a">تهران</subfield><subfield code="c">: کتابخانه اقبال</subfield><subfield
code="d">/>. ۱۳۰۷</subfield></datafield><datafield tag="215" ind1=" " ind2="
"><subfield code="a">۱۴۴ ص.</subfield><subfield
code="c"></subfield></datafield><datafield tag="312" ind1=" " ind2="
"><subfield code="a">بالاتر عنوان: مجموعه امیر.</subfield></datafield><datafield
tag="517" ind1="1" ind2=" "><subfield code="a">مجموعه امیر.</subfield></datafield><datafield
tag="700" ind1=" " ind2="1"><subfield code="a">مدرس نهانندی</subfield><subfield
code="b">/> محمد حسین</subfield><subfield code="f">۱۲۷۷</subfield></datafield><datafield
tag="801" ind1=" " ind2="0"><subfield code="a">ایران</subfield><subfield code="b">کتابخانه</subfield></datafield><datafield
tag="856" ind1=" " ind2=" "><subfield code="f">/><subfield code="e">11-
03275.pdf</subfield></datafield><datafield tag="910" ind1=" " ind2=" "><subfield
code="b">f-bagheri</subfield></datafield><datafield tag="911" ind1=" " ind2="
"><subfield code="a">emami</subfield></datafield><datafield tag="930" ind1=" "
ind2=" "><subfield code="a">SF</subfield><subfield
code="b">215332</subfield><subfield code="c">1</subfield></datafield><datafield
tag="932" ind1=" " ind2=" "><subfield code="a">a</subfield></datafield><datafield
tag="932" ind1=" " ind2=" "><subfield code="a">a</subfield></datafield></record><?xml version = '1.0' encoding = 'UTF-
8'?>
<record><leader>00714nam0022002411n04500</leader><controlfield
tag="001">954432</controlfield><controlfield
tag="005">20110703084142.0</controlfield><datafield tag="020" ind1=" " ind2="
"><subfield code="a">IR</subfield><subfield code="b">۱۲-
۳۱۳۵</subfield></datafield><datafield tag="100" ind1=" " ind2=" "><subfield
code="a">20110703 per 50 </subfield></datafield><datafield tag="101"
ind1="0" ind2=" "><subfield code="a">per</subfield></datafield><datafield
tag="102" ind1=" " ind2=" "><subfield code="a">IR</subfield></datafield><datafield
tag="200" ind1="1" ind2=" "><subfield code="a">&lt;...> دستور زیبایی</subfield></datafield><datafield
tag="210" ind1=" " ind2=" "><subfield code="a">[تهران]</subfield><subfield
code="c">: چاپخانه خودکار</subfield><subfield code="d">،
۱۳۱۶</subfield></datafield><datafield tag="215" ind1=" " ind2=" "><subfield
code="a">ص ۱۳۶</subfield></datafield><datafield tag="700" ind1=" "
ind2="1"><subfield code="9">عبدالحسین</subfield></datafield><datafield tag="801"
ind1=" " ind2="0"><subfield code="a">IR</subfield><subfield code="b">کتابخانه ملی</subfield></datafield><datafield
tag="856" ind1=" " ind2=" "><subfield code="f">/><subfield code="e">18-
01060.pdf</subfield></datafield><datafield tag="910" ind1=" " ind2=" "><subfield
code="b">sh-aslani</subfield></datafield></record></pre>

```

```
code="c">20121029</subfield></datafield><datafield tag="911" ind1=" " ind2="
"><subfield code="a">sh-aslani</subfield><subfield
code="b">20110703084142.0</subfield></datafield><datafield tag="911" ind1=" "
ind2=" "><subfield code="a">emami</subfield><subfield
code="b">20121029110447.0</subfield></datafield><datafield tag="915" ind1=" "
ind2=" "><subfield code="a">قبلی فهرست نویسی</subfield></datafield><datafield tag="930"
ind1=" " ind2=" "><subfield code="a">SF</subfield><subfield
code="b">215332</subfield><subfield code="c">1</subfield></datafield><datafield
tag="932" ind1=" " ind2=" "><subfield
code="b">Y</subfield></datafield></record><?xml version = '1.0' encoding = 'UTF-
8'?>
<record><leader>01916nam0022003731n04500</leader><controlfield
tag="001">531354</controlfield><datafield tag="020" ind1=" " ind2=" "><subfield
code="a">IR</subfield><subfield code="b">۱۱-۲۴۹۱</subfield></datafield><datafield
tag="100" ind1=" " ind2=" "><subfield code="a">20071219 per 50
</subfield></datafield><datafield tag="101" ind1="0" ind2=" "><subfield
code="a">per</subfield><subfield code="a">ara</subfield></datafield><datafield
tag="102" ind1=" " ind2=" "><subfield code="a">IR</subfield></datafield><datafield
tag="200" ind1="1" ind2=" "><subfield code="a">نفحه الیمن فیما یزول بنکره
</subfield><subfield code="f">/تالیف احمد بن محمد شروانی یمنی / الشجن
code="f">عبدالکریم بن عباس علی تفرشی معروف بمیرزا حاجی آقا واعظ ؛؛ مصحح و مترجم
code="g">بن حاج محمد علی ؛؛ کاتب</subfield><subfield code="g">متن مترجم شعر محمد وجدانی</subfield><subfield
tag="210" ind1=" " ind2=" "><subfield
code="a">[&lt;تهران=طهران]</subfield><subfield code="c">عبدالکریم بن عباس علی
</subfield><subfield code="d">۱۳۰۶ ق. = ۱۳۴۶ م. </subfield><subfield
code="e">[&lt;تهران=طهران]</subfield><subfield code="g">مطبعة استاد حاج
</subfield></datafield><datafield tag="215" ind1=" " ind2=" "><subfield
code="a">ص ۳۲۴</subfield><subfield
code="d">۲۵م ۱۷</subfield></datafield><datafield tag="300" ind1=" " ind2="
"><subfield code="a">فهرست مندرجات کتاب در سال ۱۳۰۹ و به صورت چاپ سری منتشر شده
</subfield></datafield><datafield tag="300" ind1=" " ind2=" "><subfield
code="a">متن عربی ترجمه فارسی در حاشیه
</subfield></datafield><datafield tag="330" ind1="
" ind2=" "><subfield code="a">این کتاب توسط پدر نویسنده حاج میرزا حاجی آقا واعظ از عربی به فارسی ترجمه شده و
</subfield></datafield><datafield tag="606" ind1="1" ind2=" "><subfield
code="9">مجموعه ها - ادبیات عربی
</subfield></datafield><datafield tag="680" ind1=" "
ind2=" "><subfield code="a">PJA۲۵۴۶</subfield><subfield code="b">ش ۴
</subfield></datafield><datafield tag="700" ind1=" " ind2="0"><subfield
code="9">۸۳۷م - شروانی، احمد بن محمد،
</subfield></datafield><datafield tag="702" ind1="
" ind2="0"><subfield code="9">۱۳۰۷، - تفرشی، عبدالکریم،
</subfield></datafield><datafield tag="702" ind1=" " ind2="0"><subfield
code="9">مترجم</subfield></datafield><datafield tag="801" ind1=" "
ind2="0"><subfield code="a">IR</subfield><subfield code="b">کتابخانه مجلس شورای
</subfield><subfield code="c">20121104</subfield></datafield><datafield
tag="856" ind1=" " ind2=" "><subfield code="f">13-
00208.pdf</subfield></datafield><datafield tag="856" ind1=" " ind2=" "><subfield
code="f">22-04246.pdf</subfield></datafield><datafield tag="856" ind1=" " ind2="
"><subfield code="f">/</subfield><subfield code="e">24-
05037.pdf</subfield></datafield><datafield tag="910" ind1=" " ind2=" "><subfield
code="c">20121104</subfield><subfield code="e">شاهی</subfield><subfield
code="f">۰۴/۸۱</subfield></datafield><datafield tag="911" ind1=" " ind2="
"><subfield code="a">sh-aslani</subfield><subfield
code="b">20110809095714.0</subfield></datafield><datafield tag="911" ind1=" "
ind2=" "><subfield code="a">ho-afraziabi</subfield><subfield
code="b">20111104005429.0</subfield></datafield><datafield tag="911" ind1=" "
```

```
ind2=" "><subfield code="a">ho-afraziabi</subfield><subfield
code="b">20120611094045.0</subfield></datafield><datafield tag="911" ind1=" "
ind2=" "><subfield code="a">emami</subfield><subfield
code="b">20121104091008.0</subfield></datafield><datafield tag="915" ind1=" "
ind2=" "><subfield code="a">فهرست‌نویسی قبلی</subfield></datafield><datafield tag="930"
ind1=" " ind2=" "><subfield code="a">LF</subfield><subfield
code="b">86055</subfield><subfield code="c">1</subfield></datafield><datafield
tag="930" ind1=" " ind2=" "><subfield code="a">SF</subfield><subfield
code="c">1</subfield></datafield><datafield tag="932" ind1=" " ind2=" "><subfield
code="b">Y</subfield></datafield></record><?xml version = '1.0' encoding = 'UTF-
8'?>
<record><leader>01100nam0022003131n04500</leader><controlfield
tag="001">953677</controlfield><controlfield
tag="005">20110629074513.0</controlfield><datafield tag="020" ind1=" " ind2="
"><subfield code="a">IR</subfield><subfield code="b">۱۲-
/>۳۵۱۰</subfield></datafield><datafield tag="100" ind1=" " ind2=" "><subfield
code="a">20110629 per 50 </subfield></datafield><datafield
tag="101" ind1="0" ind2=" "><subfield
code="a">ara</subfield></datafield><datafield tag="102" ind1=" " ind2="
"><subfield code="a">IR</subfield></datafield><datafield tag="200" ind1="1" ind2="
"><subfield code="a">&lt;/>مثل السائر في ادب الكاتب و الشاعر</subfield><subfield code="f /<
/>بهاشيه</subfield><subfield code="a">تأليف ابي محمد عبدالله بن مسلم ابن قتيبه</subfield><subfield code="f /<
/>تأليف ابي محمد عبدالله بن مسلم ابن قتيبه</subfield><subfield code="a">ادب</subfield></datafield><datafield tag="210" ind1=" " ind2=" "><subfield
code="a/>قاهره</subfield><subfield code="c/>مطبعة حجازي</subfield><subfield
code="d">/>.۱۳۱۴</subfield></datafield><datafield tag="215" ind1=" "
ind2=" "><subfield code="a">/>۳۴۳</subfield></datafield><datafield tag="305"
ind1=" " ind2=" "><subfield code="a/>چاپ سربی</subfield></datafield><datafield
tag="517" ind1="1" ind2=" "><subfield code="a">ادب</subfield></datafield><datafield tag="606" ind1="2" ind2=" "><subfield
code="9/>فن نگارش</subfield></datafield><datafield tag="606" ind1="1"
ind2=" "><subfield code="9/>معانی و بیان</subfield></datafield><datafield
tag="680" ind1=" " ind2=" "><subfield code="a">PJ۶۱۶۱</subfield><subfield
code="b">۲۲م الف/</subfield></datafield><datafield tag="700" ind1=" " ind2="
"><subfield code="9">۵۵۸</subfield></datafield><datafield tag="701" ind1=" " ind2="1"><subfield code="a">ابن قتيبه</subfield><subfield
code="b">عبدالله بن مسلم</subfield><subfield code="f">۲۱۳</subfield></datafield><subfield
code="3">85840</subfield></datafield><datafield tag="801" ind1=" "
ind2="0"><subfield code="a">IR</subfield><subfield code="b">کتابخانه مجلس شورای
اسلامی</subfield><subfield code="c">20121104</subfield></datafield><datafield
tag="856" ind1=" " ind2=" "><subfield code="f">/>26-
02766.pdf</subfield></datafield><datafield tag="910" ind1=" " ind2=" "><subfield
code="b">az-eskandarlo</subfield><subfield
code="c">20121104</subfield></datafield><datafield tag="911" ind1=" " ind2="
"><subfield code="a">az-eskandarlo</subfield><subfield
code="b">20110629074513.0</subfield></datafield><datafield tag="911" ind1=" "
ind2=" "><subfield code="a">emami</subfield><subfield
code="b">20121104140426.0</subfield></datafield><datafield tag="915" ind1=" "
ind2=" "><subfield code="a">فهرست‌نویسی قبلی</subfield></datafield><datafield tag="930"
ind1=" " ind2=" "><subfield code="a">SF</subfield><subfield
code="b">215332</subfield><subfield code="c">1</subfield></datafield><datafield
tag="932" ind1=" " ind2=" "><subfield
code="b">Y</subfield></datafield></record><?xml version = '1.0' encoding = 'UTF-
8'?>
<record><leader>01110nam0022003491n04500</leader><controlfield
tag="001">430153</controlfield><datafield tag="020" ind1=" " ind2=" "><subfield
code="b">۳۵۹۰</subfield></datafield><datafield tag="100" ind1=" " ind2="
```

```

"><subfield code="a">20101109 per 50
</subfield></datafield><datafield tag="101" ind1="0" ind2=" "><subfield
code="a">per</subfield></datafield><datafield tag="102" ind1=" " ind2="
"><subfield code="a">IR</subfield></datafield><datafield tag="200" ind1="1" ind2="
"><subfield code="a">[دیوان دیری]</subfield><subfield code="f">/ مصنفه پندتراجاکول عرض‌بیگی
؛ تعلیقات و حواشی از نند لعل‌کول <subfield code="g"></subfield><subfield code="g">متخلص به دیری
طالب</subfield></datafield><datafield tag="210" ind1=" " ind2=" "><subfield
code="a">[سرینگر، کشمیر]</subfield><subfield code="c">: [ محکمه تحقیق و حکومت جمون و کشمیر،
<subfield code="d">، [ ۱۹۴۵= ۱۳۲۴ اشاعت
.</subfield></datafield><datafield tag="215" ind1=" " ind2=" "><subfield
code="a">xصxi، 18، ۲۰۸.</subfield></datafield><datafield tag="225" ind1=" " ind2="
"><subfield code="a">نمبر ۷۸</subfield><subfield code="v">سلسله اشاعت مخطوطات کشمیر</subfield></datafield><datafield tag="300" ind1=" " ind2=" "><subfield
code="a">چاپ سربی</subfield></datafield><datafield tag="305" ind1=" " ind2="
"><subfield code="a">چاپ سنگی</subfield></datafield><datafield tag="312" ind1=" "
ind2=" "><subfield code="a">Diwan -i- عنوان پشت جلد و مقدمه به انگلیسی</subfield></datafield><datafield tag="500" ind1=" " ind2="1"><subfield
code="9">دیوان</subfield></datafield><datafield tag="606" ind1="1" ind2="
"><subfield code="9">شعر فارسی - پاکستان - کشمیر - قرن ۱۳</subfield></datafield><datafield tag="680" ind1=" " ind2=" "><subfield
code="a">PIR۹۲۴۹</subfield><subfield code="b">/ ۹د
۹د</subfield></datafield><datafield tag="700" ind1=" " ind2=" "><subfield
code="9">۱۹۰۰ - ۱۸۲۸</subfield></datafield><datafield tag="700" ind1=" " ind2="1"><subfield code="9">Dairi, Pandit Raja Koul
Arzbegi</subfield></datafield><datafield tag="702" ind1=" " ind2=" "><subfield
code="9">طالب کشمیری، نند لعل‌کول</subfield></datafield><datafield tag="801" ind1=" " ind2="0"><subfield
code="a">IR</subfield><subfield code="b">کتابخانه مجلس شورای
کتابخانه</subfield></datafield><datafield tag="856" ind1=" " ind2=" "><subfield
code="f">/</datafield><datafield tag="910" ind1=" " ind2=" "><subfield
code="f">۰۹/۸۴</subfield></datafield><datafield tag="915" ind1=" " ind2="
"><subfield code="a">فهرست نویسی قبلی</subfield></datafield><datafield tag="930" ind1="
" ind2=" "><subfield code="a">LF</subfield><subfield
code="c">1</subfield></datafield><datafield tag="930" ind1=" " ind2=" "><subfield
code="a">SF</subfield><subfield code="c">1</subfield></datafield><datafield
tag="932" ind1=" " ind2=" "><subfield code="b">Y</subfield></datafield><datafield
tag="999" ind1=" " ind2=" "><subfield code="r">۲۳۲۹۵۴</subfield><subfield
code="s">\</subfield></datafield><datafield tag="856" ind1=" " ind2=" "><subfield
code="e">24-01372.pdf</subfield></datafield></record></collection>

```