

Semantic Integration in Iranian Digital Libraries: a Proposing Model

Mehdi Alipour-Hafezi

PhD in Knowledge and Information Science; Department of Knowledge and Information Science; Faculty of Psychology and Education Science; Allame Tabataba'i University (ATU); Tehran, Iran meh.hafezi@gmail.com

Received: 16, Apr. 2014

Accepted: 24, May 2015

Abstract: The main purpose of this study is presenting a conceptual model to integrate Iranian digital libraries. This research has been done in three steps. Documentary research method is used in the first step. All the related documents were research population and content analysis techniques is used to analyze documents. Analytical survey research method is used in the second step of the research. All the active Iranian digital libraries were the research population in this step. In fact there are 32 active digital libraries that just 26 of them participated in this study. Researcher-made questionnaire is used to collect data in this step. Finally system analyzing research method is used in the third step.

Findings of the study demonstrated that elements such as technological infrastructure, Coding and identifier systems, Metadata standards, Legal and organizational issues, Semantic knowledge-based databases, and Knowledge Organizing Systems are needed in semantic integration. Also methods such as translators and making semantic relations can be used in semantic integration. Moreover tools such as ontology, knowledge organizing system, knowledgebase databases, translators, and descriptive and analytical languages should be used in semantic integration. Other findings related to Question 2 of the study showed that the studied digital libraries are not in suitable situation in covering digital materials and filling metadata fields. Also documenting metadata fields is not in good mood. Therefore, it is needed to strengthen all the identified weaknesses in order to make semantic integration ability. Finally, proposed model of this study presented in three layers entitled data, inference machine, and application.

Keywords: Digital Library, Semantic Integration, Interoperability, Semantic Interchange, Proposed Conceptual Model, Iran

Iranian Journal of
**Information
Processing and
Management**

Iranian Research Institute
for Information Science and Technology
(IranDoc)

ISSN 2251-8223

eISSN 2251-8231

Indexed by SCOPUS, ISC, & LISTA

Vol. 31 | No. 2 | pp. 455-482

Winter 2016

<https://doi.org/10.35050/JIPM010.2016.042>



یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتالی در ایران: مدل مفهومی پیشنهادی

مهدی علیپور حافظی

دکتری علم اطلاعات و دانش‌شناسی؛
استاد مدعو؛ دانشگاه علامه طباطبائی؛
meh.hafezi@gmail.com



دریافت: ۱۳۹۳/۰۱/۲۷ | پذیرش: ۱۳۹۴/۰۳/۰۳ | مقاله برای اصلاح به مدت ۳۳ روز نزد پدیدآور بوده است.

چکیده: هدف اصلی پژوهش حاضر ارائه مدل مفهومی پیشنهادی یکپارچه‌سازی معنایی کتابخانه‌های دیجیتالی در ایران است. این پژوهش در سه گام مستقل به اجرا درآمده است. در گام نخست، از روش سندی برای شناسایی منابع مرتبط، تولید پرسشنامه پژوهشگر ساخته و پاسخ به پرسش اساسی اول پژوهش استفاده شده است. جامعه پژوهش را تمامی منابع اطلاعاتی مرتبط در دسترس تشکیل داده و از فنون تحلیل محتوا برای تحلیل منابع استفاده شده است. در گام دوم، از پیمایش تحلیلی برای بررسی وضعیت موجود کتابخانه‌های دیجیتالی استفاده شده است. جامعه پژوهش را تمامی کتابخانه‌های دیجیتالی فعال به تعداد ۳۲ کتابخانه تشکیل می‌دهد که از این تعداد ۲۶ کتابخانه در پژوهش حاضر مشارکت داشته‌اند. ابزار گردآوری اطلاعات نیز در این بخش پرسشنامه پژوهشگر ساخته است. در گام سوم پژوهش نیز از روش تحلیل سیستم استفاده شده است.

یافته‌های پژوهش نشان داد که در یکپارچه‌سازی معنایی به عناصری مانند زیرساخت فناوریانه، نظام‌های کدگذاری و شناساگر، استانداردهای فراداده، مسائل قانونی، سازمانی و پایگاه‌های دانشی معنایی و سیستم‌های سازماندهی دانش نیاز هست. همچنین، روش استفاده از مترجم‌ها و برقراری روابط معنایی می‌تواند در یکپارچه‌سازی معنایی مورد استفاده قرار گیرند. از نظر ابزارهای مهم مورد نیاز در یکپارچه‌سازی معنایی نیز می‌توان از ابزارهایی مانند هستی‌نگاری‌ها، سیستم‌های سازماندهی دانش، پایگاه دانشی، مترجم‌ها، و زبان‌های توصیفی و تحلیلی استفاده کرد. همچنین، یافته‌های دیگر پژوهش در ارتباط با پرسش اساسی دوم پژوهش نشان داد که کتابخانه‌های دیجیتالی مورد مطالعه از نظر پوشش منابع دیجیتالی و تکمیل فیلدهای فراداده در وضعیت مطلوبی قرار ندارند. همچنین، از نظر مستندسازی محتوای فیلدهای فراداده نیز وضعیت چندان رضایت‌بخشی را شاهد نیستیم. بنابراین، برای ایجاد قابلیت

فصلنامه | علمی پژوهشی

پژوهشگاه علوم و فناوری اطلاعات ایران
(ایرانداک)

شاپا (چاپی) ۲۲۵۱-۸۲۳۳

شاپا (الکترونیکی) ۲۲۵۱-۸۲۳۱

نمایه در SCOPUS و ISI، LISTA، و

jipm.irandoc.ac.ir

دوره ۳۱ | شماره ۲ | صص ۴۵۵-۴۸۱

زمستان ۱۳۹۴

<https://doi.org/10.35050/JIPM010.2016.042>



یکپارچه‌سازی معنایی نیازمند تقویت تمامی نقاط ضعف شناسایی شده در این پژوهش هستیم. در نهایت، مدل مفهومی پیشنهادی در سه لایه داده، استنتاج ماشینی، و برنامه کاربردی پیشنهاد شده است.

کلیدواژه‌ها: کتابخانه‌های دیجیتال، یکپارچه‌سازی معنایی، مدل مفهومی پیشنهادی

۱. مقدمه

حجم بالای اطلاعات و عدم توانایی موتورهای جست‌وجو در پوشش کامل محتوای وب سطحی و عمیق و نیز جامعیت بالای بازیابی موتورهای جست‌وجو در مقابل مانعیت بسیار پایین آنها باعث شده که شکاف عظیمی میان نیازهای اطلاعاتی افراد و دستیابی به اطلاعات و منابع اطلاعاتی مرتبط ایجاد شود. همین امر باعث شده که مسیر توسعه وب به سمت وب معنایی جهت‌گیری نموده و افزایش مانعیت در نتایج بازیابی موتورهای جست‌وجو و سایر ابزارهای بازیابی جزو اهداف مهم این سیستم‌های اطلاعاتی باشد. توجه به معنا در بازیابی اطلاعات، با توجه به گسترش ورودی‌های جست‌وجو، علاوه بر توجه ویژه به مانعیت، توجه ویژه‌ای نیز بر جامعیت در راستای نیاز اطلاعات دارد. لذا، سیستم‌های اطلاعاتی کتابخانه دیجیتال به اهمیت آن واقف شده‌اند و اکثر سیستم‌های اطلاعاتی مرتبط در راستای توجه به نظام‌های معنایی هستند. بنابراین، یکپارچه‌سازی دستیابی به اطلاعات، امروزه جزء ضروری‌ترین نیازهای انسانی در دنیای فناوری 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 تا بتوانند پاسخ‌گوی نیازهای اطلاعاتی کاربران 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 سامانه‌های اطلاعاتی به‌ویژه سامانه‌های اطلاعاتی کتابخانه دیجیتال نیازمند یکپارچه‌سازی در سطح نحوی^۱ و معنایی^۲ است (Ding and Sølvberg 2004; Shen 2006). در یکپارچه‌سازی نحوی که زیربنای یکپارچه‌سازی سیستم‌های اطلاعاتی نیز قلمداد می‌شود، به لحاظ عملیاتی، سیستم‌های اطلاعاتی با یکدیگر مرتبط می‌شوند و به مبادله اطلاعات با یکدیگر اقدام

1. syntactic

2. semantic

می‌نمایند. در این سطح از یکپارچه‌سازی اطلاعات سه سطح توافقنامه محتوا، فنی، و سازمانی مطرح هستند. توافقنامه فنی شامل فرمت‌ها، پروتکل‌ها، و سیستم‌های امنیتی است. به این ترتیب، محتواهای اطلاعاتی می‌توانند مبادله شوند. توافقنامه محتوا شامل داده و فراداده است که دربرگیرنده توافقنامه‌های معنایی^۱ در تفسیر پیام‌هاست. توافقنامه سازمانی دربرگیرنده قواعد زیربنایی برای دسترسی، جایگزینی مجموعه‌ها و خدمات، پرداخت‌ها، شناسایی و غیره است (NSDL technical infrastructure white paper 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).

بحث بعدی در یکپارچه‌سازی محتوا و خدمات در کتابخانه‌های دیجیتالی 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 یکپارچه‌سازی، تسهیل بازیابی معنایی اطلاعات در سامانه‌های اطلاعاتی کتابخانه‌های دیجیتالی است تا کاربران به‌سهولت بتوانند با هر کلیدواژه جست‌وجو، 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 مختلف تهیه و ذخیره شده‌اند. همین امر باعث عدم یکدستی در محتواهای اطلاعاتی مبادله شده می‌شود. از آنجا 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. بنابراین، مسئله اصلی در پژوهش حاضر شناسایی مدل مفهومی مناسب در یکپارچه‌سازی کتابخانه‌های دیجیتالی با توجه به وضعیت موجود سامانه‌های اطلاعاتی آنهاست. در این راستا، با توجه به اهمیت استانداردهای مرتبط موجود در این زمینه سعی شده است که مدل مفهومی پیشنهادی به استانداردهای مناسب این حوزه و قابلیت کاربرد آنها در تعامل معنایی توجه داشته و

1. semantic agreements

مبتنی بر آن عمل نماید.

در این راستا، هدف اصلی پژوهش حاضر ارائه مدل مفهومی پیشنهادی یکپارچه‌سازی معنایی کتابخانه‌های دیجیتال در ایران است. برای دستیابی به هدف اصلی مذکور اهداف فرعی زیر نیز تعیین شدند:

- ◇ شناسایی عناصر دخیل در یکپارچه‌سازی معنایی کتابخانه‌های دیجیتال
- ◇ شناسایی وضعیت کتابخانه‌های دیجیتال مورد بررسی از نظر مستندسازی محتوای فراداده‌ها جهت استفاده در یکپارچه‌سازی معنایی اطلاعات^۱

با توجه به مطالب اشاره شده، پژوهش حاضر بدین شکل سازمان یافته است. در بخش دوم، به معرفی تحلیلی پیشینه‌های پژوهش پرداخته شده است. در ادامه و در بخش سوم، به بیان روش انجام پژوهش جهت وصول به اهداف ذکر شده در بالا پرداخته شده است. سپس، یافته‌ها بر مبنای پرسش‌های اساسی پژوهش ارائه گردیده و در بخش پایانی نیز نسبت به تحلیل و بحث در زمینه نتایج پژوهش و پیشنهاد مدل مفهومی یکپارچه‌سازی معنایی اقدام شده است.

۲. پیشینه پژوهش

مرور پیشینه‌های داخلی حاکی از چهار پژوهش و چهار طرح اجرایی در زمینه محور اصلی پژوهش حاضر است. یکی از پژوهش‌ها اختصاص به بررسی وضعیت مبادله اطلاعات در محیط کتابخانه‌های سنتی دارد (محقق ۱۳۷۴). دو پژوهش دیگر با مبادله اطلاعات در بستر نحوی در کتابخانه‌های دیجیتال مرتبط می‌باشند (علیپور حافظی ۱۳۸۸؛ ۱۳۹۰) و پژوهش دیگر مرتبط با کارآمدی هستی‌نگاری‌ها نسبت به اصطلاح‌نامه‌ها در بازیابی اطلاعات در سیستم‌های اطلاعاتی نوین است (صنعت‌جو و فتحیان ۱۳۹۰). از طرفی نیز از بین طرح‌های اجرا شده، دو طرح/امین و غدیر با اشتراک منابع بین کتابخانه‌ها و تسهیل در دستیابی به کتابخانه‌ها در بستر کتابخانه‌های سنتی مرتبط هستند. یک طرح مرتبط با گردآوری اطلاعات کتابشناختی در پایگاه اطلاعاتی واحد و ایجاد امکان دستیابی به آن برای کاربران است. متأسفانه پایگاه اطلاعاتی مذکور با توجه به روزآمد نبودن محتواها و عدم جامعیت، از کارآمدی مناسبی برخوردار نیست. طرح دیگر اجرا شده در یکپارچه‌سازی نیز با کنسرسیوم محتوای ملی مرتبط است. کنسرسیوم مذکور به علت نبود قابلیت

۱. برای وصول به اهداف بالا، پرسش‌هایی با همین مضمون تهیه شدند که جهت پیشگیری از تکرار مجدد محتوای آنها از بیان مجدد صرف نظر شد. پرسش‌های مذکور در بخش یافته‌های پژوهش ذکر شده‌اند.

نرم‌افزاری در پیاده‌سازی برداشت خودکار، از مدل برداشت اطلاعات به صورت سنتی و نیمه خودکار استفاده می‌کند. همین امر باعث شده که محتواهای موجود در سیستم اطلاعاتی واسط روزآمد نبوده و توانایی پاسخ‌گویی به نیازهای به‌روزشده کاربران را نداشته باشد. بنابراین، کنسرسیوم محتوای ملی به علت عدم جامعیت و عدم پوشش محتواهای اطلاعات نرم‌افزارهای کتابخانه‌ای هنوز نتوانسته در بستر نحوی پوشش مناسبی بر محتواهای کتابخانه‌ای بدهد. از طرفی نیز هنوز فعالیت جدی در راستای دستیابی معنایی به اطلاعات کتابخانه‌های دیجیتال عضو و نیز دستیابی به محتواهای اطلاعاتی، علاوه بر فراداده، انجام نگرفته است. با این وجود، با توجه به قابلیت و استعداد بالقوه موجود در این کنسرسیوم، همان‌گونه که در ادامه در مقاله حاضر بر اساس نتایج پژوهش استنباط شده، این سیستم توانایی این را می‌تواند داشته باشد که مدل پیشنهادی در این پژوهش را پیاده‌سازی نموده و بهره‌برداری از آن را برای کاربران ایرانی تسهیل نماید. علاوه بر موارد ذکر شده، یکپارچه‌سازی‌های نحوی دیگری نیز توسط نرم‌افزارهای کتابخانه‌ای شکل گرفته است. در این موارد یکپارچه‌سازی در نرم‌افزارهای متجانس شکل گرفته است. فعالیت‌های مذکور نیز از جمله فعالیت‌های مناسب در یکپارچه‌سازی نحوی هستند. مهم‌ترین مسئله‌ای که در این زمینه وجود دارد متجانس بودن سیستم‌های اطلاعاتی است و راهکار اجرایی مناسبی برای یکپارچه‌سازی نحوی در سیستم‌های اطلاعاتی نامتجانس وجود ندارد. 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 تاکنون ۱۲ پروژه در این زمینه به اجرا درآمده است. برخی از این پروژه‌ها به تولید ابزارهای مورد نیاز مانند «وُردنیت» (Szymański 2011) یا هستی‌نگاری‌ها (Chen, Finin, and Joshi 2003) در یکپارچه‌سازی معنایی اختصاص یافته‌اند. برخی دیگر از این پروژه‌ها نیز زیرمجموعه‌هایی از

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 کتابخانه سیمغ، پارس آذرخش و نرم‌افزار کتابخانه دیجیتال پیام مشرق اشاره کرد.

پروژه‌های بزرگ‌تر بوده و با یکپارچه‌سازی کتابخانه‌ها و یکپارچه‌سازی معنایی مرتبط بودند. از جمله این پروژه‌ها می‌توان به پروژه «تله‌متیکس» (Creating a European library space telematics) (for libraries programmes 2000) و «دِلاس» (DELOS 2004) اشاره کرد. 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 شناسایی 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 «دِلاس» که از سال ۲۰۰۴ آغاز شده، زیرپروژه‌های 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 داشتند (Warren and Alsmeyer 2005; Guha 2006; Issac, Schlobach, Mattheizing, and Zinn 2008)، برخی مبتنی بر فناوری خاصی مانند آردی‌اف بودند (Han 2006)، برخی به معرفی مدل خاصی در حوزه بازیابی معنایی اختصاص یافته بودند (Huang, Ke, and Yang 2005; Pasad and Madalli 2008)، و برخی نیز به روش‌های یکپارچه‌سازی پرداخته بودند (Mayer, Mutschke, and Petras 2008).

به‌طور کلی، با بررسی پژوهش‌ها و پروژه‌های اجرا شده در حوزه یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال مشخص می‌شود که تاکنون پژوهش و یا پروژه خاصی به یکپارچه‌سازی معنایی در داخل کشور پرداخته است. همچنین، در پژوهش‌ها و پروژه‌های خارجی نیز تنها یک مورد، پژوهش (Mayer, Mutschke, and Petras 2008)، 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 فضایی بکر در حوزه یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال توزیع یافته دارد و پژوهش حاضر سعی به پرداختن به این حوزه در کتابخانه‌های دیجیتال توزیع یافته در ایران دارد.

۳. روش پژوهش

۳-۱. گام نخست

در بخش اول پژوهش برای شناسایی منابع مرتبط و وصول به هدف فرعی نخست پژوهش از روش سندی استفاده شده است. برای گردآوری اطلاعات اقدام به جست‌وجو در پایگاه‌های

اطلاعاتی مهم و کاربردی مانند امرالد^۱، الزویر^۲، آی‌ای‌ای^۳، لیستا^۴، ایستا^۵، و غیره انجام گرفت. سپس با بررسی و مطالعهٔ چکیده‌های منابع اطلاعاتی بازیابی شده اقدام به انتخاب منابع مرتبط در گام اول گردید. سپس، در مرحلهٔ بعدی منابع انتخاب شده به‌طور کامل مورد بررسی و مطالعه قرار گرفت که از میان آنها منابع اطلاعاتی کاملاً مرتبط انتخاب شدند و در طول فرایند انجام پژوهش به‌ویژه در بخش مبانی نظری و ادبیات پژوهش مورد استفاده قرار گرفتند. جامعهٔ پژوهش در این گام منابع اطلاعاتی مرتبطی بودند که به‌جهت آشنایی با ادبیات پژوهش مورد استفادهٔ پژوهشگر بودند. ابزار مورد استفاده نیز برای گردآوری اطلاعات موتورهای جست‌وجو و جست‌وجو توسط آنها بود. در تحلیل اطلاعات نیز از فنون تحلیل محتوای منابع مرتبط بازیابی شده استفاده شده است.

۳-۲. گام دوم

در بخش دوم پژوهش از روش پیمایش تحلیلی استفاده شده است. برای گردآوری اطلاعات از ابزار پرسشنامه، مصاحبه و مشاهدهٔ هدفمند استفاده شد. با توجه به اینکه تاکنون پژوهشی در این سطح در کشور انجام نگرفته، و نیز اینکه ابزار مشابهی برای این منظور موجود نبود، پژوهشگر با استفاده از منابع اطلاعاتی، پژوهش‌ها و نمونه‌های موجود و نظر متخصصان این حوزه اقدام به تهیهٔ پرسشنامهٔ پژوهشگر ساخته کرده است. برای کسب روایی ابزار پژوهش، پرسشنامهٔ تهیه شده بین پنج نفر از متخصصان حوزهٔ علوم کتابداری و اطلاع‌رسانی آشنا به حوزهٔ فناوری اطلاعات توزیع شد. پس از دریافت نظرات متخصصان و اعمال اصلاحات پیشنهادی در پرسشنامهٔ اولیه، ویرایش جدیدی از پرسشنامه حاصل گردید. پرسشنامهٔ اصلاح شده در مرحلهٔ دوم در پنج کتابخانهٔ دیجیتالی که حائز تمام ویژگی‌های جامعهٔ پژوهش بودند، نیز توزیع گردید و از آنها خواسته شد تا ضمن پاسخ‌گویی به پرسش‌ها موارد مبهم و یا ناقص را مشخص نمایند. با دریافت پرسشنامه‌ها و اعمال اصلاحات پیشنهادی، ویرایشی جدید از پرسشنامه حاصل شد که دارای روایی مناسبی بود. پس از دریافت نظرات متخصصان و نمونه‌های پاسخ داده‌شده، اصلاحات نهایی روی پرسشنامه انجام گرفت و پرسشنامهٔ نهایی تهیه گردید. 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.

پرسشنامهٔ نهایی شده برای گردآوری اطلاعات در اختیار جامعهٔ پژوهش قرار داده شد و پس از تکمیل در مراحل بعدی پژوهش جهت توصیف و تحلیل مورد استفاده قرار گرفت. برای کسب

1. Emerald
4. LISTA

2. Elsevier
5. ISTA

3. IEEE

اطمینان از یافته‌های پژوهش، پژوهشگر سعی نموده است که تا حد امکان با مراجعه به محل کتابخانه‌های دیجیتال مورد بررسی با استفاده از ابزار مشاهده مستقیم و مصاحبه هدفمند اقدام به تصحیح پاسخ‌ها نماید. بنابراین، در این فرایند برخی از داده‌های گردآوری شده با توجه به مشاهده صورت گرفته در این مرحله اصلاح شدند. این فرایند باعث شد که یافته‌های پژوهش از اطمینان بیشتری برخوردار باشند.

جامعه پژوهش حاضر را در این گام کتابخانه‌های دیجیتال فعال تشکیل می‌دادند. 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 تهیه فهرست ۱۱ نرم‌افزار زیر شد:^۱ نرم‌افزار کتابخانه دیجیتال آذرخش^۲، نرم‌افزار کتابخانه دیجیتال ارم^۳، نرم‌افزار کتابخانه دیجیتال پایروس^۴، نرم‌افزار کتابخانه دیجیتال پروان^۵، نرم‌افزار کتابخانه دیجیتال پیام^۶، نرم‌افزار کتابخانه دیجیتال تیان^۷، نرم‌افزار کتابخانه دیجیتال دید^۸، نرم‌افزار کتابخانه دیجیتال سنا^۹، نرم‌افزار کتابخانه دیجیتال سیمرغ^{۱۰}، نرم‌افزار کتابخانه دیجیتال نورلیب^{۱۱}، و نرم‌افزار کتابخانه دیجیتال وستا^{۱۲}.

بررسی اولیه و بگام 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، تیان، سنا، ارم و وستا را نام برد. سایر نرم‌افزارها مانند سیمرغ، آذرخش، پیام، پروان و پایروس 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^{۱۳} استفاده کرده‌اند. به طور کلی، بررسی در این زمینه منتج به تهیه فهرستی از کتابخانه‌های دیجیتال به تعداد ۴۶ شد. در ادامه، اقدام به بررسی اولیه ۴۶ کتابخانه دیجیتال شد.

۱. فهرست به ترتیب الفبا مرتب شده است.

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

3. <http://www.dlib.ir/fa/index.aspx>

۴. فاقد سایت قابل بازدید

5. <http://www.parvanpajooch.ir>

6. <http://www.payamnet.com>

7. <http://www.tebyan.net/Library.html>

8. www.did.ir/

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

10. <http://www.nosa.com/nosasoftware/webUI/Home.aspx>

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

12. <http://www.vestasoft.ir/CyberLib.aspx>

13. open source

بررسی اولیه نشان داد که ۱۴ مورد از ۴۶ کتابخانه در تعریف کتابخانه دیجیتالی پژوهش حاضر نمی‌گنجند، یا در حال حاضر کتابخانه دیجیتالی نیستند و یا در حال نصب و راه‌اندازی کتابخانه دیجیتالی هستند. از این رو، جامعه پژوهش حاضر به ۳۲ کتابخانه دیجیتالی محدود شد. سپس، پرسشنامه پژوهش به کتابخانه‌های دیجیتالی مورد نظر ارسال شد و حسب مورد مشاوره‌هایی نسبت به تکمیل محتوای پرسشنامه داده شد. در انتها ۲۶ پرسشنامه تکمیل شد. با توجه به اینکه موفق به گردآوری اطلاعات مربوط به ۸۱/۲۵ درصد از جامعه پژوهش شدیم، بر همین اساس توصیف و تحلیل‌ها را بر اساس ۲۶ پرسشنامه گردآوری شده انجام دادیم.

۳-۳. گام سوم

در بخش سوم پژوهش از روش تحلیل سیستم استفاده شده است. پس از تحلیل یافته‌های پژوهش و نیز با استفاده از یافته‌های پژوهش‌های پیشین اقدام به طراحی مدل مفهومی پیشنهادی یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتالی شده است. لذا، با استفاده از یافته‌های این پژوهش و نیز با بهره‌گرفتن از تجربیات پژوهش‌های پیشین در این بخش سعی شده است که مدل مفهومی مناسب با وضعیت موجود کتابخانه‌های دیجیتالی ایران تهیه و ارائه شود. از این رو، برای بهره‌گیری و آزمون کاربردی بودن مدل مفهومی پیشنهادی، مدل مذکور به جهت اعتبارسنجی مورد آزمون قرار گرفته است.

۴. یافته‌های پژوهش

در این بخش از مقاله یافته‌های پژوهش ارائه شده است. این یافته‌ها بر اساس پرسشنامه ذکر شده به دست آمده است. ساختار این بخش از مقاله بر اساس اهداف فرعی ذکر شده در مقدمه مقاله می‌باشد. بنابراین، ابتدا هدف مورد نظر در قالب پرسش اساسی ذکر شده و در ادامه یافته‌های مرتبط با آن ارائه گردیده است.

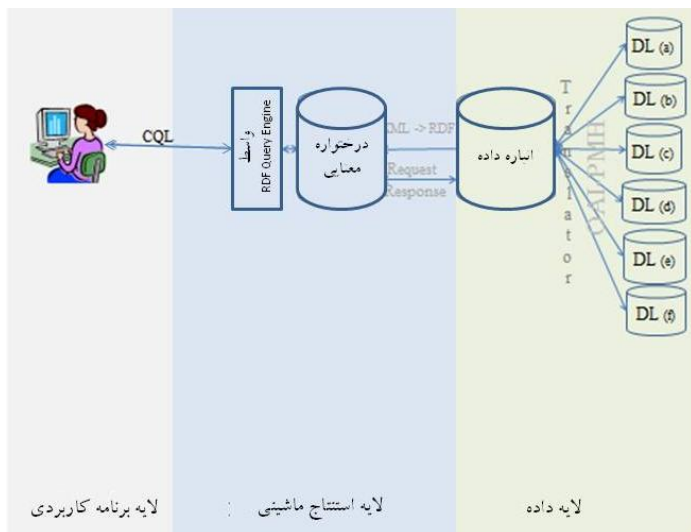
۴-۱. چه عناصری در یکپارچه‌سازی معنایی کتابخانه‌های دیجیتالی دخیل هستند؟

برای پاسخ به این پرسش، داده‌ها از منابع دانشی موجود گردآوری 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، راه‌اندازی و مورد استفاده قرار گرفته بودند. با توجه به یافته‌های ذکر شده، عناصر دخیل در یکپارچه‌سازی معنایی با رویکرد (Shen 2006) و متجانس یا نامتجانس بودن سیستم‌های اطلاعاتی حاضر در شبکه یکپارچه‌سازی معنایی بستگی 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 که مجموعه‌ی واژگان و روابط معنایی بین آنها با استفاده از یک اصطلاح‌نامه یا هستی‌نگاری و یا مجموعه‌ای از آنها (برای پوشش کامل موضوعی) برقرار گردد. در سیستم‌های اطلاعاتی نامتجانس (که بیشترین کاربرد را در یکپارچه‌سازی سامانه‌های اطلاعاتی دارد) برای ایجاد یکپارچه‌سازی معنایی، به زیرساخت فناوریانه مانند پروتکل‌های مبادله‌ی اطلاعات (Alipour-Hafezi, Horri, Shiri, and 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) نیاز هست (Patel, Koch, Doerr, and Tsinaraki 2005).

از طرفی نیز یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال با توجه به پژوهش‌ها و پروژه‌های پیشین به دو روش اصلی انجام می‌گیرد. روش اول عبارت است از استفاده از مترجم‌ها برای ترجمه‌ی استانداردهای فراداده به استاندارد مورد نظر در سیستم و ماشین استنتاج^۱ (Vdovjak, 2006; Guha 2006; Shen 2006; Patel, Koch, Doerr, and Tsinaraki 2005; and Houben 2001). 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 که فراداده‌ها با استفاده از استانداردهای مختلف به استاندارد مورد نظر در سیستم تبدیل شوند. پس از ایجاد پایگاه اطلاعاتی واسط و ترسیم فرایند 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 پرس‌وجوهای کاربران می‌شود. در شکل ۱ روش انجام چنین یکپارچه‌سازی نشان داده شده است.

1. inference machine



شکل ۱. روش استفاده از مترجم‌ها و ماشین استنتاج

شکل ۱ نشانگر این است که لایه داده^۱ مربوط به یکپارچه‌سازی نحوی می‌شود. در این لایه فراداده‌ها با استفاده از تفاهم‌نامه آی‌آی^۲ گردآوری شده و در انباره داده^۳ نگهداری می‌شود. در این لایه کتابخانه‌های دیجیتال عضو شبکه یکپارچه کتابخانه‌ای امکان مرور پایگاه فراداده خود را به خزشگر^۴ مورد استفاده در انباره داده می‌دهند (علی‌پور حافظی ۱۳۹۰). در مسیر انتقال فراداده‌ها، مترجمی وجود دارد که فراداده‌های مورد استفاده در کتابخانه‌های دیجیتال را به استاندارد مورد نیاز در انباره داده ترجمه می‌کند.

در گام بعدی داده‌های معنایی مستخرج از فیله‌های فراداده با ساختار سلسله‌مراتبی موجود در انباره معنایی^۵ مطابقت داده شده و به انضمام روابط سلسله‌مراتبی معنایی در انباره مذکور نگهداری می‌شود. جست‌وجوی کاربر در این انباره شامل فراداده‌های کتابخانه‌های دیجیتالی عضو شبکه یکپارچه کتابخانه‌ای و روابط معنایی محتواهای فراداده‌ها برای کاربر بازیابی و نمایش داده می‌شود. از جمله پروژه‌هایی که بر این مبنا استوار شده‌اند می‌توان به پروژه DELOS (2004)، پروژه «تی‌اس‌آی‌ام‌آی‌اس» (Chawathe, et al. [1996])، و پروژه «یورپانا» (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

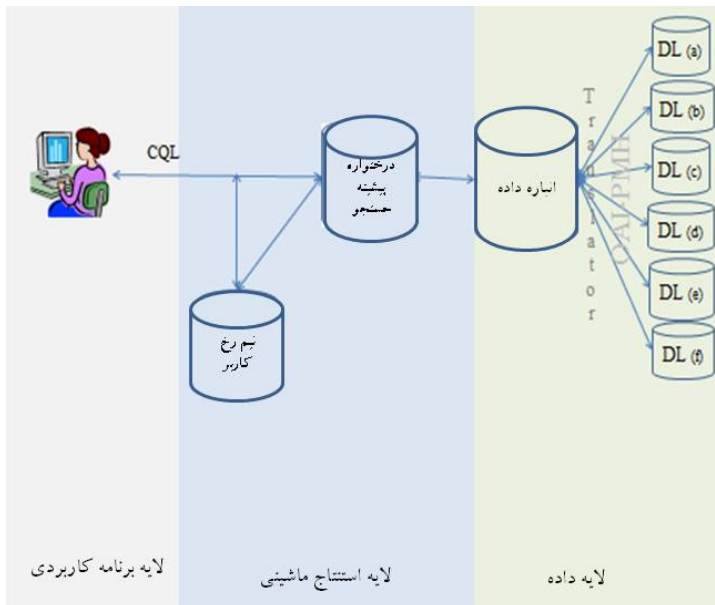
1. data layer
4. crawler

2. OAI (Open Achieve Initiatives)
5. semantic store

3. data warehouse

داده‌ها رسیده است. از این رو، روش پیشنهادی می‌تواند مشابه هم باشد. همچنین، Huang, Ke, (2005) and Yang (2008) و Issac, Schlobach, Mattheizing, and Zinn (2008) و Pasad and Madalli (2008) نیز به استفاده از روش مشابهی رهنمون شده‌اند.

در روش دوم که معماری آن در شکل ۲ نشان داده شده، برقراری روابط معنایی برعکس روش پیشین مبتنی بر معماری پس‌همارایی است. تأکید اصلی در این روش بر جست‌وجوی 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 جست‌وجوی کاربر وجود دارد و روابط معنایی میان کلیدواژه‌های جست‌وجو و منابع بازیابی شده در آن برقرار و ثبت می‌شود. هنگام جست‌وجوی کاربر در این سیستم، جست‌وجوی او با روابط ثبت شده در پایگاه اطلاعاتی مذکور برقرار شده و بر اساس وزن حاصل از حجم رابطه بین کلیدواژه‌ها و فراداده‌ها یافته‌های معنایی مرتبط بازیابی شده و به کاربر نمایش داده می‌شود.



شکل ۲. معماری مبتنی بر ثبت پیشینه جست‌وجو

پروژه «ان‌اس‌دی‌ال» (Lagoze, et al. 2002) از جمله پروژه‌هایی است که مبتنی بر پروفایل کاربر بوده و 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 دیجیتالی نبودند، در اینجا به این نمونه‌ها اشاره‌ای نمی‌شود.

برای اجرای چنین روش‌هایی نیاز به استفاده از ابزارهایی است که در معماری‌های بالا نیز به آنها اشاره شد. در این زمینه بیان شد که پیش‌نیازهای زیادی از جمله زیرساخت فناوریانه (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, and Tsinaraki 2005)، برای یکپارچه‌سازی معنایی و محقق‌ساختن 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 «وُردنیت» اشاره کرد.
- ◇ سیستم‌های سازماندهی دانش. همان‌گونه که پیش‌تر ذکر شده، سیستم‌های سازماندهی دانش به واژگان کنترل‌شده و نیز سیستم‌ها، ابزارها یا خدماتی اشاره دارد که به‌منظور سازماندهی دانش ایجاد شده‌اند (ابزارهایی که تفسیر سازماندهی شده‌ای از ساختار دانش ارائه می‌نمایند) (Patel, Koch, Doerr, and Tsinaraki 2005). از جمله این سیستم‌ها می‌توان به «اسکاس»^۱ اشاره کرد.
- ◇ پایگاه دانشی. این پایگاه شامل واژگان و روابط معنایی بین آنها بوده و در حقیقت دربرگیرنده هر دو ابزار ذکر شده در بالاست. این پایگاه همواره در نظام‌های یکپارچه‌اطلاعاتی در لایه استنتاج قرار می‌گیرد و مغز متفکر آن را تشکیل می‌دهد. پایگاه دانشی 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 را بر عهده دارند. تمامی مترجم‌ها مبتنی بر نقشه‌هایی هستند که از قبل برای استانداردها و یا حتی فراداده‌های غیراستاندارد تهیه و در

1. Simple Knowledge Organization System (SKOS)

سیستم تعبیه می‌شوند. نوع فعالیت این مترجم‌ها بر اساس استانداردسازی منفعل (Shen 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 «ایکس‌ام‌ال» است. امروزه بدون آن امکان انتقال داده‌ها امکان‌پذیر نیست. بنابراین، می‌توان اظهار داشت که زبان توصیفی «ایکس‌ام‌ال» جزء لازم و ضروری انتقال داده‌هاست. علاوه بر این، برای ایجاد شبکه‌های معنایی و هستی‌نگاری‌ها نیاز به استفاده از زبان‌های تحلیلی مانند آردی‌اف^۱، دمل+اوایل^۲ و یا ادبلیوال^۳ است. این زبان‌ها قابلیت برقراری روابط معنایی را در سیستم‌های دانشی بر عهده دارند.

۴-۲. وضعیت کتابخانه‌های دیجیتالی مورد بررسی از نظر مستندسازی محتوای فراداده‌ها جهت استفاده در یکپارچه‌سازی معنایی اطلاعات چگونه است؟

در این بخش و در راستای پاسخ به پرسش مذکور و همچنین پیشگیری از اطاله کلام، یافته‌های مرتبط با این پرسش به‌طور خلاصه ارائه شده است. بنابراین، پس از ارائه خلاصه‌ای از یافته‌های پژوهش اقدام به تحلیل و پاسخ‌گویی به این پرسش شده است.

◇ یافته‌ها در این زمینه حاکی از تأکید ویژه کتابخانه‌های دیجیتال مورد بررسی بر منابع متنی (کتاب ۱۰۰ درصد، پایان‌نامه ۶۲ درصد، و مقالات ۵۸ درصد) است.

◇ در کتابخانه‌های دیجیتال مورد بررسی برای توصیف منابع اطلاعاتی از فراداده‌هایی که شامل فیلدهای اطلاعاتی متعددی هستند، بهره‌برداری می‌شود و از بین فیلدهای اطلاعاتی مورد استفاده محتوای فیلدهای عنوان (۱۰۰ درصد)، پدیدآور (۹۶ درصد)، پدیدآور همکار (۹۲ درصد)، موضوع (۹۲ درصد)، ناشر (۹۶ درصد)، و شابک برای کتاب‌ها (۷۷ درصد) بیشتر تکمیل می‌گردند.

◇ در زمینه مستندسازی اسامی پدیدآورندگان، یافته‌ها حاکی از این است که از فهرست مستند اسامی مشاهیر و مؤلفان (۶۲ درصد) و پایگاه اطلاعاتی کتابخانه ملی (۷۷ درصد) و برای مستندسازی نام ناشران بیشتر از قسمت شناسنامه اثر در دست بررسی (۲۷ درصد) استفاده می‌شود.

◇ در تحلیل موضوعی منابع اطلاعاتی بیشتر از موضوع‌دهی کنترل‌شده با استفاده از سرعنوان‌های موضوعی (۸۵ درصد) و نمایه‌سازی کنترل‌شده با استفاده از اصطلاح‌نامه‌ها (۵۴ درصد)

بهره‌برداری می‌شود. همچنین، موضوع‌دهی (۹۲ درصد) در بیشتر موارد در مورد کتاب‌ها انجام می‌گیرد.

◇ در توصیف موضوعی با استفاده از کلیدواژه‌ها، پایان‌نامه‌ها (۵۰ درصد)، کتاب‌ها و مقالات (هر کدام با ۲۱ درصد) بیشترین بهره را برده‌اند.

◇ در بیشتر موارد از فهرست‌نویسی به شیوهٔ نسخه‌برداری (۸۵ درصد) استفاده می‌شود. جهت نسخه‌برداری کتابشناسی ملی (۸۸ درصد) بیشترین کاربرد را دارد. برای منابع انگلیسی‌زبان از پایگاه اطلاعاتی کتابخانهٔ کنگره آمریکا (۸۸ درصد) استفاده می‌شود. در اکثر موارد نسبت به اصلاح محتواهای بازیابی شده از پایگاه‌های اطلاعاتی مذکور (۸۵ درصد) استفاده می‌شود.

◇ در فهرست‌نویسی بنیادی از سرعنوان‌های موضوعی کتابخانهٔ کنگرهٔ آمریکا (۷۲ درصد)، فهرست‌نویسی پیش‌ازانتشار (۵۸ درصد) و سرعنوان‌های موضوعی فارسی (۶۹ درصد) استفاده می‌شود.

◇ در نمایه‌سازی منابع اطلاعاتی نیز از اصطلاح‌نامه‌های موضوعی (۵۰ درصد) استفاده می‌شود. در این زمینه اکثر اصطلاح‌نامه‌های موضوعی موجود مورد استفاده قرار می‌گیرند. اصطلاح‌نامهٔ سه‌زبانۀ اصفا (۳۱ درصد) نیز از جمله اصطلاح‌نامه‌های عمومی پراستفاده است.

با توجه به یافته‌های ذکرشده در مورد وضعیت محتواهای پیشینه‌های کتابشناختی کتابخانه‌های دیجیتال مورد بررسی و نیز نظر 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 آیندهٔ نزدیک در نظر داشت. این یافته با یافتهٔ پژوهش «علی‌پور حافظی» (۱۳۸۸ و ۱۳۹۰) که به منظور شناسایی وضعیت سازماندهی منابع اطلاعاتی کتابخانه‌های دیجیتال انجام گرفته، همسو می‌باشد. همچنین، متأسفانه اغلب فیلدهای اطلاعاتی مربوط به منابع اطلاعاتی در کتابخانه‌های دیجیتال تکمیل نمی‌شوند و تنها فیلدهای ضروری و اصلی معدودی برای توصیف و تحلیل منابع اطلاعاتی حاوی محتواهای اطلاعاتی هستند. این یافته نیز با یافتهٔ «علی‌پور حافظی» (۱۳۹۰) همسوست. در زمینهٔ مستندسازی نیز که در یکپارچه‌سازی محتواهای اطلاعاتی بسیار مهم است، برای نام پدیدآورندگان در برخی از موارد از «فهرست مستند اسامی مشاهیر و مؤلفان» استفاده می‌شود. در مقابل، در برخی موارد دیگر مستندسازی در زمینهٔ اسامی افراد صورت نمی‌گیرد. در مورد مستندسازی اسامی ناشران وضعیت از این هم بدتر است و در اکثر موارد نسبت به مستندسازی اسامی ناشران اقدام نمی‌شود. همین مسئله باعث ایجاد مشکلاتی در یکپارچه‌سازی معنایی و محتوایی منابع پراکنده در کتابخانه‌های دیجیتال است.

در زمینه سازماندهی نیز با توجه به اینکه از دو شیوه بنیادی و نسخه‌برداری به‌طور 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 آمریکا» انجام می‌گیرد و سایر منابع اطلاعاتی با استفاده از اصطلاح‌نامه‌های متعدد موجود نمایه‌سازی می‌شوند. در صورتی که سازماندهی به‌صورت نسخه‌برداری نیز باشد، در این شرایط، کتابشناسی ملی و کتابخانه کنگره آمریکا که هر دو از سرعنوان‌های موضوعی استفاده می‌کنند، 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 تأثیرگذاری مناسبی در برقراری تعامل معنایی میان منابع اطلاعاتی داشته باشد. با وجود این، هنوز ابزار استاندارد تعامل معنایی در محیط دیجیتال مانند هستی‌نگاری‌های موضوعی و زبان‌های توصیفی مانند آدلیووال یا آردی‌اف توسط کتابخانه‌های دیجیتال مورد بررسی استفاده نمی‌شوند. بر این اساس، همین مسئله می‌تواند محدودیتی در راستای ایجاد یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال مورد بررسی به‌وجود آورد.

مسئله دیگر، توجه محدود به منابع چندرسانه‌ای در ارائه خدمات کتابخانه‌های دیجیتال مورد بررسی است. این منابع روزبه‌روز جایگاه بهینه‌ای در ارائه خدمات اطلاعاتی به کاربران به‌دست می‌آورند همین مسئله توجه جدی کتابخانه‌های دیجیتال به منابع غیرمتنی را می‌طلبد. بر این اساس، خواه‌ناخواه کتابخانه‌های دیجیتال باید به این سمت حرکت کنند و قابلیت‌های مورد نیاز را ایفاد نمایند. بنابراین، 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. مدل مفهومی پیشنهادی

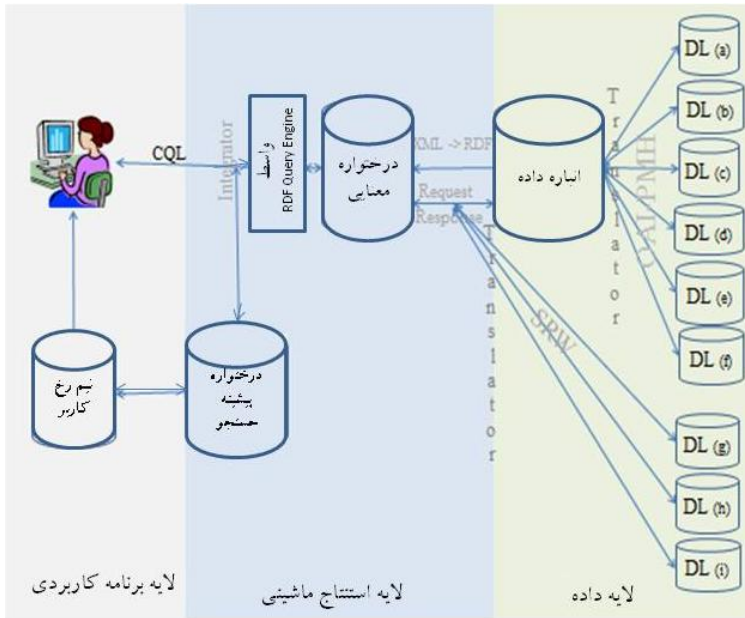
با توجه به یافته‌های پژوهش حاضر متوجه می‌شویم که عناصر، ابزارها و روش‌های متعددی در یکپارچه‌سازی معنایی در سیستم‌های اطلاعاتی کتابخانه‌های دیجیتال مورد استفاده قرار می‌گیرند. به‌طور کلی، تمامی این عناصر در قالب الگوهای متعددی می‌توانند مطرح شوند. روش‌های موجود در این زمینه در پاسخ به پرسش ۲ این پژوهش معرفی شدند. 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 دیجیتال ایران از منظر قابلیت‌های مبادله اطلاعات (نحوی و معنایی) مدلی متناسب با کتابخانه‌های دیجیتال ایران در پاسخ به این پرسش طراحی شد.

مدل مفهومی پیشنهادی برای یکپارچه‌سازی معنایی ناگزیر نیازمند اشاره به تعامل نحوی نیز می‌باشد. بنابراین، الگوی پیشنهادی باید به‌طور کلی پوشش دهنده معماری یکپارچه‌سازی باشد. علت اصلی طرح این مسئله 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 جست‌وجوی همزمان^۱ و برخی دیگر به شیوه برداشت اطلاعات^۲ است. این بدان معنی است که برخی از نرم‌افزارهای کتابخانه دیجیتال قادر به ارائه دسترسی به اطلاعات پایگاه اطلاعاتی خود به‌صورت باز با مرورگرهای شناخته شده هستند. برخی دیگر نیز این قابلیت را در اختیار قرار نمی‌دهند و تنها از طریق پروتکل Z39.50 یا «اس آریو» یا «اس آردلیو» این اجازه را به سایر سیستم‌های اطلاعاتی مشابه می‌دهند تا به محتواهای اطلاعاتی (در بیشتر موارد فراداده) دسترسی پیدا کنند. بنابراین، در چنین شرایطی نیاز به سیستمی دو وجهی است که هر دو شیوه دسترسی به اطلاعات را بتواند پوشش دهد. بر این اساس، مدل پیشنهادی پژوهش حاضر با توجه به نتایج پژوهش علیپور حافظی (۱۳۸۸) از نظر نحوی طوری طراحی شده که پوشش دهنده هر دو روش جست‌وجوی همزمان و برداشت اطلاعات باشد و در نتیجه، ساختار نحوی دو وجهی مورد توجه قرار گرفته است.

همچنین، سیستم‌های اطلاعاتی عضو شبکه یکپارچه، سیستم‌های اطلاعاتی نامتجانسی هستند که دارای معماری و ساختار متفاوتی از هم هستند (علیپور حافظی ۱۳۸۸). بنابراین، مسئله اصلی دیگر در این زمینه ترجمه و تبدیل زبان‌ها، استانداردهای فراداده 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 یکپارچه‌سازی وجود دارد. این مسئله زمانی اهمیت خود را دوچندان نشان می‌دهد که هدف، ایجاد یکپارچه‌سازی معنایی در ارائه خدمات به کاربران باشد. بنابراین، همان‌گونه که در مدل نیز اشاره شده در لایه داده، نیاز به واسطی برای ترجمه استانداردها و زبان‌هاست. بر این اساس، مدل مفهومی پیشنهادی پژوهش حاضر با توجه به ویژگی‌های مذکور و نتایج حاصل از این پژوهش به شکل ۳ در زیر ارائه می‌شود:

1. federated search

2. harvesting



شکل ۳. مدل مفهومی پیشنهادی یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال ایران

مدل پیشنهادی، همان‌گونه که از نتایج ذکر شده در پاسخ به پرسش اساسی ۱ مشخص شده، از سه لایه اصلی داده، استنتاج ماشینی و کاربرد تشکیل شده است. هر یک از سه لایه مذکور وظایف متعددی را برعهده دارد. 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 نیز نشان داده شده، نیاز به مترجم‌هایی برای تبدیل این استانداردها و فراداده‌ها به یکدیگر وجود دارد. نتایج حاصل از پاسخ به پرسش ۲ پژوهش و نیز مدل نحوی دو وجهی پیشنهاد شده در پژوهش علیپورحافظی (۱۳۸۸) توجیه‌کننده چنین ساختاری در لایه داده هستند.

در شکل ۳ 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 یا مرورگر نظام واسط پیشنهاد شده در این پژوهش می‌دهند. این نوع کتابخانه‌ها با مدل برداشت اطلاعات و با استفاده از پروتکل آی‌آی (که در پاسخ به پرسش اساسی ۱ این پژوهش به آن اشاره شد و گفته شد که پروتکل مناسبی در شبکه‌های باز و دسترس‌پذیر از طریق مرورگرها یا خزنده‌هاست) به سیستم واسط تعبیه شده در لایه استنتاج متصل می‌شوند و فراداده خود را در اختیار

سیستم واسط مذکور قرار می‌دهند. در مسیر انتقال فراداده‌ها همان‌گونه که از یافته‌های مرتبط با پرسش اساسی ۲ حاصل شد، مترجمی قرار می‌گیرد که وظیفه اصلی آن ترجمه استانداردها و یا فراداده‌های مختلف به استاندارد مورد نیاز سیستم واسط است. با توجه به تنوع استانداردها و فرمت‌های مورد استفاده توسط کتابخانه‌های دیجیتال عضو در شبکه نیاز به یکدست‌سازی آنها و استفاده از استاندارد فراداده ساده‌ای بود تا پاسخ‌گوی نیازهای یکپارچه‌سازی 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 پیش‌بینی شده که تمامی استانداردها و فرمت‌های فراداده به استاندارد «دابلین‌گر»^۱ ساده با ۱۴ فیلد زیر تبدیل شوند. شمای تبدیل تمامی استانداردها و فیلدهای ابر داده‌ای در مترجم مورد استفاده در این فرایند تعبیه می‌شود. فیلدهای فراداده مورد نظر عبارت‌اند از: عنوان، پدیدآور، موضوع، توصیف، ناشر، همکار، تاریخ، نوع، قالب، شناسگر، منبع، زبان، رابطه، حقوق. چهارده فیلد پیشنهادی در این بخش که مبتنی بر استاندارد فراداده «دابلین‌گر» است، از نتایج پاسخ به پرسش اساسی ۲ این پژوهش حاصل شده است و مطابق با نتایج پژوهش علیپور حافظی (۱۳۹۰) است.

همین مترجم در فرایند تبدیل فراداده‌ها در یکپارچه‌سازی سیستم‌های اطلاعاتی توزیع یافته نیز مورد استفاده قرار می‌گیرد. در بخش پایین لایه داده مدل مفهومی پیشنهادی، کتابخانه‌های دیجیتال توزیع یافته‌ای قرار دارند که اجازه دسترسی به محتوای فراداده را به سیستم واسط نمی‌دهند (دسته دوم). همان‌گونه که در مدل پیشنهادی نیز به این موضوع پرداخته شده، نحوه تعامل با این سیستم‌ها با استفاده از مدل جست‌وجوی همزمان خواهد بود. بر اساس نتایج پژوهش در پاسخ به پرسش اساسی ۲، در این مدل پیشنهاد می‌شود که از پروتکل «اس آردلیو» استفاده شود. این پروتکل نسل جدید پروتکل Z39.50 برای برقراری ارتباط همزمان در کتابخانه‌های دیجیتال است. در فرایند 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 همزمان با ارسال دستور دریافت اطلاعات کاربر به سیستم اطلاعاتی واسط ارسال می‌شود و خروجی و پاسخ سیستم‌های اطلاعاتی توزیع یافته نیز همزمان با دریافت پاسخ از سیستم واسط توسط پایگاه معنایی دریافت می‌شود.

در لایه میانی، لایه استنتاج، بر اساس نتایج این پژوهش در پاسخ به پرسش‌های ۱ و ۲، دو پایگاه اطلاعاتی مربوط به پایگاه اطلاعاتی شبکه معنایی فراداده‌ها و پایگاه اطلاعاتی شبکه ارتباطی

1. Dublin Core

پیشینه‌های جست‌وجوی کاربران قرار دارند. منطق فعالیت این پایگاه‌های اطلاعاتی برقراری و ایجاد شبکه‌های معنایی مورد نیاز برای بازیابی معنایی است. در انبار معنایی ساختار موجود مبتنی بر روابط سرعنوان‌های موضوعی، اصطلاح‌نامه‌های مورد استفاده در نمایه‌سازی، و مستند اسامی مورد استفاده در مستندسازی اسامی خواهد بود. جهت غنای بیشتر این شبکه معنایی پیشنهاد می‌شود که روابط بین ساختارهای معنایی مذکور با یکدیگر از طریق پیوندهای موجود در شبکه‌های علمی و داده‌های پیوندی برقرار شود. برای این منظور، نیاز به تبدیل فراداده‌ها از فرمت «ایکس‌ام‌ال» به «آردی‌اف» برای استفاده در شبکه معنایی مذکور، بر اساس نتایج حاصل از پرسش اساسی ۱ پژوهش حاضر، وجود دارد. ساختار این شبکه معنایی با استفاده از «آردی‌اف» درست مانند شبکه معنایی «اسکاس» خواهد بود. یافته‌های ارائه‌شده در پاسخ به پرسش‌های اساسی ۱ و ۲ پژوهش توجیه‌کننده چنین ساختار معنایی است. در این مرحله تنها فیلدهای ضروری مورد استفاده در شبکه معنایی وارد این پایگاه اطلاعاتی می‌شوند. فیلدهای مذکور عبارت‌اند از: پدیدآور، موضوع، توصیف، همکار و شناساگر. همچنین، 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. به بیانی دیگر، این انبار معنایی برای شناسایی دقیق محتواهای مورد نیاز کاربران در بخش جست‌وجوی همزمان در مسیر ارسال پرس‌وجو و نیز در مسیر دریافت پاسخ سیستم‌های اطلاعاتی پراکنده قرار می‌گیرد تا بتواند پاسخ معنایی را ارائه دهد. بنابراین، هم کلیدواژه‌های جست‌وجوی کاربر وارد این سیستم می‌شوند و کلیدواژه‌های مرتبط شناسایی و به کتابخانه‌های دیجیتال پراکنده ارسال می‌شود و هم پاسخ‌های دریافتی از کتابخانه‌های دیجیتال حاضر در بخش جست‌وجوی همزمان با انبار معنایی و نیز با پاسخ بخش دیگر کتابخانه‌های دیجیتال، که داده‌های آنها از طریق برداشت اطلاعات حاصل می‌شوند، تلفیق می‌گردند.

با توجه به یافته‌های حاصل از پرسش اساسی ۱ پژوهش پیشنهاد می‌شود که از زبان «سی کیوال»^۱ برای دریافت پرس‌وجوهای کاربر استفاده شده و این پرس‌وجوها با توجه به نیاز به تبدیل آنها به ساختار مناسب مورد نیاز در انبار معنایی پیشنهادشده در سیستم، در میان‌افزاری تبدیل به پرس‌وجوی «آردی‌اف» می‌شود. برای غنی‌سازی و دقت در پاسخ‌های ارائه‌شده به کاربران در این سیستم پیشنهاد می‌شود که از فناوری سیستم‌های یادگیرنده^۲ و تصمیم‌یار^۳ استفاده شود. پایگاه اطلاعاتی مربوط به ثبت پیشینه جست‌وجوی کاربران نیز مبتنی بر سیستم‌های

1. CQL

2. learning systems

3. Decision Support System (DSS)

یادگیرنده 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 غیرقابل تصور هستند، در سیستم‌های یادگیرنده به‌وجود می‌آید.

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 توصیفی برای توفیق و کارایی سیستم‌های یادگیرنده نقش حیاتی دارد. قوانین شبکه‌های معنایی و قالب‌ها 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، باید قابل فهم توسط ماشین و انسان باشد.

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

در هنگامی که کاربر در حال جست‌وجوی منابع مورد نظر خود در این سیستم است، پس از دیدن نتایج حاصله، کلمات کلیدی جدیدی در ذهن او شکل می‌گیرد. لذا، سلسله‌کلماتی که او به سیستم وارد می‌کند، دارای نظم منطقی می‌تواند باشد و ناشی از برقراردن ارتباطات معنایی بین آنها از سوی فرد جست‌وجوکننده است؛ به این مفهوم که احتمالاً مجموعه این کلمات از نظر معنایی با یکدیگر ارتباطی دارند که جست‌وجوگر بعد از جست‌وجو در مورد هر یک از آنها و دیدن بازخورد سیستم، از بین کلمات کلیدی ثانویه‌ای که سیستم به او بازخورد می‌دهد کلمات خاص دیگری را انتخاب کرده و در مورد آنها جست‌وجو می‌کند.

اما مسئله این است که ممکن است در این فرایند، جست‌وجوگر کلمات نامرتبیطی را هم که به نظر او مرتبط می‌رسند، در جست‌وجوی خود وارد کند؛ بدون اینکه این کلمات اخیر ارتباطی با

موضوع اصلی داشته باشند. اگر بنا داشته باشیم که سیستمی یادگیرنده بسازیم که با استفاده از کلماتی که جست و جویگران مختلف در یک بازه زمانی در مورد آنها جست و جو می کنند، 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 می کنند استفاده کرده و کلماتی را که توسط کاربران مختلف به کرات و به توالی در موردشان جست و جو می شود، «دارای ارتباط معنایی با یکدیگر» تلقی کنیم، و ثانیاً بتوانیم نامرتب بودن برخی کلمات را که در این توالی ها دیده می شوند، تشخیص داده و آنها را وارد این رده از کلمات مرتبط نماییم.

یادگیرنده بودن سیستم، داشتن ویژگی سومی را نیز برای آن تکلیف می کند، و آن این است که انتظار می رود که با توجه به پویا بودن مجموعه های علمی، کلمه ای که ممکن است قبلاً در رده ای 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 رده بندی را به صورتی متناسب با آن تغییر دهد (اصلاح کند).

با این تفاسیر، روش کار در این مرحله برای ایجاد روابط معنایی با استفاده از کلیدواژه ها و نتایج حاصل از جست و جوی کاربران و با استفاده از پایگاه پرو فایل کاربران به این صورت است:

1. دوره ای خاص به عنوان دوره یادگیری تعیین می شود؛ مثلاً یک سال یا شش ماه.
2. در طی این یک سال یا شش ماه، دنباله کلمات کلیدی که توسط کاربران مختلف در موردشان به توالی جست و جو شده، مورد بررسی قرار می گیرند.
3. از روی این بررسی، گروه هایی که کاربران مختلف به اجماع قائل به مرتبط بودن آنها با یکدیگر هستند در یک خانواده قرار داده می شوند، و به عبارتی دارای ارتباط معنایی با یکدیگر در نظر گرفته می شوند. به این ترتیب، اثر جست و جوهای نامرتب که طی این دوره توسط کاربران محدودی به اشتباه مرتبط با کلمه اصلی قلمداد شده اند، خنثی شده، و این کلمه اخیر در خانواده قرار داده نمی شود.
4. وقتی کاربری در مورد کلمه ای خاص به جست و جو می پردازد، با توجه به دسته خانواده های ساخته شده در بالا، پیشنهادهایی در خصوص کلمات مرتبط دیگری که این کاربر می تواند در مورد آنها نیز به جست و جو پردازد، به او ارائه می شود.
5. وقتی دوره به کارگیری سیستم از شش ماه یا یک سال بیشتر شد، هر روز اطلاعات جمع آوری شده از به عنوان مثال ۳۶۶ روز به قبل کنار گذاشته می شود و اطلاعات جمع آوری شده از روز آخر وارد فرایند خانواده بندی مندرج در بند ۳ می گردد. به این ترتیب، پویایی سیستم در شبکه سازی همواره حفظ می شود.

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

در لایه برنامه کاربردی که در سمت چپ الگوی پیشنهادی ذکر شده در شکل ۳ نشان داده شده، ارتباط کاربران با پورتال جامع شبکه یکپارچه برقرار می‌شود. پورتال 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 پورتال شبکه معنایی پیشینه جست‌وجوی کاربران به تدریج شکل می‌گیرد و با گذشت یک سال از شروع، می‌تواند نتایج خود را در بازیابی معنایی به کاربران نشان دهد. این درگاه می‌تواند درگاهی شبیه به درگاه‌های کنسرسیومی باشد که در آن تمامی کاربران کتابخانه‌های دیجیتال عضو می‌توانند از یک درگاه واحد وارد این سیستم شوند. با توجه به شکل‌گیری کنسرسیوم محتوای ملی در کشور در چند سال گذشته و فعال‌بودن آن و نیاز به مدیریت و توسعه ساختار و مدل پیشنهادی در این پژوهش پیشنهاد می‌شود که کنسرسیوم محتوای ملی متولی‌گری این حوزه را بر عهده گرفته و نسبت به پیاده‌سازی و مدیریت آن اقدام نماید. هر چند بهره‌گیری از چنین نظامی و شناسایی ساختار مدیریتی و اجرایی این نظام نیازمند پژوهش مستقل دیگری است، ولی پیشنهاد مذکور با توجه به حضور فعال کنسرسیوم محتوای ملی به عنوان تنها کنسرسیوم فعال در این زمینه منطقی به نظر می‌رسد.

۶. اعتبارسنجی مدل پیشنهادی

ارائه مدل، بدون آزمون و اعتبارسنجی آن نمی‌تواند کارآمدی خود را به اثبات برساند. برای این منظور، لازم است که ارائه هر مدل، الگو، یا راهکاری پیش از انتشار از نظر اعتبار و کارآمدی مورد آزمون قرار گیرد. لازم به ذکر است که هر الگو، مدل، یا راهکاری برای حل یا تسهیل مسئله‌ای ارائه می‌شود. مهم‌ترین عنصر در آزمون مدل، راهکار، یا الگویی توجه به مسئله و هدف از ارائه آن است. برای آزمون مدل‌ها هفت روش تاکنون مطرح شده است. روش‌های مذکور عبارت‌اند از: ارائه^۱، آزمون تجربی^۲، شبیه‌سازی^۳، استفاده از معیارها^۴، معیارسازی^۵، استدلال منطقی^۶ و، اثبات ریاضی^۷ (Vaishnavi and Kuechler Jr. 2008).

1. demonstration
4. using metrics
7. mathematical proofs

2. experimentation
5. Benchmarking

3. simulation
6. logical reasoning

با توجه به مسئله مورد نظر و راهکار ارائه شده لازم است 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 که راهکار پیشنهادی در عمل قابلیت ارائه داشته باشد. برخی دیگر مانند شبیه‌سازی و معیارسازی 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 که ابتدا مدل پیشنهادی تهیه شود. این مدل نشانگر آن است که در راستای نیاز طرح شده در ابتدای پژوهش طراحی شده و پاسخ‌گوی آن است. در ادامه، باید نشان داد که این مدل می‌تواند مسئله طرح شده را مرتفع نماید. به این منظور ابتدا در پژوهش حاضر با تحلیل محتواها و شناسایی وضعیت موجود کتابخانه‌های دیجیتال ایران از منظر یکپارچه‌سازی معنایی اقدام به طراحی مدل مفهومی پیشنهادی شد. سپس در گام بعدی اقدام به مطابقت آن با مسئله طرح شده در ابتدای پژوهش شد. در این مرحله از نظر متخصصان حوزه علم اطلاعات و دانش‌شناسی آشنا به حوزه فناوری اطلاعات و متخصصان علوم کامپیوتر آشنا به یکپارچه‌سازی سیستم‌های اطلاعاتی و وب معنایی استفاده گردید. پس از جلسات متعدد با متخصصان مذکور و مطابقت مدل با نیاز موجود، اهداف و مسئله پژوهش، در نهایت به قالب نهایی مدل پیشنهادی دست یافتیم که متناسب با شرایط موجود کتابخانه‌های دیجیتال ایران، قادر به یکپارچه‌سازی معنایی بوده و امکان توسعه متناسب با شرایط موجود را داراست. در نهایت، بر همین مبنای مدل پیشنهادی این پژوهش به لحاظ اعتبار، مورد سنجش قرار گرفته و کارآمدی آن برای حل مسئله پژوهش مورد تأیید قرار گرفت.

۷. پیشنهادهای پژوهش

در ادامه، با توجه به نتایج پژوهش، برخی پیشنهادهای اجرایی جهت بهبود وضعیت کتابخانه‌های دیجیتال به لحاظ برقراری روابط معنایی در یکپارچه‌سازی کتابخانه‌های دیجیتال ایران ارائه می‌شوند:

◇ نتایج پژوهش حاضر در زمینه شناسایی عناصر دخیل در یکپارچه‌سازی معنایی کتابخانه‌های دیجیتال 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 دیجیتال اجرا شود. برای این منظور، به ایجاد قابلیت‌هایی برای تعامل مانند پروتکل‌های یکپارچه‌سازی کتابخانه دیجیتال و تبعیت از استانداردهای فراداده و تولید پایگاه‌های دانشی نیاز است.

◇ نتایج این پژوهش در زمینه روش‌های مورد استفاده برای یکپارچه‌سازی معنایی در کتابخانه‌های دیجیتال نشان داد که به استفاده از مترجم‌ها و ماشین‌های استنتاج نیاز وجود دارد. لذا، پیشنهاد می‌شود که شمای تبدیل استانداردهای فراداده به یکدیگر شناسایی و نرم‌افزارهایی برای این منظور تولید شوند. در این راستا، نرم‌افزارهای منبع باز مانند «مارک‌آدیت»^۲ 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 دیجیتال نشان داد که در این زمینه به پایگاه‌های دانشی، هستی‌نگاری‌ها، سیستم‌های سازماندهی دانش و مترجم‌ها نیاز 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 گیرند.

◇ نتایج پژوهش در زمینه وضعیت موجود محتواهای پیشینه‌های کتابشناختی کتابخانه‌های دیجیتال مورد بررسی نشان داد که تأکید بر منابع متنی در این کتابخانه‌ها وجود دارد. در این زمینه توجه به سایر منابع به‌ویژه منابع چندرسانه‌ای و غیرمتنی با توجه به تنوع اطلاعات موجود در دسترس می‌تواند مهم باشد. علاوه بر این، توصیف داده‌ها در فراداده‌های تولیدی کتابخانه‌های دیجیتال 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.

۱. با توجه به اینکه در این پژوهش لازم بود مدلی برای یکپارچه‌سازی معنایی تهیه شود و یکی از ابزارهای مهم مورد نیاز، هستی‌نگاری‌های مناسب برای این منظور بود، لازم است پژوهش‌هایی در زمینه ایجاد هستی‌نگاری‌ها و کاربرد آنها انجام پذیرد. از آنجا که لازم است پژوهش مجزایی در این زمینه صورت گیرد، این پژوهش، فقط به استفاده از هستی‌نگاری‌ها در بخشی از مدل پیشنهادی اشاره کرده و به شیوه‌ها و نحوه ایجاد آنها نپرداخته و فقط به جایگاه و ضرورت و کاربرد آنها پرداخته است.

2. MARCEDIT

نتایج حاصل از این پژوهش در زمینه مدل مفهومی مناسب برای یکپارچه‌سازی نیز نشان داد که با توجه به اینکه این مدل برگرفته از مدل‌های موجود و شخصی‌سازی آن منطبق با توانمندی‌ها و ویژگی‌های کشورمان است، پیشنهاد می‌شود این مدل 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.

فهرست منابع

- صنعت‌جو، اعظم و اکرم فتحیان. ۱۳۹۰. مقایسه کارآمدی اصطلاح‌نامه و هستی‌شناسی در بازنمون دانش (طراحی و ساخت نمونه هستی‌شناسی اصفهان). *پژوهشنامه کتابداری و اطلاع‌رسانی* ۱ (۱): ۲۱۹-۲۴۰.
- علیپور حافظی، حامد. ۱۳۹۰. بررسی وضعیت سازماندهی منابع اطلاعاتی در کتابخانه‌های دیجیتال ایران و ارائه راهکار پیشنهادی. پایان‌نامه کارشناسی ارشد به راهنمایی عباس حرّی. همدان، دانشگاه آزاد اسلامی همدان.
- _____، مهدی. ۱۳۸۸. بررسی نحوه مبادله اطلاعات بین سیستم‌های اطلاعاتی کتابخانه‌های دیجیتال در ایران و ارائه الگوی پیشنهادی. پایان‌نامه دکتری به راهنمایی عباس حرّی. دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران.
- _____ . ۱۳۹۱. بررسی هزینه-سازگاری سیستم‌های اطلاعاتی کتابخانه‌های دیجیتال در ایران و ارائه الگوی پیشنهادی. *پژوهشنامه پردازش و مدیریت اطلاعات* ۲۸ (۲): ۵۲۳-۵۴۳.
- محقق، نیلوفر. ۱۳۷۴. بررسی وضعیت مبادله اطلاعات و منابع میان کتابخانه‌های دانشگاهی شهر تهران. پایان‌نامه کارشناسی ارشد به راهنمایی عباس حرّی. دانشگاه تهران.
- Alipour-Hafezi, Mehdi., Abbas Horri, Ali Shiri, and Amir Ghaebi. 2010. Interoperability models in digital libraries: an overview. *The Electronic Library* 28 (3): 438-452.
- Chawathe, Sudarshan, Hector Garcia-Molina, Joachim Hammer, Kelly Ireland, Yannis Papakonstantinou, Jeffery D. Ullman, and Jennifer Widom. [1996]. The TSIMMIS Project: Integration of Heterogeneous Information Sources. <http://db.ucsd.edu/publications/tsimmis-overview.ps> (accessed July 10, 2013).
- Chen, Harry, Tim Finin, and Anupam Joshi. 2003. An Ontology for Context-Aware Pervasive Computing Environments. *Knowledge engineering review* 18 (3): 197-207.
- Creating a European library space telematics for libraries programmes 1990-1998. 2000. <http://cordis.europa.eu/libraries/en/intro.html> (accessed July 10, 2013)
- DELOS. 2004. Welcome to the DELOS network of excellence. http://www.delos.info/index.php?option=com_frontpage&Itemid=1 (accessed July 10, 2013)
- Ding, Hao, and Ingeborg T. Sølberg. 2004. *Towards the schema heterogeneity in distributed digital libraries. ICEIS 2004 – human computer interaction*. Porto: Universidade Portucalense and INSTICC: 307-312.
- Guha, Nabonitea. 2006. Semantic Digital Library Services. http://www.I3s.de/kweb/kwepsy2006/FinalSubmissions/kwepsy2006_guha.pdf (accessed July 10, 2013)
- Han, Yan. 2006. RDF-based digital library system. *Library Hi Tech* 24 (2): 234-240.
- Huang, Su-Hsien., Hao-Ren Ke, and Wei-Pang Yang. 2005. Enhancing semantic digital library query

- using a content and service inference model (CSIM). *Information processing and management* 41: 891-908.
- Introduction Europeana. 2011. <http://www.europeana.eu/portal/api-introduction.html> (Accessed July 10, 2013)
- Issac, Antoine, Stefan Schlobach, Henk Mattheizing and Claus Zinn. 2008. Integrated access to cultural heritage resources through representation and alignment of controlled vocabularies. *Library Review* 57 (3): 187-199.
- Lagoze, Carl, Walter Hoehn, David Millman, William Arms, Stoney Gan, Diane Hillmann, Christopher Ingram, Dean Krafft, Richard Marisa, Jon Phipps, John Saylor, Carol Terrizzi, James Allan, Sergio Guzman-Lara, and Tom Kalt. 2002. *Core Services in the Architecture of the National Digital Library for Science Education (NSDL)*. JCDL 2002. Portland: ACM. <http://arxiv.org/ftp/cs/papers/0201/0201025.pdf> (accessed July 15, 2013)
- Mayer, Philipp, Peter Mutschke, and Vivien Petras. 2008. Reducing semantic complexity in distributed digital libraries: treatment of term vagueness and document re-ranking. *Library Review* 57 (3): 213-224.
- NSDL technical infrastructure white paper version 2.0. 2004. NSDL. http://nsdl.com.nsd.org/meeting/archives/smete/workgroups/technical/nsdl_tech_arch20.doc (accessed July 15, 2013)
- Pasad, A.R.D. and Devika P. Madalli. 2008. Faceted infrustructure for semantic digital libraries. *Library Review* 57 (3): 225-234.
- Patel, Manjula, Traugott Koch, Martin Doerr, and Chrisa Tsinaraki. 2005. Semantic interoperability in digital library systems. DELOS: A Network of Excellence on Digital Libraries (Project no. 507618). <http://delos-wp5.ukoln.ac.uk/project-outcomes/SI-in-DL/> (Accessed July 10, 2013)
- Shen, Rao. 2006. *Applying the 5S Framework To Integrating Digital Libraries*. Virginia Polytechnic Institute. Virginia: Virginia Polytechnic Institute.
- Shi, Rong. 2005. *Lightweight federation of non-cooperating digital libraries*. Computer Science. Old Dominion University. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.114.5754&rep=rep1&type=pdf> (accessed July 15, 2013)
- Szymański, Julian. 2011. *Cooperative WordNet Editor for Lexical Semantic Acquisition*. IC3K 2009. Berlin: Springer-Verlag: 187-196.
- Vaishnavi, Vijay K.; William Kuechler Jr. 2008. *Design science research methods and patterns: innovating information and communication technology*. Boca Raton: Auerbach publications.
- Vdovjak, Richard & Geert-Jan Houben. 2001. RDF Based Architecture for Semantic Integration of Heterogeneous Information Sources. http://iwayan.info/Research/Interoperability/Papers_Research/SemanticMediation/wiiv01.pdf (accessed July 15, 2013)
- Warren, Paul, David Alsmeyer. 2005. Applying semantic technology to a digital library: a case study. *Library management* 26 (4-5): 196-205.

مهدی علی‌پور حافظی

متولد سال ۱۳۵۲، دارای مدرک دکتری در رشته علم اطلاعات و دانش‌شناسی از دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران است. ایشان هم‌اکنون استاد مدعو گروه علم اطلاعات و دانش‌شناسی دانشگاه علامه طباطبائی است. کتابخانه‌های دیجیتالی، مبادله اطلاعات، امانت الکترونیکی و روش‌شناسی پژوهش در علم اطلاعات و دانش‌شناسی از جمله علایق پژوهشی وی است.

