Explaining the Concept and Models of Information Fusion and its Application in the Thesaurus-based Information Management Systems

Morteza Mohammadi Ostani
PhD in Knowledge and Information Science; Assistant Professor; University of Isfahan; Isfahan, Iran Email: m.ostani@edu.ui.ac.ir

Mohammad Karimi
M.A. in Information Technology Engineering; Islamic Sciences and Culture Academy; Qom, Iran Email: mo.karimi@isca.ac.ir

Ali Mirarab*
PhD in Information Technology Management; Assistant Professor; Islamic Sciences and Culture Academy; Qom, Iran;
Email: alimirarab@isca.ac.ir

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. 39 | No. 2 | pp. 565-598
Winter 2024
https://doi.org/10.22034/jipm.2023.706406

Received: 30, Mar. 2023 **Accepted: 17, Jun. 2023**

Abstract: Information fusion is a vital process in information management systems which aims to merge information from multiple sources to provide a more comprehensive and accurate view of specialized domain. In structured information management systems there are several methods for aggregating, consolidating, and fusioning information, but their approaches have not yet provided a clear pattern. The aim of the current research is to explain the concept and appropriate models of information fusion in information management systems and its use in relationship-based databases such as thesaurus. The research method is a conceptual type with an analytical approach. The research population consisted of texts and outputs in the field of information fusion, and the data collected by the library method. The findings showed that information fusion models are in four general categories of models based on information flow, workflow and activity, roles and functions of entities, and understanding concepts. The Omnibus workflow and activity model and the model based on the roles and functions of Endsley's entity as selected models according to the basic JDL are suitable for use in relationship-based information management systems such as thesauruses and ontologies, and they have the ability to be used with some changes. The main disadvantage of the reviewed information fusion models is not paying attention to the characteristics of specific information structures such as thesauruses.

* Corresponding Author

Not considering role of expert users in information systems based on user decisions that have the task of extracting specialized information and ignoring the collective participation system to solve complex problems, including finding similarities and solving it in the mass of information, not describing the required functions to solve the problems, managing the effects in case of system or user errors and restoring fusioned and consolidated information, as well as non-implementation of methods with operational examples in such structures, can be counted among these problems. In general, in each model of information fusion in information management systems based on thesaurus, the general processes of finding similarities, examining similarities, aggregating parameters and information fusion and managing the effectiveness of sub-systems should be considered. Unlike traditional information fusion methods that focus more on merge data, semantic information fusion emphasizes fusioning related knowledge and concepts stored in thesauruses instead of merging data and information. Therefore, it is suggested to pay attention to semantic approaches in the process of integrating and fusioning new models.

Keywords: Data Fusion, Information Fusion, Information Fusion Models, Information Management Systems, Thesaurus

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

تبیین مفهوم و مدل‌های ادغام اطلاعات و کاربست آن در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای

مرتضی محمدی استانی
دکتری علم اطلاعات و دانش‌شناسی؛ استادیار؛
گروه علم اطلاعات و دانش‌شناسی؛ دانشگاه اصفهان؛
اصفهان ایران m.ostani@edu.ui.ac.ir

محمد کریمی
کارشناسی ارشد مهندسی فناوری اطلاعات؛
گروه اشاعه اطلاعات و تبادل دانش؛ پژوهشگاه علوم و فرهنگ اسلامی؛ قم، ایران mo.karimi@isca.ac.ir

علی میرعرب
دکتری مدیریت فناوری اطلاعات؛ استادیار؛
گروه اشاعه اطلاعات و تبادل دانش؛
پژوهشگاه علوم و فرهنگ اسلامی؛ قم، ایران؛
پدیدآور رابط alimirarab@isca.ac.ir

پژوهش‌نامه پردازش و مدیریت اطلاعات

دریافت: ۱۴۰۲/۰۱/۱۰ | پذیرش: ۱۴۰۲/۰۳/۲۷ | مقاله برای اصلاح به مدت ۱۵ روز نزد پدیدآوران بوده است.

نشریهٔ علمی | رتبهٔ بین‌المللی
پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)
شاپا (چاپی) ۸۲۲۳-۲۲۵۱
شاپا (الکترونیکی) ۸۲۳۱-۲۲۵۱
نمایه در LISTA، ISC، و SCOPUS
jipm.irandoc.ac.ir
دورهٔ ۳۹ | شمارهٔ ۲ | ص‌ص ۵۶۵-۵۹۸
زمستان ۱۴۰۲
https://doi.org/10.22034/jipm.2023.706406

چکیده: روش‌های مختلفی برای تجمیع، تلفیق و ادغام اطلاعات به‌عنوان فرایندی حیاتی در سیستم‌های مدیریت اطلاعات وجود دارد، اما رویکرد و الگوی مشخصی در حوزه سیستم‌های مدیریت اطلاعات مبتنی بر روابط (اصطلاحنامه‌ای) ارائه نشده است. هدف پژوهش حاضر تبیین مفهوم و مدل‌های مناسب ادغام اطلاعات و کاربست آن در سیستم‌های مدیریت اطلاعات مبتنی بر رابطه (اصطلاحنامه‌ای) است. پژوهش حاضر از نوع مفهومی با رویکرد تحلیلی است. جامعه پژوهش، متون و بروندادهای حوزه ادغام اطلاعات در پایگاه‌های اطلاعاتی مختلف بود. روش گردآوری داده‌ها، کتابخانه‌ای بود که به‌صورت هدفمند انتخاب گردید. یافته‌ها نشان داد که ادغام، تلفیق و ترکیب اطلاعات در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای، نتایج و تأثیرات متفاوتی ایجاد می‌کنند. مدل‌های ادغام اطلاعات در چهار دسته کلی مدل‌های مبتنی بر گردش اطلاعات، گردش کار و فعالیت، نقش‌ها و عملکرد موجودیت‌ها، و درک مفاهیم قرار دارند. مدل گردش کار و فعالیت Omnibus و مدل مبتنی بر نقش‌ها و عملکرد موجودیت Endsley's به‌عنوان مدل‌های انتخابی با توجه به مدل پایه JDL جهت به‌کارگیری در سیستم‌های مدیریت اطلاعات مبتنی بر

روابط نظیر اصطلاحنامه‌ها و هستی‌شناسی‌ها مناسب است و این قابلیت را دارند که با برخی تغییرات استفاده شوند. عیب اصلی مدل‌های ادغام اطلاعات، عدم توجه به ویژگی‌های ساختارهای اطلاعاتی مشخص نظیر اصطلاحنامه‌هاست. کم‌توجهی به کاربران خبره در سیستم‌های اطلاعاتی مبتنی بر تصمیمات کاربری که وظیفه استخراج اطلاعات تخصصی را دارند و نادیده‌گرفتن سیستم مشارکت جمعی برای حل مسائل پیچیده از جمله تشابه‌یابی و رفع آن در حجم انبوه اطلاعاتی، عدم شرح ریز عملکردهای مورد نیاز برای حل مسائل، مدیریت تأثیرات در صورت بروز اشتباهات سیستمی یا کاربری و بازگردانی اطلاعات تجمیعی و تلفیقی، و همچنین عدم اجرای روش‌ها با نمونه‌های عملیاتی در چنین ساختارهایی را می‌توان از جمله این مشکلات برشمرد. به‌طور کلی، باید در هر مدل ادغام اطلاعات در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای، فرایندهای کلی تشابه‌یابی، بررسی تشابهات، تجمیع پارامترها و ادغام اطلاعات و مدیریت تأثیرپذیری زیرسیستم‌ها مورد توجه قرار گیرد. بر خلاف روش‌های سنتی ادغام اطلاعات که بیشتر بر ترکیب داده‌ها تمرکز دارند، ادغام اطلاعات معنایی به‌جای ترکیب داده‌ها و اطلاعات، بر ترکیب دانش‌های مرتبط و مفاهیم ذخیره‌شده در اصطلاحنامه‌ها تأکید دارد. از این‌رو، توجه به رویکردهای معنایی در فرایند ترکیب و ادغام مدل‌های جدید پیشنهاد می‌گردد.

کلیدواژه‌ها: ادغام داده، ادغام اطلاعات، مدل‌های ادغام اطلاعات، سیستم‌های مدیریت اطلاعات، اصطلاحنامه

۱. مقدمه

امروزه، حجم عظیمی از اطلاعات و دانش در قالب‌های مختلف از جمله قالب متنی به‌صورت منسجم و ساختارمند و یا غیرمنسجم و غیرساختارمند از طریق سیستم‌های مدیریت اطلاعات[1] در دسترس است. این سیستم‌ها عمل جمع‌آوری، کنترل، پالایش، بازیابی و مدیریت داده و اطلاعات مورد نیاز سازمان‌ها را توسط انسان و ماشین بر عهده دارند. عوامل متعددی می‌توانند منجر به اخلال در عملکرد اطلاعات در سیستم‌های مدیریت اطلاعات گردد که مهم‌ترین آن‌ها عدم توجه به سیستم‌های مدیریت یکپارچه[2] است. به‌رغم وجود سیستم‌های پردازش مکانیزه و بسته‌های نرم‌افزاری کاربردی درون‌سازمانی در کنار سیستم‌های مدیریت اطلاعات، مفهوم یکپارچگی داده و اطلاعات در سازمان‌ها با ورود سیستم‌های هوش تجاری[3] و فناوری‌های معنایی تغییرات مناسبی در زمینه دسترسی و ایجاد داده‌های یکپارچه، پاک‌سازی‌شده و دسته‌بندی‌شده با رویکردهای معنایی برای مدیران و کاربران خبره سازمان ایجاد کرده است. حجم انبوه

1. information management systems 2. integrated management systems 3. business intelligence

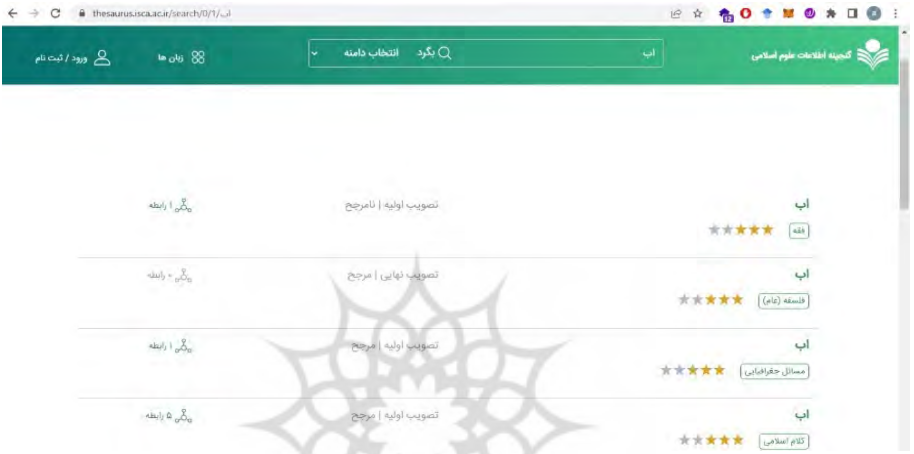
داده‌ها و اطلاعات در این‌گونه سیستم‌ها، دقت و صحت در بازیابی صحیح اطلاعات را طلب می‌نماید که از نظر محتوایی به دو دسته ایستا و پویا تقسیم‌بندی می‌شوند. اطلاعات ایستا، اطلاعاتی تغییرناپذیر است. اطلاعاتی که در کتاب، تصویر، صفحه فشرده و مانند آن نگهداری می‌شوند، از جمله اطلاعات ایستا به‌شمار می‌آیند. اطلاعاتی که در جریان کنش یا تعامل، اصلاح، تحلیل مورد پردازش قرار می‌گیرند از جمله اطلاعات پویا به‌شمار می‌آیند. نگهداری و پردازش اطلاعات پویا برای مدیریت و کاوش برای سازمان‌هایی که در حوزه مدیریت و طبقه‌بندی داده و اطلاعات فعالیت می‌نمایند، از اهمیت بسیار بالایی برخوردار است.

دستیابی به اطلاعاتِ دربردارنده محتوای ارزشمند و مدیریت‌شده در کمترین زمان و با صرف کمترین هزینه برای سازمان‌ها دارای اولویت زیادی است. از نگاهی دیگر، جهت اخذ تصمیم‌های اضطراری، راهبردی و حتی تصمیمات کلان، سازمان‌هایی که در حوزه پردازش و سازماندهی اطلاعات فعالیت می‌کنند، وجود تحلیل‌های صحیح از وضعیت عملکردی اطلاعات امری ضروری محسوب می‌شود. این تحلیل‌ها نیازمند وجود داده‌های باکیفیت و در دسترس است. در اینجا روش‌های سازماندهی و پردازش اطلاعات بستر مناسبی برای ساختارمندکردن داده و اطلاعات برای مدیریت، تحلیل، مدل‌سازی و مهندسی پا به عرصه فعالیت می‌گذارند.

اصطلاحنامه‌ها یکی از مؤثرترین روش‌های سازماندهی و پردازش اطلاعات به‌صورت سلسله‌مراتبی و کنترل‌شده در چند دهه گذشته بوده است که نقش مهمی در یکپارچگی داده و اطلاعات دارند (Pai, Yang and Chung 2017). مدیریت اصطلاحنامه از گذشته تا به‌حال یکی از مسائل مهم حوزه اصطلاح‌شناسی است. در تدوین و نگهداری اصطلاحنامه‌ها، بررسی مسائلی نظیر میزان هم‌ارزی اصطلاحات، میزان واژه‌های پیش‌همارا و پس‌همارا، و تعیین توضیح‌گر برای کلمات مشابه از موارد مهم و اساسی است. بررسی و تحلیل این موارد سبب حل مسائل و سهولت سازگاری یا یکپارچگی در زبان نمایه‌سازی می‌گردد. جایگاه یکپارچه‌سازی داده‌ها در ارتباط مستقیم با مدیریت داده‌های اصلی[1] است و برای داشتن داده‌های صحیح و باکیفیت نیازمند یکپارچه‌سازی داده‌ها، اطلاعات، صفات و حتی مقادیر صفات است. به‌عنوان نمونه، یکی از مواردی که فقدان یکپارچگی داده‌ها در سیستم‌های مدیریت اطلاعات مبتنی بر روابط نظیر اصطلاحنامه‌ها را به‌دنبال دارد، وجود اطلاعات مشابه

1. master data management

در یک دسته‌بندی مشخص است (مانند شکل ۱). این امر می‌تواند به‌دلیل فقدان کنترل داده‌های مشابه توسط کاربران خبره در حجم انبوهی از اطلاعات ایجاد شود. رفع رکوردهای اطلاعاتی مشابه به‌دلیل به‌کارگیری آن‌ها در سایر نظام‌های بازیابی اطلاعات نظیر نمایه‌سازی در گذر زمان کاری بسیار دشوار، پیچیده، زمان‌بر و برای افراد خبره، فرسایشی است که سرانجام، نتایج دلخواهی را نیز برای سازمان یا کاربر به‌دنبال نخواهد داشت.



شکل ۱. نمونه اطلاعات مشابه (اصطلاح «آب») در دسته‌بندی علوم اسلامی

به دلایل فوق، به‌کارگیری فناوری‌های نرم‌افزاری در این‌گونه ساختارها امری لازم و ضروری محسوب می‌شود. پیش از ورود فناوری اطلاعات، ساخت و نگهداری به‌روز اصطلاحنامه‌ها یکی از چالش‌های حوزه سازماندهی و پردازش اطلاعات بود. نرم‌افزارهای مدیریت اصطلاحنامه از سال ۱۳۴۵ تاکنون در داخل و خارج از کشور پوشش درست و مشخصی نسبت به رعایت قوانین مورد نیاز جهت کنترل اصطلاحات در زمان تجمیع، تلفیق و ادغام رکوردهای اطلاعاتی و تأثیرپذیری آن‌ها در سایر بخش‌های مورد استفاده ارائه نداده‌اند. بررسی‌ها نشان از آن دارد که در حوزه ادغام اطلاعات، پژوهشگران تاکنون مدل‌های مختلفی را نظیر مدل پایه مدیران مشترک آزمایشگاه‌ها ارائه داده‌اند. این مدل‌ها در بافت‌های مختلفی نظیر نظامی، پزشکی، مخابراتی، و فضای مجازی به‌کار گرفته شده‌اند. توسعه مدل‌های ادغام و ترکیب را می‌توان به پیشرفت‌های فناوری‌های محاسباتی، نظیر یادگیری ماشین، هوش مصنوعی و فناوری‌های معنایی نسبت داد. این فناوری‌ها امکان ایجاد مدل‌هایی را فراهم کرده‌اند که می‌توانند به‌طور مؤثر داده‌ها را از منابع متعدد، ادغام و تجزیه‌وتحلیل کنند. این

امر منجر به بهبود تصمیم‌گیری و حل مسئله خواهد شد. افزون بر این، ظهور اینترنت اشیا و تکثیر دستگاه‌های دیجیتال منجر به افزایش قابل توجهی در حجم و تنوع داده‌ها و اطلاعات شده است.

مدل‌های کنونی ادغام اطلاعات، قابلیت کاربست و پیاده‌سازی در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای را دارند؛ زیرا در جریان مدیریت اصطلاحات اصطلاحنامه‌ای مباحث تجمیع، ادغام و تلفیق دخیل هستند. اکثر مدل‌های ادغام به‌منظور خودکارسازی فرایندها برای پشتیبانی از تصمیم استفاده می‌شوند و در مسائلی که نیازمند استنباط و تصمیم‌گیری پیچیده‌ای هستند، عملکرد مناسبی از خود نشان نمی‌دهند (رشیدی ۱۳۹۵؛ Raol 2009). مدل‌های موجود ادغام و ظرفیت‌های آن‌ها با توجه به مقتضیات سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای نیازمند بررسی و تحلیل است. وجود مباحث نحوی و معنایی در سیستم‌های مدیریت اصطلاحنامه از یک سو، و حرکت به سمت نظام‌های معنایی بسیاری از سیستم‌های کنونی مدیریت اطلاعات از سوی دیگر، لزوم توجه به نیازمندی‌ها و مسائل سیستم‌های اصطلاحنامه‌ای را در مدل‌های ادغام اطلاعات دوچندان می‌کند. ادغام اطلاعات در سیستم‌های مبتنی بر اصطلاحنامه شامل ترکیبی از تکنیک‌ها، نظیر تجزیه‌وتحلیل آماری، یادگیری ماشین، پردازش زبان طبیعی، و فناوری‌های معنایی به‌منظور استخراج، تبدیل، و یکپارچه‌سازی داده‌هاست. نتیجه توجه به این امر، نمایشی جامع و دقیق‌تر از اطلاعات اساسی است که می‌تواند برای پشتیبانی از برنامه‌های کاربردی مختلف از جمله بازیابی اطلاعات، مدیریت دانش و تصمیم‌گیری استفاده شود. با توجه به مسائل و مشکلات پیش‌گفته در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای، و همچنین معرفی انواع مدل‌های ادغام اطلاعات و دخیل بودن مواردی نظیر تجمیع و ادغام در فرایند مدیریت اصطلاحنامه‌ها، پژوهش حاضر در پی بررسی مدل‌های ادغام اطلاعات با توجه به کاربست آن‌ها در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای است. مدل‌های ادغام اطلاعات کنونی و همچنین مفهوم ادغام نیز نیازمند تبیین صحیح در کارکرد کنونی به جهت تشابه معنایی و کاربرد در بافت‌های گوناگون است. از این‌رو، پژوهش حاضر ضمن تبیین مفهوم ادغام، اصطلاحات مرتبط و مدل‌های ادغام اطلاعات، این مدل‌ها را از منظر کاربست در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای بررسی و تحلیل می‌کند.

به‌منظور نیل به اهداف پژوهش، پرسش‌های زیر مد نظر قرار گرفت:

۱. مفهوم ادغام و اصطلاحات مرتبط آن نظیر تجمیع، ترکیب و تلفیق در بافت اصطلاحنامه‌ای چیست؟

۲. انواع مدل‌های ادغام اطلاعات کدام‌اند؟
۳. آیا مدل‌های ادغام اطلاعات در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای قابلیت کاربرد دارند؟

۲. پیشینه تحقیق

بررسی پیشینه‌های این حوزه نشان از آن دارد که می‌توان آن‌ها را در دو دسته کلی ارائه کرد. دسته اول، پژوهش‌های مرتبط با اصطلاحنامه و نرم‌افزارهای مدیریت آن به‌عنوان سیستم‌های مدیریت اطلاعات رابطه‌مدار، و دسته دیگر پژوهش‌های مرتبط با کاربرد و مدل‌های ادغام داده/ اطلاعات است. در دسته اول، پژوهش‌های گوناگونی از ارزیابی نرم‌افزارهای مدیریت اصطلاحنامه و تعیین معیارهای ارزیابی تا چگونگی تأثیرپذیری ورود اصطلاحات جدید، و مسائل و مشکلات آن‌ها وجود دارد. در ادامه، به برخی از آن‌ها اشاره شده است.

«میرزابیگی و حری» در پژوهشی به بررسی تأثیرپذیری اصطلاحنامه‌ها از ورود اصطلاحات جدید در اصطلاحنامه «اریک» پرداختند. یافته‌ها نشان داد که نسبت این‌گونه تغییرات در اصطلاحنامه «اریک» با توجه به حذف و تغییر برخی روابط، بسیار اندک است. عمده‌ترین تأثیر ورود اصطلاحات در افزودن حجم وسیعی از روابط در هر سه ساختار اصطلاحنامه، به‌ویژه در ساختار هم‌ارزی است. کمترین تأثیر در ساختار همبسته در تعداد فراوان آن‌ها مشاهده شد و ساختار سلسله‌مراتبی در تعداد کم آن‌ها، اما بیشتر از ساختار همبسته تحت تأثیر قرار گرفت. به‌طور کلی، صِرف وجود روابط زیاد یا کم دلیلی بر تأثیرپذیری کم یا زیاد آن‌ها از ورود اصطلاحات جدید نبود (۱۳۸۵). «ولی‌نژاد» و همکاران با توجه به فقدان خوشه طب سنتی ایران در ساختار اَبراصطلاحنامه نظام زبان واحد پزشکی[1] به بررسی ورود و ادغام این خوشه پرداختند. نتایج این ادغام، تخصیص یك شناسگر واحد مفهوم، دو اصطلاح مترادف با شناسگرهاي واحد اصطلاح، دو مفهوم والد و نیا، نُه مفهوم که به همراه سرشاخه طب سنتی ایران به‌عنوان مفاهیم فرزند دو مفهوم والد و نیای فوق بودند، هجده مفهوم هم‌نیا، شش مفهوم اخص و پنج مفهوم مرتبط دیگر برای سرشاخه پیشنهادی طب سنتی ایران یافت شد و گونه معنایی «رشته یا حرفه زیست‌پزشکی» بود (۱۳۸۷).

1. unified medical language system (UMLS)

«پوراسدالهی‌نژاد» نیز از دو بعد عملکردی و عمومی پنج نرم‌افزار ساخت و مدیریت اصطلاحنامه‌ای علوم اسلامی، تزاروس‌بیلدر، پارس‌آذرخش، قاموس، و نسخه تولیدی «ابراهیم‌پور» را مورد بررسی قرار داد. نتایج نشان داد که در معیار عملکردی و عمومی، نرم‌افزار تزاروس‌بیلدر دارای عملکرد بهتری نسبت به سایرین بود (۱۳۸۸). در پژوهش دیگری، پنج نرم‌افزار ساخت و مدیریت اصطلاحنامه (علوم اسلامی، پارس‌آذرخش، تزاروس بیلدر، قاموس و نوسا) توسط «حسن‌زاده، محمدخانی و پاک‌نژاد» از منظر سه معیار عمومی، عملکردی، و مدیریتی مورد بررسی قرار گرفت. یافته‌ها نشانگر تطابق نرم‌افزارها با معیارهای عملکردی بیش از دو معیار دیگر بود. نرم‌افزار پارس‌آذرخش نسبت به سایر نرم‌افزارها بیشترین میزان رعایت معیارها را داشت (۱۳۹۰).

«گانزمان» در پژوهشی به بررسی معیارهای ارزیابی نرم‌افزارهای اصطلاحنامه پرداخت. بررسی‌های وی نشان داد که کنترل چند مدخل از یک اصطلاح مشابه (تکراری) در زمان ورود یا ویرایش مدخل جدید از الزامات سیستم‌های اصطلاحنامه است. وی بی‌توجهی برخی از سیستم‌های مدیریت اصطلاحنامه‌ای از قوانین و شروط مورد نیاز در کنترل ورودی‌ها را یکی از مشکلات آن‌ها دانست. وی معیارهای ارزیابی نرم‌افزارهای اصطلاحنامه‌ای را در پنج دسته کلی «ساخت و نگهداری»، «خروجی»، «نمایه‌سازی و بازیابی»، و «مبادله»، «یکپارچگی و سازگاری» تقسیم‌بندی کرد (Ganzmann 1990). همچنین، تحلیل و ارزیابی نُه نرم‌افزار اصطلاحنامه[1] توسط «هابیتزل و گرگور» در مقاله‌ای با عنوان «تجزیه‌و تحلیل و طرح کلی نظام‌های نمایه‌سازی» نشان داد که معیارهای مدیریتی و عملکردی نرم‌افزارها به جهت رهگیری تغییرات مفاهیم اصطلاحات در فرایند مدیریت اصطلاحنامه باید مورد بازنگری دقیق‌تری قرار گیرد (Habitzel and Gregor 1999).

پژوهشی دیگر با عنوان «مدل‌های داده برای پایگاه داده اصطلاحنامه یکپارچه» توسط «سورگل» انجام شد. وی به ارائه دو مدل برای ذخیره داده‌ها در اصطلاحنامه‌های چندگانه در یک پایگاه داده یکپارچه برای ساخت و نگهداری اصطلاحنامه‌های فردی پرداخت. این پژوهش سعی کرده مشکل اصطلاحاتی را که از نظر مفهومی با یکدیگر یکسان هستند و در پایگاه داده‌های مختلف نگهداری می‌شوند، به‌نوعی حل نماید. وی نرم‌افزار یا روش‌های هوشمند سیستمی ارائه نکرده، اما از منظر نظری، ساختاری مورد نظر

1. BEAT, Carbox, Cindex, Dtseach, Multithes, Stride, TAT, TermTree, Adlib

را جهت حفظ یکپارچگی پیشنهاد کرد (Soergel 2001). همچنین «ویل» ده‌ها نرم‌افزار مدیریت اصطلاحنامه موجود را مورد بررسی قرار داد. با بررسی انجام‌شده بر روی امکانات معرفی‌شده نرم‌افزارها، تنها شماری از این نرم‌افزارها نظیر آی‌بی‌ام[1] با سکوی InfoSphere و TopBraid Enterprise Vocabulary Net (TopBraid EVN)، OObian، Mondeca ITM T3 و IC index مباحث تطبیق و ادغام اصطلاحات را به‌نوعی در داخل خود در نظر گرفته‌اند. رویکرد اکثر این نرم‌افزارها بیشتر در ادغام دو اصطلاحنامه کنترل‌شده از یک یا چند زبان متفاوت است (Will 2013).

در زمینه مدل‌های ادغام داده / اطلاعات پژوهش‌های چندی در زمینه‌های مختلف صورت گرفته است. پژوهش‌هایی به کاربرد ادغام داده / اطلاعات در موضوعاتی نظیر راه‌حل‌های ادغام داده چندحسگری در تولید هوشمند توسط Tsanousa et al. (2022)، سیستم‌های ماهواره‌ای پیش‌بینی آب‌وهوا به کمک الگوریتم‌های یادگیری ماشین توسط Mokhtari and Akhoondzadeh (2021)، اینترنت اشیا توسط Krishnamurthi et al. (2020) و Qi et al. (2020)، سامانه‌های حمل‌ونقل هوشمند توسط El Faouzi, Leung and Kurian (2011)، شناسایی در جنگ و امور نظامی توسط Schick, Hunter and Wilson (2009)، شبکه‌های حسگر بی‌سیم توسط Nakamura, Loureiro and Frery (2007)، امور دفاعی و هوافضایی توسط Harris, Bailey and Dodd (1998) و جز این‌ها انجام شده است. بخش عمده پژوهش‌ها در این حوزه قرار دارند. همچنین برخی دیگر از پژوهش‌ها به تبیین مفهوم، روش‌ها و فنون ادغام مبادرت کردند. Beddar-Wiesing and Bieshaar (2020) به تبیین دقیق و روشن مفهوم ادغام، عناصر و سطوح آن جهت کاربرد در بافت‌های گوناگون و همچنین فنون ادغام متداول پرداختند. «کاستاندو» نیز به مرور پژوهش‌های حوزه ادغام داده / اطلاعات، فنون و الگوریتم‌های آن پرداخت. وی فنون و الگوریتم‌های ادغام داده‌ها را در سه دسته وابستگی داده[2]، تخمین حالات، ادغام تصمیم دسته‌بندی کرد (Castanedo 2013).

«اشتاینبرگ، باومن، و وایت» در پژوهشی به بیان تلاش‌های جاری در زمینه بازنگری و گسترش مدل ادغام اطلاعات JDL جهت تسهیل توسعه مقرون‌به‌صرفه، فراهم‌آوری، یکپارچه‌سازی و بهره‌بردای از سیستم‌های چندحسگر / چند منبع پرداختند. بخش عمده بازنگری‌های مدل JDL بر گسترش مدل کارکردی، بسط به حوزه‌های غیرنظامی، ترکیب

1. IBM
2. data association

مـدل معمـاری درختـی ادغـام داده[1] بـرای توصیـف، طراحـی و توسـعه سیسـتم‌ها متمرکـز بـود (Steinberg, Bowman and White 1999).

در حـوزه توجـه بـه مسـائل معنایـی و نحـوی، پژوهش‌هـای بسـیار اندکـی صـورت گرفتـه اسـت. «علیـزاده نوقابـی، کاهانـی و به‌کمـال» در پژوهشـی بـه ارائـه مدلـی موسـوم بـه SemFus[2] جهـت ادغـام معنایـی مبتنـی بـر مـدل JDL پرداختنـد. یکـی از مشـکلات مـدل پایـه JDL عـدم توجـه بـه مباحـث نحـوی و معنایـی اسـت. تضادهـای معنایـی زمانـی رخ می‌دهـد کـه دو بافـت گوناگـون از یـک اطلاعـات، تفسـیر مشـابهی نداشـته باشـند. به‌کارگیـری سیسـتم‌های معنایـی نظیـر هستی‌شناسـی‌ها می‌توانـد مشـکلات تضادهـای معنایـی در سیسـتم‌های مدیریـت اطلاعـات را برطـرف کنـد. مـدل SemFus بـرای غلبـه بـر مشـکلات معنایـی در سیسـتم‌های ناهمگـن پیشـنهاد شـده اسـت (Alizadeh Noughabi, Kahani and Behkamal 2013).

مرور پیشینه‌ها نشان داد که بخش عمده آن‌ها به کاربرد ادغام داده / اطلاعات در بافت‌ها و سامانه‌های مختلف پرداخته‌اند. در پژوهش‌های حوزه اصطلاحنامه و نرم‌افزارهای مدیریت آن، «گانزمان» اشاره‌ای ضمنی به توجه به مباحث ادغام در ارزیابی نرم‌افزارهای مدیریت اصطلاحنامه داشت (Ganzman 1990). «ویل» نیز برخی از نرم‌افزارهای مدیریت اصطلاحنامه را که به بحث ادغام داده / اطلاعات توجه داشتند، برشمرد. یکی از دلایل نپرداختن در سیستم‌های مبتنی بر روابط نظیر اصطلاحنامه‌ها، وجود روابط چندگانه میان اصطلاحات و به‌کارگیری آن‌ها در سامانه‌های تولید و بازنمایی اطلاعات است. بر همین اساس، ادغام اطلاعات در سیستم‌های مبتنی بر اصطلاحنامه با توجه به وجود مواردی نظیر تجمیع، تلفیق و ادغام اطلاعات با یکدیگر، نیازمند پیاده‌سازی مدل‌های ادغام اطلاعات در آن‌هاست. پژوهش حاضر با توجه به خلأ موجود به بررسی مدل‌های ادغام داده / اطلاعات از منظر کاربست در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای مبادرت کرد (Will 2013).

۳. روش پژوهش

پژوهـش حاضـر به‌علـت تبییـن مفهـوم و بررسـی مدل‌هـای ادغـام داده/ اطلاعـات و کاربسـت آن در سیسـتم‌های مدیریـت اطلاعـات مبتنـی بـر روابـط از نـوع مفهومـی اسـت. بـه گفتـه «طاهـری»، پژوهش‌هـای مفهومـی می‌تواننـد مبنـای بسـیاری از ایده‌هـا و مباحـث باشـند و از بسـیاری از کجروی‌هـا و اشـتباهات و برداشـت‌های نادرسـت جلوگیـری کننـد، تـا جایـی کـه

1. data fusion tree architecture
2. semantic information fusion system based on JDL (SemFus)

مقالات مفهومی، پژوهش‌های یک حوزه را به سمت درست هدایت می‌کنند (۱۳۹۹). بسط نظریه‌ها، بومی‌سازی مفاهیم و کاربرد در یک بافت خاص، ایجاد زمینه‌های پژوهشی جدید و بازنمون روابط مفهوم یک حوزه با حوزه دیگر از مزایای پژوهش‌های مفهومی است. رویکرد مورد توجه در پژوهش حاضر، تحلیلی بود. همچنین با توجه به بررسی کاربرد مدل‌های ادغام داده / اطلاعات در نرم‌افزارهای مدیریت اصطلاحنامه از نوع کاربردی نیز محسوب می‌شود.

به‌منظور گردآوری ادبیات و داده‌های اولیه پژوهش از روش کتابخانه‌ای (مطالعات اسنادی) با مراجعه به منابع علمی و مقالات هر دو حوزه ادغام داده / اطلاعات و سیستم‌های مدیریت اطلاعات مبتنی بر روابط و اصطلاحنامه استفاده شد. کاوش در وب‌سایت‌های عمومی و تخصصی این حوزه نظیر مجله Information Fusion، «گوگل اسکولار»، «جویشگر علم‌نت» و سایر پایگاه‌های اطلاعات علمی مانند «ساینس دایرکت»[1]، «اشپرینگر»[2] و «محقق معنایی»[3] انجام شده است. این منابع به‌صورت هدفمند انتخاب شد. همچنین از تجربیات سایر سازمان‌ها، شرکت‌ها و جلسات و مصاحبه‌های متعدد کارشناسی استفاده گردید.

بررسی کتب و مقالات معتبر داخلی و خارجی در زمینه ادغام سیستمی اطلاعات با کلیدواژه‌هایی صورت پذیرفته است. این کلیدواژه‌ها در دو دسته مرتبط با اصطلاحنامه و مدیریت آن و دیگری اصطلاحات مرتبط با تجمیع، تلفیق و ادغام داده / اطلاعات بودند که در جدول ۱، ذکر شده است.

جدول ۱. کلیدواژه‌های کاوش‌شده در منابع داخلی و خارجی

کلیدواژه‌های فارسی	کلیدواژه‌های انگلیسی
اصطلاحنامه، نرم‌افزارهای اصطلاحنامه، اصطلاح، ارزیابی اصطلاحنامه، اصطلاحنامه‌نویسی، سایت‌های اصطلاحنامه، نقش اصطلاحنامه در بازیابی، تهیه اصطلاحنامه، استانداردهای اصطلاحنامه‌نویسی	Term, Thesaurus, Thesaurus software's, Term software's, Thesaurus Website
ادغام در اصطلاحنامه، سیستم‌های اطلاعاتی و ادغام، ترکیب اصطلاحات، تأثیر ادغام اطلاعات در بازیابی، ادغام و تلفیق، تجمیع اطلاعات، مدل‌های ادغام داده / اطلاعات، نقش داده‌کاوی در ادغام اطلاعات، مفهوم ادغام داده / اطلاعات	Integration, Contraction, Merges, Information, Integration Data, Data Fusion, Data Fusion Model, Information Fusion, Information Integration, Thesaurus Software Fusion, Thesaurus Software Merges, Term Software Merges

1. www.siencedirect.com 2. https://www.springer.com 3. https://www.semanticscholar.org

٤. یافته‌ها

پرسش اول: مفهوم ادغام و اصطلاحات مرتبط آن نظیر تجمیع، ترکیب و تلفیق در بافت اصطلاحنامه‌ای چیست؟

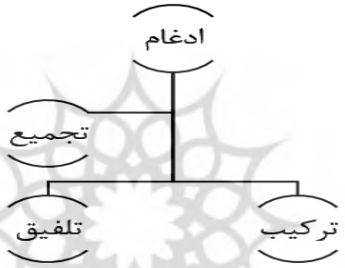
سؤال مهمی که در این زمینه پرسیده می‌شود، این است که آیا افراد خبره یا سازمان‌های متخصص در سیستم‌های مدیریت اطلاعات مبتنی بر روابط، درک یکسانی از مفهوم ادغام اطلاعات دارند؟ ادغام داده / اطلاعات دارای واژه‌های وابسته‌ای نظیر تجمیع، ترکیب یا تلفیق بوده که وابستگی زیادی با مفهوم ادغام دارند. مفهوم ادغام با توجه به محیط به‌کارگیری می‌تواند متفاوت باشد که به‌تبع آن، این تفاوت در واژه‌های جایگزین، مرتبط و وابسته نیز وجود دارد. به‌عنوان مثال، در حوزه رایانه در برخی از نرم‌افزارها نظیر محصولات شرکت «مایکروسافت آفیس» از اصطلاح Merge (Microsoft 2023)، در حوزه ادبیات از اصطلاح Umlaut، و در حوزه‌های نظامی و سایبری از اصطلاح Fusion برای توصیف مفهوم ادغام استفاده شده است. واژه‌های جایگزین، مرتبط و وابسته با مفهوم ادغام در ادبیات این حوزه مشتمل بر Integration، Interaction، Collaboration، Partnership، Synergy، Acting together، Composition, Cooperation، Merge، Fusion، Assimilation، Diphthong، Umlaut، و Affricate است (آبادیس، بی‌تا). برای واژه تجمیع، ترکیب و تلفیق نیز همین شرایط حاکم است. برای درک جنبه‌های مختلف تجمیع، تلفیق و ادغام با همکاری ماشین و افراد خبره، دانش پایه از معماری این‌گونه سیستم‌ها و روند کلی و نحوه نقش‌آفرینی آن درون یک سیستم، امری الزامی است (لامعی ۱۳۸۶). تنها زمانی که از ادغام و واژگان مرتبط به آن مانند تجمیع، ترکیب و تلفیق یك تعریف روشن و فنی ارائه شود، می‌توان بر مبنای آن تعریف، ماهیت و چیستی و فرایند ادغام سیستمی را تبیین نمود.

از منظر لغوی در «فرهنگ وبستر»، «ادغام» را تشکیل، هماهنگی یا ترکیب [اجزا] به‌صورت یک کلِ یکپارچه یا یک کلِ کاری تعریف کرده‌اند. همچنین «فرهنگ لغات آکسفورد»، «ادغام» را عمل یا فرایند ترکیب دو یا چند چیز تعریف می‌کند، به‌طوری که این‌ها باهم کار کنند؛ ترکیب داده‌هایی که از منابع مختلف به روشی محسوس به دست می‌آیند. تعریف ادغام اطلاعات در صفحه اصلی انجمن بین‌المللی ادغام اطلاعات[1] به این شکل ذکر شده است: تلفیق اطلاعات شامل تئوری، تکنیک‌ها و ابزارهایی است که برای

1. http://www.inforfusion.org/

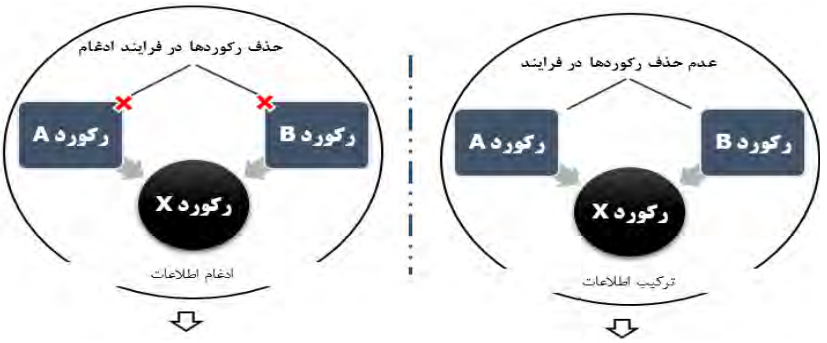
بهره‌برداری از هم‌افزایی در اطلاعاتی که از منابع مختلف (حسگر، پایگاه داده‌ها، اطلاعات جمع‌آوری‌شده توسط انسان و غیره) گرفته شده، مورد استفاده قرار می‌گیرد. به‌طوری که تصمیم یا اقدامات صورت‌گرفته در نتیجه به نوعی جهت بهبود (کیفی یا از لحاظ دقت، استحکام، و غیره) است.

منظور از ادغام در پژوهش حاضر که مورد توجه در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای است، ادغام سرشاخه خوشه‌ای از یک اصطلاحنامه در اصطلاحنامه دیگر یا یک اَبراصطلاحنامه نیست؛ چنان‌که در پژوهش «ولی‌نژادی» و همکاران (۱۳۸۷) این مورد بررسی شده است. شکل ۲، ساختار مفهوم ادغام در کنار سایر مفاهیم مرتبط و وابسته را منطبق با بافت پژوهش حاضر نشان می‌دهد.



شکل ۲. ساختار مفهوم ادغام در کنار سایر مفاهیم مرتبط و وابسته

واژه ادغام با توجه به شکل ۲، و مفاهیم وابسته به آن در درون خود، بسته به شرایط مختلف، ترکیب و تلفیق اطلاعات را نیز دربردارد. تمایز دو مفهوم ادغام و ترکیب در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای حائز اهمیت است. در فرایند ادغام، مفهوم «ترکیب - تلفیق» و «تجمیع» مفاهیمی هستند که باید مورد توجه قرار گیرند. برای ادغام دو عنصر با یکدیگر، شاید این به ذهن خطور نماید که ابتدا باید اصطلاحات با یکدیگر «ترکیب» گردند و سپس ادغام صورت گیرد.



شکل ۳. نمای کلی ادغام و ترکیب رکوردهای اطلاعاتی

بلوک واژگان در اصطلاحنامه‌ها که به رکورد اطلاعاتی موسوم است، در فرایند ترکیب جایگاه خود را حفظ می‌نمایند و رکورد اطلاعاتی جدید، مستقل و واحد از سایر رکوردهای مشارکت‌کننده ایجاد می‌گردد؛ در صورتی که در فرایند ادغام، رکوردهای اطلاعاتی مشارکت‌کننده حذف می‌شوند و رکورد اطلاعاتی مستقلی ایجاد می‌گردد. رکورد اطلاعاتی جدید از نظر عنوان می‌تواند همانند یکی از رکوردهای اطلاعاتی حذف‌شده در فرایند ادغام باشد. در ترکیب رکوردهای اطلاعاتی، تأثیری بر سایر بخش‌های استفاده‌کننده از اصطلاحات ایجاد نمی‌شود؛ دلیل آن را می‌توان ایجاد رکورد اطلاعاتی جدید دانست؛ چرا که رکورد اطلاعات جدید از ترکیب دو یا چند اصطلاح به‌وجود آمده است و هنوز استفاده نشده و ارتباطاتی نیز با سایر اصطلاحات ندارد. اما در فرایند ادغام سایر بخش‌ها به‌واسطه وجود روابط تعیین‌شده از قبل تحت تأثیر قرار خواهد گرفت. از این‌رو، می‌توان نتیجه گرفت که ادغام سیستمی ماهیتی دوگانه شامل ترکیب و ادغام را در خود جای خواهد داد.

پرسش دوم: انواع مدل‌های ادغام اطلاعات کدام‌اند؟

مدل‌های تلفیق داده و اطلاعات از منابع جداگانه و متفاوتی گردآوری شده تا اطلاعات مرتبطی برای رسیدن به نتیجه مورد نیاز را با توجه به محیطِ مسئله فراهم کنند (Alizadeh Noughabi, Kahani & Behshid 2013). بررسی متون نشان می‌دهد که متداول‌ترین مدل‌های به‌کاررفته در این زمینه در چهار دسته مدل‌های مبتنی بر گردش اطلاعات[1]، مدل‌های گردش کار و فعالیت[2]، مدل‌های مبتنی بر نقش‌ها و عملکرد موجودیت‌ها[3]، و مدل‌های مبتنی بر درک مفاهیم قابل دسته‌بندی هستند که در جدول ۲، ارائه شده‌اند.

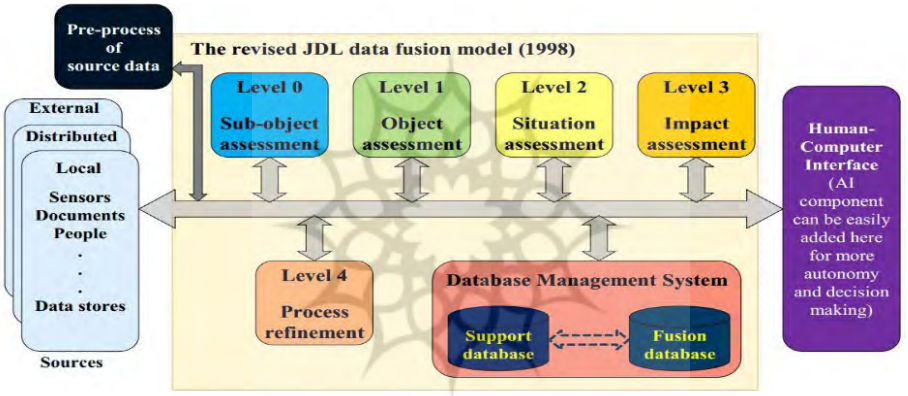
1. information-based models 2. activity-based models 3. role-based models

جدول ۲. انواع مدل‌های ادغام اطلاعات

دسته	مدل‌ها
مدل‌های مبتنی بر گردش اطلاعات	Joint Directors of Laboratories (JDL) Data Fusion Information Group (DFIG) Waterfall Fusion Process (WFFP) Transformations of “requirements for information” process (TRIP) ۸- JDL
مدل‌های گردش کار و فعالیت	Intelligence Cycle (IC) Dasarathy’s (DFD) OODA Loop Boyd control Loop Laboratorie d’Analyse et d’Architecure des Systemes (LAAS) Omnibus Model
مدل‌های مبتنی بر نقش‌ها و عملکرد موجودیت‌ها	Object Oriented Model Endsley’s Mode of Situational Awareness
مدل‌های مبتنی بر درک مفاهیم	Semantic information fusion system based on JDL (SemFus)

مدل‌های مبتنی بر گردش اطلاعات، مراحل ادغام را بر اساس سطوح انتزاع اطلاعات تغییریافته توسط سیستم ادغام بنا می‌کنند. مدل‌های مبتنی بر گردش کار و فعالیت بر اساس فعالیت‌هایی که باید توسط یک سیستم ادغام انجام شود، مشخص می‌شوند. در چنین مدل‌هایی فعالیت‌ها و ترتیب صحیح اجرای آن‌ها به صراحت مشخص شده است. مدل‌های مبتنی بر نقش، نشانگر تغییر تمرکز در نحوه مدل‌سازی و طراحی سیستم‌های ادغام اطلاعات هستند. در چنین مدل‌هایی سیستم‌های ادغام اطلاعات بر اساس نقش‌های ادغام و روابط بین آن‌ها مشخص می‌شوند که مدل دقیق‌تری برای سیستم ادغام نظیر مدل JDL ارائه می‌کنند. مدل‌های مبتنی بر نقش نیز یک دیدگاه سیستمی از ادغام اطلاعات ارائه می‌کنند. با این حال، بر خلاف مدل‌های قبلی، مدل‌های مبتنی بر نقش وظایف یا فعالیت‌های تلفیقی را مشخص نمی‌کنند. در عوض، مجموعه‌ای از نقش‌ها را ارائه می‌دهند و روابط بین آن‌ها را مشخص می‌کنند. در مدل‌های مبتنی بر درک مفاهیم، ادغام اطلاعات با استفاده از ابزارهای معنایی نظیر هستی‌شناسی‌ها صورت می‌پذیرد. در این نوع مدل‌ها بر اساس فناوری‌های وب معنایی، امکان درک مسائل معنایی وجود دارد. با توجه به پایه‌بودن مدل JDL به معرفی و تبیین این مدل و همچنین کاربست آن در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای مبادرت می‌شود.

مدل پایه JDL: این مدل، یکی از مدل‌های بسیار رایج و معروف بوده که با همکاری مشترک مدیران آزمایشگاه‌ها و وزارت دفاع آمریکا در سال‌های ۱۹۹۱ و ۱۹۹۲ طراحی و توسعه یافته است. مدل پایه JDL جهت استفاده در حوزه‌های مختلف نظیر صنایع نظامی و غیرنظامی و بر پایه گردش اطلاعات طراحی شده است (Hall and Llinas 2007). این مدل از پنج سطح برای پردازش داده‌ها به‌ترتیب شامل پیش‌پردازش، تشخیص و شناسایی عناصر و موجودیت‌ها، تجمیع داده‌ها و بررسی وضعیت محیط، ارزیابی تهدیدات و فرصت‌ها، و اصلاح ساختارهای منابع داده‌ای تشکیل شده است. خروجی آن تشخیص و شناسایی موجودیت‌ها، عناصر مهم و حساس در مسئله مورد بررسی است (Hall and Llinas 1997; Elmenreich 2002). شکل ٤، این مدل را به تصویر کشیده است.



شکل ٤. مدل پایه JDL (Devlin 2006; Hall and Llinas 2007)

منابع ورودی از حسگرها، پایگاه‌های داده‌ای مختلفی نظیر منابع توزیع‌شده، اطلاعات دانشی سایر سیستم‌ها، و داده‌های بیرون از مسئله تأمین می‌شود. قبل از پردازش داده‌ها در سطح یک، پیش‌پردازش صورت می‌پذیرد. داده‌های اولیه بنا به مسئله مورد نظر از منابع و سطوح اهمیت متفاوت از محیط پیرامونی گردآوری می‌شود که به لحاظ کیفی و محتوایی شرایط مشابهی ندارند. به همین دلیل، در گام اول و قبل از ورود به پردازش در مدل JDL، به پیش‌پردازش نیاز است تا به لحاظ کیفی و کمی داده‌های اولیه به شرایط مطلوبی برسند. در سطح صفر داده‌ها به شرایطی رسانده می‌شود تا بتوانند طی این مدل برای تشخیص شرایط مسئله و تهدیدات متناظر به‌خوبی مورد استفاده قرار گیرند. سطح یک برای شناسایی عناصر و اشیای مهم در مسئله به‌کار گرفته می‌شود. سطح دو، وظیفه ترکیب دانش مرحله قبل و ارزیابی و وضعیت

کلی محیط پیرامون را بر عهده دارد. سطح سه، شناسایی تهدیدات و فرصت‌هایی را که در محیط وجود دارد، مدیریت می‌کند و سرانجام، سطح چهار کنترلی بر روی سطوح قبلی دارد تا بهینه‌سازی فرایند ترکیب اطلاعات و ارائه کارآمد تشخیص را برای ادغام‌های آینده داشته باشیم. برای این کار می‌توانیم از سیستم‌ها و پایگاه داده‌هایی که دانش قبلی و اطلاعات موجود در آن است، استفاده کرد (Hall and Llinas 2007). در مدل پایه JDL تمامی سطوح می‌توانند به‌صورت موازی در حال تعامل با یکدیگر باشند. برای این کار از یک باس[1] مشترک استفاده می‌کنند. در واقع، نیازی نیست که سطوح پردازشی به‌صورت ترتیبی اجرا شوند. بین پنج سطح پردازشی مدل JDL، سطوح ۱، ۲ و ۳ سطوح اصلی و حساس این مدل محسوب می‌شوند (Jitendra 2009; Salemo 2002).

نقش ادغام اطلاعات در شکل‌گیری و کاربرد مدل JDL بنا بر سنجش نمونه‌های استفاده از آن حدود ۲۵-۳۰ درصد کاری است که در این مدل انجام می‌پذیرد. بقیه مربوط به تحلیل داده‌ها، ادغام و برنامه‌ریزی‌هایی است که عمدتاً بر اساس تحلیل‌های مبتنی بر هوش مصنوعی و شناسایی و تشخیص الگوها و موجودیت‌ها در مسئله صورت می‌گیرد. توسعه مدل پایه JDL در سیستم‌های ادغام اطلاعات در سال‌های بعد و بر مبنای مدل‌های مبتنی بر شناخت یا مبتنی‌بر تصمیم صورت پذیرفته است. این توسعه و گسترش‌ها حسب نیاز و ضرورت‌هایی که در استفاده کارآمد از مدل JDL در طی مسایل و دامنه‌ها که کاربردهای مختلفی دارند، طراحی شده است. این مدل‌ها بر اساس توسعه مدل پایه JDL، سعی در رفع معایب و افزودن قابلیت‌های جدید متناسب با مسئله مورد نظر نموده‌اند. محوریت، تمرکز و گسترش‌های صورت‌گرفته بر روی گردش اطلاعات، طی فرایندهای مختلفی است. این فرایندها به نحوی تعریف شده‌اند که طی آن با گردش اطلاعات خام بتوان در نهایت، به داده‌های پردازش‌شده رسید. در این صورت امکان تصمیم‌گیری در بهترین حالت ممکن با توجه به شرایط مختلف مسئله وجود خواهد داشت.

پرسش سوم: آیا مدل‌های ادغام اطلاعات در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای قابلیت کاربرد دارد؟

مدل پایه JDL بنا به شرایط مسئله، مورد بازنگری و ارائه یک مدل دیگر قرار گرفته است. سایر مدل‌های ادغام داده / اطلاعات سعی در رفع مشکلات و معایب این مدل نظیر

1. bus

محور نبودن فعالیت‌ها یا عدم پوشش کامل شرایط، نیازها و ضرورت‌های مسئله مورد نظر، عدم توجه به رفتارهای چرخه‌ای، فعال‌نبودن/ بازبینی فرایندهای پردازشی و تحلیلی- به معنای فقدان امکان اعمال نتایج تصمیم‌گیری‌ها بر روی محیط مسئله یا اعمال بازخورد آن در مشاهدات بعدی- عدم پشتیبانی از مسائل معناشناختی (Alizadeh Noughabi, Kahani & Behshid 2013) و عدم نقش‌آفرینی درست افراد خبره برآمده‌اند. در این مدل‌ها، قابلیت‌های جدید با توجه به محیط مسئله‌ای خاص از طریق طراحی و پیاده‌سازی مدل‌های خاص اضافه شده است. در حالی که مدل پایه JDL یک رویکرد محبوب و پرکاربرد برای یکپارچه‌سازی اطلاعات از منابع متعدد است، اما با توجه به سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای و همچنین توجه به رویکردهای معنایی دارای معایبی به‌شرح زیر است:

- ◇ محدودیت انعطاف‌پذیری: مدل پایه JDL بر اساس مجموعه‌ای ثابت از سطوح و فرایندها شکل گرفته که ممکن است برای همه برنامه‌ها یا دامنه‌ها مناسب نباشد. این امر می‌تواند سفارشی‌سازی مدل را جهت رفع نیازهای خاص کاربر یا دامنه دشوار کند؛
- ◇ فقدان چارچوب: این مدل بر اساس مجموعه‌ای از دستورالعمل‌ها و اصول است، نه یک چارچوب ریاضی یا آماری رسمی. از این‌رو، تعریف دقیق روابط و تعاملات بین سطوح و فرایندهای مختلف دشوار است؛
- ◇ تمرکز بالا بر جنبه‌های فنی: مدل پایه JDL در درجه اول بر جنبه‌های فنی ترکیب داده‌ها نظیر مدیریت حسگر، پیش‌پردازش داده‌ها و الگوریتم‌های ادغام تمرکز دارد. این مهم می‌تواند ترکیب سایر عوامل مهم نظیر نیازهای کاربر، دانش دامنه و ملاحظات اخلاقی را در فرایند ادغام دشوار کند؛
- ◇ توجه محدود به عدم قطعیت: این مدل به صراحت به عدم قطعیت یا کیفیت منابع داده ورودی نمی‌پردازد که می‌تواند عامل مهمی در فرایند ترکیب داده‌ها باشد. اگر منابع داده ورودی کیفیت پایینی داشته باشند یا دارای خطا باشند، منجر به نتایج ترکیبی نادرست یا غیرقابل اعتماد می‌شود؛
- ◇ عدم مقیاس‌پذیری: این مدل به‌احتمال، برای برنامه‌های کاربردی ترکیب داده‌های مقیاس بزرگ که شامل تعداد زیادی از منابع داده و تعاملات پیچیده بین آن‌هاست، مقیاس‌پذیر نخواهد بود. این امر می‌تواند کاربرد مدل JDL را در برنامه‌های کاربردی دنیای واقعی که به الگوریتم‌های ترکیب داده‌های مقیاس‌پذیر و کارآمد نیاز دارند، دشوار کند.

در مدل‌های مبتنی بر روابط نظیر اصطلاحنامه، مدل JDL به‌دلیل تخصصی‌بودن ساختارهای علم و نیروی متخصص نقش مهمی ایفا می‌کند که منجر به یادگیری در سیستم می‌شود. خروجی در این مدل از طریق واسط رابط کاربری در هر مرحله در اختیار کاربر قرار می‌گیرد. کاربر می‌تواند از این واسط‌ها تصمیم‌گیری‌هایی را در رابطه با عملکرد مطلوب در شرایط تشخیصی و در مسئله مورد نظر استفاده نماید. همچنین در این مدل، اطلاعات بر روی یک باس با یکدیگر به اشتراک گذاشته شده و به‌صورت ترتیبی نیاز به اجرای آن‌ها نیست. در مدل‌های رابطه‌ای نیز می‌توان متناسب با شاخص‌هایی نظیر نوع روابط، وجود روابط زیاد و غیره مشخص نماییم که کدام‌یک از داده‌ها در ابتدا و با توجه به چه شاخصی باید با یکدیگر پردازش و در صورت نیاز ادغام شوند. به‌طور کلی، در حالی که مدل پایه JDL چارچوب مفیدی برای ادغام داده / اطلاعات ارائه می‌کند، اما محدودیت‌هایی نیز دارد که باید هنگام اعمال آن در برنامه‌های دنیای واقعی به‌دقت در نظر گرفته شوند. در ادامه، ضمن معرفی مختصر هر مدل به تبیین و تشریح آن از منظر کاربست در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای مبادرت شده است:

مدل WFFP[1]: این مدل سعی دارد اشکال عمده مدل JDL را که عدم توجه به رفتارهای چرخه‌ای و بازبینی فرایندهای پردازشی و تحلیلی است، با یکسری بازگشت به مراحل قبل و چرخش بین فرایندهای مختلف آبشاری جبران نماید. مبنا در مدل WFFP بر اساس پردازش‌هایی که روی داده‌ها و نتایج تحلیل‌های آن‌ها در سطوح مختلف یا گاهی در سطوح پایین مدل JDL اتفاق می‌افتد، در نظر گرفته می‌شود (Markin et al. 1997; Bedworth and O'Brien 2000). برگشت به عقب در فرایندهای این مدل در سطوح مختلف می‌تواند به بهینه‌سازی اطلاعات کمک شایانی نماید. به‌عنوان نمونه بر حسب ارزیابی‌هایی که روی داده‌ها و موجودیت‌های مسئله صورت می‌گیرد، امکان اصلاح فرایند پردازش داده‌های خام وجود خواهد داشت یا بر حسب کنش‌هایی که در محیط مسئله اتفاق می‌افتد، تشخیص موجودیت‌ها و روابط بین آن‌ها مورد بازبینی و اصطلاح قرار می‌گیرد. این کار باعث می‌شود که برای انجام پردازش‌های بعدی از جمله تعیین و انتخاب حسگرهای متناظر با نیازهای مسئله و پردازش‌هایی که روی سیگنال‌های خام صورت می‌گیرد و تشخیص موجودیت‌ها، ادغام اطلاعات را جهت تشخیص وضعیت محیطی انجام داد که سرانجام، تمامی تصمیم‌گیری‌ها

1. waterfall fusion process (WFFP)

بـر اسـاس فهرسـت نیازمندی‌هـای اطلاعاتـی اولیـه جهـت پیـدا خواهـد کـرد. در گام آخـر نیـز انتخـاب تصمیمـات مناسـب بـر اسـاس داده‌هـای دریافتـی، مبنـا و محـور پـردازش قـرار می‌گیـرد.

یکـی از معایـب اصلـی ایـن مـدل آن اسـت کـه بـرای مدیریـت منابـع داده پویـا مناسـب نیسـت؛ زیـرا ایـن مـدل متکـی بـه یـک فراینـد متوالـی از ترکیـب داده‌هاسـت کـه می‌توانـد زمان‌بـر بـوده و نتوانـد بـا سـرعت تغییـرات در منابـع داده هماهنـگ شـود. عیـب دیگـر در نظـر نگرفتـن عـدم قطعیـت در منابـع داده ورودی اسـت. ایـن امـر منجـر بـه نتایـج نادرسـت و غیرقابـل اعتمـاد می‌شـود، به‌ویـژه در شـرایطی کـه منابـع داده ورودی مسـتعد خطـا هسـتند. افـزون بـر ایـن، مـدل WFFP بـرای مدیریـت منابـع داده‌هـای بـزرگ و پیچیـده مقیاس‌پذیـر مناسـب نیسـت. بـا افزایـش تعـداد منابـع داده ورودی، پیاده‌سـازی مـدل دشـوارتر بـوده و بـه منابـع محاسـباتی قابـل توجهـی بـرای کار نیـاز دارد.

مـدل (TRIP)[1]**:** مـدل ادغـام TRIP، به‌عنـوان فراینـد ادغـام سـه‌گانه نیـز شـناخته می‌شـود. ایـن مـدل داده‌هـا را از سـه منبـع مختلـف بـا اسـتفاده از یـک فراینـد متوالـی ادغـام می‌کنـد. اولیـن گام در مـدل TRIP، پیش‌پـردازش داده‌هـا (از منابـع مختلـف) اسـت تـا از سـازگاری قالـب آن‌هـا اطمینـان پیـدا شـود. سـپس، داده‌هـا بـا اسـتفاده از روش رأی اکثریـت بـرای ایجـاد یـک مجموعـه داده ترکیبـی اولیـه ادغـام می‌شـوند. در مرحلـه دوم، مجموعـه داده‌هـای ادغام‌شـده اولیـه بـا اسـتفاده از یـک روش آمـاری، نظیـر تجزیه‌وتحلیـل مؤلفه‌هـای اصلـی یـا تحلیـل عاملـی، بـرای کاهـش افزونگـی و نوفـه در داده‌هـا، بیشـتر پـردازش می‌شـوند. ایـن عمـل منجربـه یـک مجموعـه داده ادغام‌شـده می‌شـود کـه دقیـق و قابـل اعتمادتـر از مجموعـه داده‌هـای ادغام‌شـده اولیـه اسـت. در مرحلـه آخـر، مجموعـه داده‌هـای ادغام‌شـده را بـا اسـتفاده از یـک الگوریتـم یادگیـری ماشـین، نظیـر شـبکه عصبـی یـا درخـت تصمیـم، پـردازش کـرده تـا الگوهـا از داده‌هـا اسـتخراج شـود. آنـگاه امـکان پیش‌بینـی یـا اطلاع‌رسـانی در تصمیم‌گیـری از الگوهـا فراهـم می‌شـود. افـزودن اطلاعـات اولیـه یـا نیازمندی‌هـای اولیـه[2] یـا کلان مسـئله از مهم‌تریـن بخش‌هـای مـدل TRIP اسـت. بـا افـزودن نیازهـای کاربـر تـوان تصمیم‌گیـری مسـئله مـورد نظـر افزایـش می‌یابـد. در حالـی کـه مـدل TRIP دارای مزایایـی نظیـر توانایـی در مدیریـت چندیـن منبـع داده و رویکـرد پـردازش متوالـی اسـت، امـا دارای معایـب ضعـف در محاسـبات فشـرده، حساسـیت بـه کیفیـت داده، انعطاف‌پذیـری محـدود، پیچیدگـی و فقـدان شـفافیت، و نیـاز بـه تخصـص دامنـه اسـت.

دو مـدل WFFP و TRIP سـعی در حـل مشـکل رویکـرد برگشـت بـه عقـب و بازنگـری در

1. Transformations of “Requirements for Information” Process (TRIP)
2.User Requirement (UR)

ساختارهای اطلاعات و نحوه به‌کارگیری داده‌ها برای حل مشکل مدل پایه JDL دارند. این امر در اصطلاحنامه‌ها که ماهیت پویایی دارند، امری الزام‌آور و حیاتی است؛ چرا که نوع اصطلاحات در گذر زمان و روابط آن می‌تواند دست‌خوش تغییرات زیادی گردد. به همین دلیل فرایند کنترل و پردازش داده‌های خام در زمان ثبت اطلاعات جدید می‌تواند از ایجاد داده‌های تکراری در همان ابتدا جلوگیری نماید تا سیستم نیاز به پردازش حجم عظیمی از اطلاعات و روابط بین آن‌ها نداشته باشد.

مدل Jdlc -۸[1]: مدل ادغام Jdlc -۸ رویکردی به‌نسبت جدید به ادغام داده‌ها دارد که بر اساس مدل JDL طراحی شده است. تمرکز این مدل بر نتایج پردازش‌های اولیه و تعیین راهکارهاست. ترکیب اطلاعات در این مدل به‌عنوان ترکیبی از ارزیابی و شناسایی اشیا و موجودیت‌ها و همین‌طور شناسایی وضعیت‌ها که در عمل، روابط بین موجودیت‌ها در کنار اطلاعات ثانویه است، مورد استفاده قرار می‌گیرد. در این مدل تمامی سطوح بر روی یک باس، داده‌ها و نتایج پردازشی خود را به اشتراک می‌گذارد.

به‌منظور ادغام اطلاعات در سیستم‌های مدیریت اطلاعات مبتنی بر روابط، افزون بر اینکه اطلاعات را باید بر روی یک باس برای پردازش و اولویت‌بندی قرار داد، از طریق ضرب دکارتین[2] این اطلاعات را باید در یک ماتریس جایگذاری کرد و اشتراکات آن‌ها را برای ادغام مورد بررسی قرار داد. به‌طور خلاصه، مدل Jdlc -۸ رویکردی امیدوارکننده برای ادغام داده‌های مکانی بوده، اما بدون معایب بالقوه آن نیست. از معایب این مدل با توجه به کاربست آن در سیستم‌های مدیریت اطلاعات مبتنی بر روابط می‌توان به چالش‌های یکپارچه‌سازی، کاربرد محدود، نیاز به تخصص دامنه و مشکلات اعتبارسنجی اشاره کرد. سازمان‌هایی که در نظر دارند از این مدل استفاده کنند باید تناسب آن را برای نیازهای خود به‌دقت ارزیابی کنند و منابع و تخصص مورد نیاز برای اجرای مؤثر آن را در نظر بگیرند.

مدل IC[3]: مدل ادغام IC یک مدل مبتنی بر فعالیت است و رویکردی جهت ترکیب داده‌ها در برنامه‌های اطلاعاتی و امنیتی است. این مدل سعی دارد اشکال عمده مدل JDL را که عدم توجه به رفتارهای چرخه‌ای و بازبینی فرایندهای پردازشی و تحلیلی است، به‌گونه‌ای جبران نماید. پنج مرحله اصلی این مدل مشتمل بر برنامه‌ریزی و تعیین جهت در پردازش‌ها

1. joint directors of laboratories cycle (JDLC) 2. Cartesian 3. Intelligence Cycle (IC)

و محاسبات بر حسب نیازها، گردآوری داده‌های خام از حسگرهای مختلف و منابع انسانی، تجمیع و تلفیق داده‌های گردآوری‌شده در مرحله قبل، ارزیابی و تجزیه‌وتحلیل نتایج حاصل از تجمیع / تلفیق داده‌ها، و در نهایت، اعلام و انتشار نتایج برای گرفتن تصمیمات مورد نیاز مسئله است. به این مدل می‌توان معایبی نظیر کاربرد محدود، تخصص بالا، چالش کیفیت و یکپارچه‌سازی داده، وجود عدم قطعیت و نیازمندی‌های منابع اشاره کرد.

برای کاربست این مدل در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای، متخصصان باید روابط بین اطلاعات را پردازش و ایجاد نمایند و ماشین نمی‌تواند در این مرحله کمک شایانی انجام دهد. مراحل یک و دو در این مدل زمانی مورد استفاده قرار می‌گیرد که بخواهیم داده‌ها را از منابع مختلف مورد بررسی قرار دهیم و بعد از تبدیل داده‌ها به اطلاعات اصطلاحنامه‌ای (اصطلاحات)، روابط بین آن‌ها توسط ماشین پیشنهاد خواهد شد. سرانجام، با توجه به داده‌های موجود میزان تشابه اصطلاحات، شناسایی و توسط نیروی متخصص نتایج اعلام، اعمال و انتشار می‌یابد.

مدل Dassarath's یا DFD[1]: مدل «داساراتی» یا مدل داده - ویژگی - تصمیم یک مدل مبتنی بر فعالیت نظیر مدل IC است و از جزئیات و ساختار آن تبعیت می‌نماید. این مدل، سه سطح ادغام سطح پایین، سطح میانی و سطح بالا را در خود جای داده است. در سطح صفر دریافت اطلاعات به‌صورت داده‌های خام چه از طریق حسگرهای پایه محیطی یا از طریق منابع ثانویه انجام می‌گیرد. در سطح دوم بر اساس داده‌های سازگارِ شناسایی‌شده مرحله قبل، به تخمین و پیش‌بینی موجودیت‌های مسئله مورد نظر پرداخته می‌شود و در عمل، پردازش داده‌ها در این سطح انجام می‌شود. در سطح سوم تخمین و پیش‌بینی روابط بین موجودیت‌ها انجام می‌شود (Dassarathy 1997; Raol 2009). توانایی مدل جهت نمایش جریان‌های داده‌ای و فرایندهای ادغام پیچیده از مزایای آن محسوب می‌شود.

این مدل به‌دلیل پردازش اطلاعات در مرحله اولیه می‌تواند در کشف خودکار اصطلاحات و روابط بین آن‌ها، به محققان و پژوهشگران کمک نموده و در شناسایی اولیه داده‌هایی که می‌توانند تبدیل به اطلاعات اصطلاحنامه‌ای شوند، راهگشای خوبی باشد. از معایب این مدل با توجه به کاربست در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای می‌توان به مواردی نظیر راهنمایی محدود در پیاده‌سازی فرایند ادغام، تمرکز زیاد بر جنبه‌های فنی ترکیب

1. Data Fusion Diagram (DFD)

داده‌ها، کاربرد محدود، فقدان توجه به عدم قطعیت، و مقیاس‌پذیری محدود اشاره کرد.

مدل Boyd's OODA: مدل/حلقه Boyd's OODA[1] یک مدل ادغام داده به‌خودی خود نیست، بلکه یک چارچوب تصمیم‌گیری است که می‌تواند در ارتباط با روش‌های ترکیب داده‌ها استفاده شود. این مدل رویکردی محبوب و پرکاربرد برای یکپارچه‌سازی اطلاعات از منابع متعدد است (Boyd 2018). «بوید» یک مدل چرخه‌ای چهارمرحله‌ای بوده و شامل مشاهده، ادغام، تصمیم‌گیری و عمل است. تفاوت مدل Boyd's با JDL این است که مدل Boyd's یک مدل مبتنی بر فعالیت است که نتایج تصمیم‌گیری‌ها را می‌توان روی سیستم و روی محیط مسئله اعمال کرد و بازخورد آن را در مشاهدات بعدی داشت؛ در صورتی که در مدل غیرفعال JDL، این کار صورت نمی‌پذیرفت. یکی دیگر از ویژگی‌های این مدل وجود ساختار چرخشی در آن است که یکی از نواقص پایه مدل JDL محسوب می‌شود. انتقاداتی نظیر کاربرد محدود، فقدان توجه به عدم قطعیت و مقیاس‌پذیری محدود به این مدل وارد است. همچنین این مدل بیشتر بر کاربردهای نظامی تأکید دارد. از این‌رو، ممکن است برخی از مفاهیم و اصطلاحات آن به‌طور مستقیم برای سایر دامنه‌ها یا کاربردها قابل استفاده نباشد. این امر می‌تواند تطبیق چارچوب با سایر زمینه‌ها را دشوار سازد.

مدل LAAS[2] (سیستم تطبیقی و خودمختار محلی): LAAS یک مدل ادغام داده است که از مدل‌های محلی برای ادغام اطلاعات از چندین منبع استفاده می‌کند. این مدل در حالی که یک ابزار خوب برای تقسیم سیستم‌های بزرگ به ماژول‌ها را فراهم می‌کند، اما ارتباطات زمان واقعی و نمایش داده‌ها را در سطوح بالاتر از سطح عملکردی فراهم نمی‌کند. در مقایسه با مدل پایه JDL، معماری LAAS، طراحی خوبی را در اجرای ماژول‌های قابل استفاده مجدد به‌عنوان بخشی از یک برنامه واقعی مورد استفاده قرار می‌دهد. این مدل دارای مزایایی نظیر قابلیت سازگاری با محیط‌های در حال تغییر و استفاده از مدل‌های محلی برای کاهش پیچیدگی محاسباتی دارد. مدل LAAS مبتنی بر مجموعه‌ای از مدل‌های محلی است که بر روی منابع داده‌ای خاص آموزش داده شده‌اند که می‌توانند توانایی مدل را برای تعمیم به منابع داده جدید محدود کند. هنگامی که منابع داده ورودی به‌خوبی توسط مدل‌های محلی نمایش داده نمی‌شوند، این مدل می‌تواند منجر به نتایج ترکیبی نادرست یا غیرقابل اعتماد شود.

1. observe, orient, decide, act
2. local adaptive and autonomous system (LAAS)

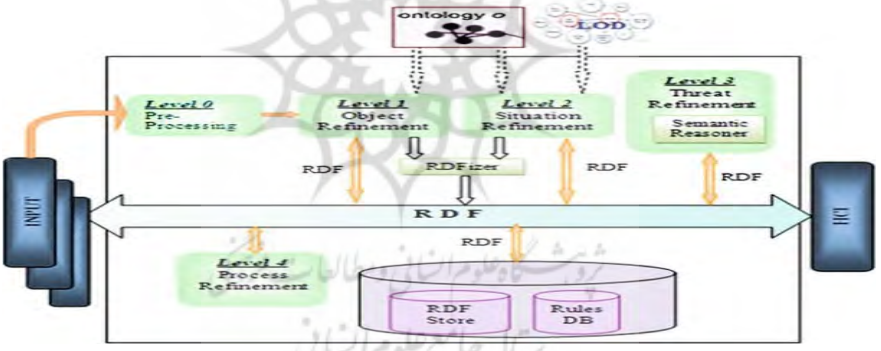
از دیگر معایب این مدل می‌توان به عدم شفافیت، فقدان توجه به عدم قطعیت، راهنمایی محدود در مورد پیاده‌سازی و مقیاس‌پذیری محدود اشاره کرد.

مدل Omnibus: این مدل چندین الگوریتم را جهت بهبود دقت ادغام به‌کار می‌برد. بر خلاف مدل JDL، مدل Omnibus مراحل یک سیستم ادغام اطلاعات را نظیر مدل Boyd و IC در یک دنباله چرخه‌ای سازماندهی می‌کند. در این مدل سعی شده تلفیقی از مزایا و قابلیت‌های کلیدی سایر مدل‌های ترکیب اطلاعات به نحوی مورد استفاده قرار گیرد. این مدل را می‌توان مدل گسترش‌یافته Boyd's دانست؛ زیرا سطوح چهارگانه آن به نحوی بین گام‌های این مدل در محیط مسئله وجود دارد (Raol 2009). این همپوشانی برای این است که بتوان برای رفع خطاها و به اصطلاح، شکاف‌های پردازشی غیرممکن ورود کرده و نتیجه مطلوب‌تر و قابل اطمینان‌تری نسبت به مسئله مورد نظر به‌دست آورد. استفاده از الگوریتم‌های متعدد در مدل Omnibus می‌تواند درک منابع خطاها یا سوگیری‌ها در فرایند ادغام را دشوار کند. این امر شفافیت و تفسیرپذیری مدل را به‌ویژه برای برنامه‌هایی که نیازمند پاسخگویی یا انطباق با مقررات است، محدود می‌کند. از معایب دیگر این مدل می‌توان به افزایش پیچیدگی فرایند ادغام، فقدان توجه به عدم قطعیت، راهنمایی محدود در مورد انتخاب الگوریتم و مقیاس‌پذیری محدود اشاره کرد.

مدل Endsley's Mode of Situational Awareness: مدل Endsley که به‌عنوان مدل ادغام سه سطحی نیز شناخته می‌شود، یک مدل ترکیب داده است که اطلاعات را در سه سطح ادغام می‌کند: سطح شیء، سطح موقعیت و سطح تهدید. ایده این مدل برای رفع خلأ به‌کارگیری فردی خبره در مدل JDL است. مدل دارای دو بخش است: بخش هسته یا آگاهی از موقعیت و بخش بعدی عوامل متعدد مؤثر برای آگاهی از موقعیت. بخش هسته شامل سه سطح درک، فهم و طرح‌ریزی است. بخش دوم عوامل مختلف مؤثر بر آگاهی از موقعیت را بسیار مفصل‌تر با جزئیات توصیف می‌کند. مدل JDL افزون بر همه تلاش‌هایی که در فرایند ادغام اطلاعات بر مبنای شواهد و داده‌های دریافتی انجام می‌دهد، این امکان را به اندازه کافی فراهم نمی‌کند که یک فرد خبره‌ای که قرار است از نتایج تحلیل مدل JDL استفاده کند، بتواند دید کافی از محیط پیرامون مسئله مورد نظر داشته باشد. مدل Endsley's سعی در جبران این نقیصه دارد. توانایی این مدل در ارائه یک دید سلسله‌مراتبی از فرایند ادغام، از مزایای این مدل است. همچنین انعطاف‌پذیری محدود، شفافیت کم، فقدان

توجه به عدم قطعیت، راهنمایی محدود در پیاده‌سازی، و مقیاس‌پذیری محدود نیز از معایب این مدل است.

مدل SemFus (ادغام معنایی داده): این مدل برای رفع مسایل معناشناختی و نحو مدل JDL ارائه شده است. مدل‌های مبتنی بر درک مفاهیم در اصل وظیفه درک سایر مدل‌ها را با استفاده از فناوری‌های معنایی نظیر هستی‌شناسی‌ها دربردارند. در این مدل‌ها از هستی‌شناسی به‌عنوان دانش حوزه زیربنایی جهت ادغام داده‌ها از منابع متعدد استفاده می‌شود. مدل SemFus جهت مدیریت منابع داده گوناگون و پیچیده طراحی شده است و می‌تواند دید قابل تفسیر و شفاف‌تری از داده‌ها ارائه کند. یکی از مشکلات مدل‌های ادغام اطلاعات در سیستم‌های ناهمگن، درک معانی اطلاعاتی در گذر زمان است که سیستم باید قادر به درک آن باشد. این درک بیشتر زمانی که یک تفسیر مشابه از اطلاعات وجود دارد، ملموس‌تر است. یکی دیگر از این کاربردها در حوزه فهم مدل‌های ادغام اطلاعات بوده و در شکل ۵، نشان داده شده است.



شکل ۵. چارچوب مفهومی درک معنایی برای درک مدل‌های ادغام اطلاعات
(Alizadeh Noughabi, Kahani & Behshid 2013)

مطابق شکل ۵، فناوری‌های معنایی نظیر اصطلاحنامه‌ها یا هستی‌شناسی‌ها به یک مدل ادغام اطلاعات نظیر JDL در فرایند مراحل ۱ و ۲ به‌منظور درک مدل و به‌تبع آن درک سیستم افزوده می‌شود (علیزاده و همکاران ۱۳۹۰؛ Alizadeh Noughabi, Kahani & Behshid 2013; Llinas et al. 2004). در حالی که مدل‌های ادغام مبتنی بر اصول معنایی دارای مزایایی هستند، اما محدودیت‌هایی نیز دارند. به‌عنوان مثال، این مدل‌ها ممکن است از منظر فرایندهای محاسباتی پیچیده و مقیاس‌بندی آن‌ها برای کاربردهای ادغام داده‌ها در مقیاس بزرگ دشوار باشند. افزون بر این، طراحی و توسعه

هستی‌شناسی‌ها می‌توانند چالش‌برانگیز باشد و کیفیت هستی‌شناسی تأثیر قابل توجهی بر دقت و قابلیت اطمینان فرایند ترکیب داده‌ها دارد. از معایب این مدل می‌توان به افزایش پیچیدگی محاسباتی زمان استدلال مبتنی بر هستی‌شناسی، تفسیرپذیری و مقیاس‌پذیری محدود، و فقدان توجه به عدم قطعیت اشاره کرد. همچنین در مورد طراحی هستی‌شناسی برای بازنمون دانش حوزه خاص، راهنمایی محدودی وجود دارد و هیچ‌گونه راهکاری در مورد نحوه طراحی یا توسعه هستی‌شناسی ندارد. این امر می‌تواند انتخاب مناسب‌ترین هستی‌شناسی برای یک برنامه یا مجموعه داده معین را دشوار سازد. ایجاد یک هستی‌شناسی که به‌گونه‌ای اصطلاحنامه توسعه‌یافته در ابعاد روابط معنایی است، در صورتی که به‌درستی تنظیم نگردد به‌دلیل انتقال مفاهیم و روابط معنایی گسترده در آن می‌تواند منجر به ایجاد پیچیدگی‌هایی گردد که امکان انحراف اطلاعات و ساختارهای رابطه‌ای در آن را دارد. این امر منجر به ترکیب و ادغام ناصحیح خواهد شد.

جدول ۳، نقاط قوت و ضعف هر یک از مدل‌های ادغام اطلاعات را از منظر کاربست در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای نشان می‌دهد.

جدول ۳. ارزیابی مدل‌های ادغام اطلاعات از منظر سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای

مدل	نقاط قوت	نقاط ضعف
WFFP	حل مشکل عدم توجه به رفتارهای چرخه‌ای و بازبینی فرایندهای پردازشی و تحلیلی در مدل پایه جِی‌دی‌اِل؛ امکان اصلاح فرایند پردازش داده‌های خام	عدم مدیریت منابع داده پویا؛ در نظر نگرفتن عدم قطعیت در منابع داده ورودی؛ عدم مدیریت منابع داده‌های بزرگ و پیچیده مقیاس‌پذیر
TRIP	مدیریت چندین منبع داده؛ امکان پردازش متوالی؛ استفاده از الگوریتم یادگیری ماشین	ضعف در محاسبات فشرده؛ حساسیت به کیفیت داده؛ انعطاف‌پذیری محدود؛ پیچیدگی و فقدان شفافیت؛ نیاز به تخصص دامنه
λ- Jdlc	ارزیابی و شناسایی اشیا و موجودیت‌ها؛ شناسایی وضعیت اشیا و موجودیت‌ها؛ قرارگیری داده‌ها و نتایج پردازشی بر روی یک باس	مشکل در یکپارچه‌سازی؛ کاربرد محدود؛ نیاز به تخصص دامنه؛ مشکلات اعتبارسنجی
IC	حل مشکل عدم توجه به رفتارهای چرخه‌ای؛ امکان بازبینی فرایندهای پردازشی و تحلیلی	کاربرد محدود؛ نیاز به تخصص دامنه؛ مشکلات کیفیت و یکپارچه‌سازی داده؛ عدم قطعیت
DFD	حل مشکل عدم توجه به رفتارهای چرخه‌ای؛ نمایش جریان‌های داده‌ای؛ امکان انجام فرایندهای ادغام پیچیده	راهنمایی محدود در پیاده‌سازی فرایند ادغام؛ تمرکز زیاد بر جنبه‌های فنی ترکیب داده‌ها؛ کاربرد محدود؛ فقدان توجه به عدم قطعیت؛ مقیاس‌پذیری محدود
Boyd's OODA	یکپارچه‌سازی اطلاعات از منابع متعدد؛ وجود ساختار چرخشی؛ امکان بازبینی فرایندهای ادغام	کاربرد محدود؛ فقدان توجه به عدم قطعیت؛ مقیاس‌پذیری محدود

مدل	نقاط قوت	نقاط ضعف
LAAS	تقسیم سیستم‌های بزرگ به ماژول‌ها؛ قابلیت استفاده مجدد؛ قابلیت سازگاری با محیط‌های در حال تغییر؛ استفاده از مدل‌های محلی برای کاهش پیچیدگی محاسباتی	عدم شفافیت؛ فقدان توجه به عدم قطعیت؛ راهنمایی محدود در مورد پیاده‌سازی؛ مقیاس‌پذیری محدود
Omnibus	حل مشکل عدم توجه به رفتارهای چرخه‌ای؛ امکان بازبینی فرایندهای پردازشی و تحلیلی؛ یکپارچه‌سازی اطلاعات از منابع متعدد	عدم شفافیت؛ افزایش پیچیدگی فرایند ادغام؛ فقدان توجه به عدم قطعیت؛ راهنمایی محدود در مورد انتخاب الگوریتم؛ مقیاس‌پذیری محدود
Endsley's	ارائه یک دید سلسله‌مراتبی از فرایند ادغام؛ امکان بازبینی فرایندهای ادغام	انعطاف‌پذیری محدود؛ شفافیت کم؛ فقدان توجه به عدم قطعیت؛ راهنمایی محدود در پیاده‌سازی؛ مقیاس‌پذیری محدود
SemFus	رفع مسائل معناشناختی و نحو؛ استفاده از هستی‌شناسی	فرایندهای محاسباتی پیچیده و مقیاس‌بندی؛ مشکلات طراحی و توسعه هستی‌شناسی‌ها؛ افزایش پیچیدگی محاسباتی؛ مقیاس‌پذیری محدود؛ فقدان توجه به عدم قطعیت؛ راهنمایی محدود

۵. نتیجه‌گیری و بحث

پژوهش حاضر به‌علت تبیین مفهوم و بررسی مدل‌های ادغام داده / اطلاعات و کاربست آن در سیستم‌های مدیریت اطلاعات مبتنی بر روابط از نوع مفهومی است. به گفته «طاهری»، پژوهش‌های مفهومی می‌توانند مبنای بسیاری از ایده‌ها و مباحث باشد و باعث جلوگیری از بسیاری کجروی‌ها و اشتباهات و برداشت‌های نادرست شود، تا جایی که مقالات مفهومی پژوهش‌های یک حوزه را به سمت درست هدایت می‌کند (۱۳۹۹). پژوهش حاضر با رویکرد تحلیلی به بررسی و تبیین مفهوم ادغام داده / اطلاعات و اصطلاحات مرتبط در قالب پژوهشی مفهومی پرداخت. بررسی هزینه و فایده، مهم‌ترین پارامتر در تولید و توسعه سیستم‌ها محسوب می‌گردد. به همین دلیل، شرکت‌ها در حوزه توجه به مدل‌های ادغام اطلاعات و کاربست آن در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای، ورود چشمگیری نکرده و راهکار مشخصی در این‌گونه سیستم‌ها معرفی ننموده‌اند. همچنین پایگاه‌های تجمیعی زیادی در زمینه مدیریت اصطلاحنامه وجود ندارد که افزون بر مدیریت اصطلاحنامه، مدیریت سایر بخش‌ها نظیر نمایه‌سازی و فرهنگنامه را در خود پیاده‌سازی کرده باشد تا نیاز باشد تأثیرات ادغام اطلاعات را در این‌گونه سیستم‌ها مدیریت و کنترل نمایند.

یافته‌ها نشان داد که ادغام، تلفیق و ترکیب اطلاعات در سیستم‌های مدیریت اطلاعات

رابطه‌محور نظیر سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای می‌تواند نتایج و تأثیرات متفاوتی ایجاد کند. اصطلاحات موجود در سامانه‌های مدیریت اصطلاحنامه‌ها در فرایند ترکیب اطلاعات جایگاه خود را حفظ کرده و اصطلاح جدیدی خلق می‌شود؛ در حالی که در فرایند ادغام، اصطلاح جدید در جایگاه یکی از اصطلاح‌های اولیه است. این امر به‌واسطه وجود روابط با سایر اصطلاحات نیازمند کنترل تأثیرات ناشی از فرایند ادغام است. فرایند ادغام می‌تواند توسط سامانه و یا نیروی متخصص به کار رود. از این‌رو، ادغام سیستمیِ اطلاعات ماهیتی دوگانه، شامل هم ترکیب و هم ادغام اطلاعات را دارد. تجمیع، تلفیق و ادغام اطلاعات می‌تواند برتری اطلاعات را به‌دنبال داشته باشد. برتری اطلاعاتی عامل مهمی برای تصمیم‌گیرندگان است و آن‌ها را قادر به انجام قضاوت‌های صحیح در محیط‌های سازماندهی و پردازش اطلاعاتِ در حال تغییر می‌کند. هدف ارائه برتری اطلاعات از طریق هماهنگی اطلاعات جهت رسیدن به برتری تصمیم در کمترین زمان ممکن در حجم انبوهی از اطلاعات است.

وجود اصطلاحات تکراری یا مشابه در سیستم‌های اصطلاحنامه، افزون بر اینکه منجر به سردرگمی و ابهاماتی برای کاربران خبره دست‌اندرکار مدیریت اطلاعات می‌شود، برای کاربران عمومی نیز ناخوشایند است. چشم‌پوشی از رفع تشابه و عدم مدیریت بهینه تأثیرات آن در صورت رفع تشابه در گذر زمان توسط افراد خبره می‌تواند ساختارهای اطلاعاتی را دچار ضعف نماید. بدین منظور، رفع خلأ کنترل رکوردهای مشابه و کنترل تأثیرپذیری آن در سایر ساختارهای اطلاعاتی، نیازمند تمرکز بر روی مدل‌های تجمیع، تلفیق و ادغام اطلاعات نظیر مدل پایه JDL است.

بررسی‌ها نشان داد که مدل‌های ادغام اطلاعات در چهار دسته کلی مدل‌های مبتنی بر گردش اطلاعات، مدل‌های گردش کار و فعالیت، مدل‌های مبتنی بر نقش‌ها و عملکرد موجودیت‌ها، و مدل‌های مبتنی بر درک مفاهیم قرار دارند. مدل‌های ادغام اطلاعات بسته به محیط مسئله، مراحل ثابتی شامل پیش‌پردازش داده‌ها و اطلاعات، پردازش داده‌ها و اطلاعات و تحلیل محتوا، تجمیع و تلفیق اطلاعات و استخراج شباهت‌ها و دانش نهان اطلاعات و رهگیری تأثیرات را در خود دارند. مقایسه این مدل‌ها حاکی از این بود که پیش‌نیازهای مفروض با توجه به محیط مسئله مورد نظر کافی نیست. به همین دلیل با توجه به خلأهای مدل‌های قبلی، مدل‌های ادغام اطلاعات سعی در تکامل مدل متناسب با محیط مسئله مورد نظر خود نمودند. این تکامل منجر گردید تا ویژگی‌های جدیدی در مدل‌های ادغام اطلاعات از ترکیب یا افزودن به ویژگی‌های مدل‌های گذشته به‌وجود آیند. اکثر مدل‌های ادغام داده / اطلاعات مورد بررسی دارای معایبی نظیر کاربرد محدود، انعطاف‌پذیری پایین، تمرکز زیاد بر جنبه‌های فنی، راهنمایی محدود در حوزه پیاده‌سازی فرایند ادغام، مقیاس‌پذیری محدود و فقدان توجه به عدم قطعیت بودند.

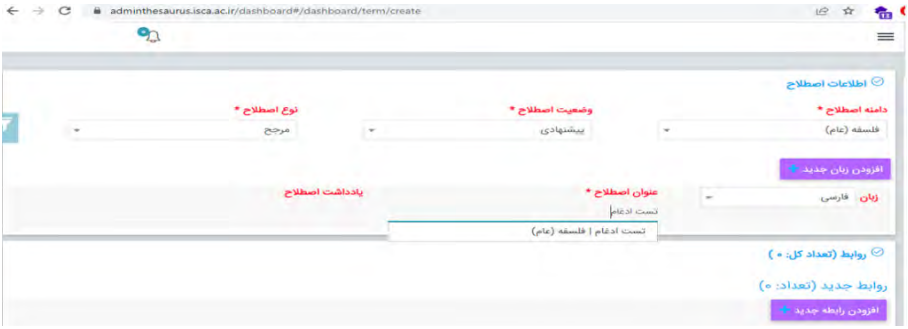
مـدل گـردش کار و فعالیـت Omnibus و مـدل مبتنـی بـر نقش‌هـا و عملکـرد موجودیت‌هـا Endsley's به‌عنـوان مدل‌هـای انتخابـی بـا توجـه بـه مـدل پایـه JDL جهـت به‌کارگیـری در سیسـتم‌های مدیریـت اطلاعـات اصطلاحنامـه‌ای و هستی‌شناسـی‌ها مناسـب اسـت و ایـن قابلیـت را دارنـد کـه بـا برخـی تغییـرات بـه‌کار رونـد. عیـب اصلـی مدل‌هـای ادغـام اطلاعـات بررسی‌شـده، عـدم توجـه بـه ویژگی‌هـای سـاختارهای اطلاعاتـی مشـخص نظیـر اصطلاحنامه‌هاسـت. همچنیـن تمرکـز عمـده ایـن مدل‌هـا بـر اصطلاحـات و واژگان به‌جـای مفاهیـم اسـت. در ایـن مدل‌هـا عـدم توجـه کافـی بـه کاربـران خبـره در سیسـتم‌های اطلاعاتـی مبتنـی بـر تصمیمـات کاربـری کـه وظیفـه اسـتخراج اطلاعـات تخصصـی را دارنـد و نادیده گرفتـن سیسـتم مشـارکت جمعـی بـرای حـل مسـائل پیچیـده از جملـه تشـابه‌یابی و رفـع آن در حجـم انبـوه اطلاعاتـی، عـدم شـرح ریزعملکردهـای مـورد نیـاز بـرای حـل مسـائل، مدیریـت تأثیـرات در صـورت بـروز اشـتباهات سیسـتمی یـا کاربـری و بازگردانـی اطلاعـات تجمیعـی / تلفیقـی و همچنیـن عـدم اجـرای روش‌هـا بـا نمونه‌هـای عملیاتـی در چنیـن سـاختارهایی از جملـه مهم‌تریـن مشـکلات اسـت. ادغـام موفقیت‌آمیـز رکوردهـای اطلاعاتـی اصطلاحنامـه می‌توانـد تغییـر سـاختاری و مفهومـی در اصطلاحنامـه به‌وجـود آورد. از ایـن‌رو، توجـه بـه مبحـث ادغـام اطلاعـات بسـیار حائـز اهمیـت اسـت.

بـا توجـه بـه پیشـنهاد مـدل ادغـام گـردش کار و فعالیـت Omnibus، ایـن مـدل بـا کمـی تغییرات در یـک سیسـتم مدیریـت اطلاعـات اصطلاحنامـه‌ای تحـت وب پیاده‌سـازی گردیـد. در ایـن مثـال، اصطـلاح «تسـت ادغـام» در ابتـدا در پایـگاه جسـت‌وجو می‌شـود. شـکل ۶، نشـان می‌دهـد کـه ایـن اصطـلاح در پایـگاه وجـود دارد.

شکل ٦. جست‌وجوی اصطلاح در سیستم مدیریت اطلاعات اصطلاحنامه‌ای مبتنی بر وب

در صورتی‌کـه اصطـلاح جدیـدی وارد پایـگاه شـود و یـا اصطـلاح دیگـری ویرایـش شـود کـه

مشابه با یکی از اصطلاحات موجود در سیستم باشد، اصطلاح موجود به کاربر نشان داده می‌شود. در شکل ۷، کاربر قصد ایجاد اصطلاح «تست ادغام» جهت افزودن در پایگاه را دارد که با سیستم پیشنهاددهنده پایگاه مواجه می‌شود.



شکل ۷. سیستم تشخیص اصطلاحات موجود در سیستم مدیریت اطلاعات اصطلاحنامه

در صورت درج اصطلاح جدید یا تغییر اصطلاح موجود در پایگاه به اصطلاحی که قبلاً وجود دارد، سامانه اصطلاح تکراری را تشخیص داده و هشدار صادر می‌کند. در صورت تأیید اصطلاح فعلی با اصطلاحی که از قبل وجود داشته، دیگر این اصطلاح به‌عنوان یک اصطلاح جدید در سیستم ذخیره نمی‌شود و فرایند ادغام در اصطلاح موجود صورت می‌گیرد. مدل‌های ادغام مورد بررسی قرار گرفته و در اینجا به‌طور خاص مدل ادغام Omnibus در سیستم مدیریت اطلاعات اصطلاحنامه فعال شده و از ورود اطلاعات تکراری و مشابه جلوگیری کرده و فرایند ادغام را انجام می‌دهد.

شکل ۸. تشخیص تکراری بودن سیستم و هشدار با به‌کارگیری مدل ادغام اطلاعات Omnibus

چالش‌های ادغام اطلاعات را جدا از پیاده‌سازی، محیط مسئله، امکانات نرم‌افزاری و سخت‌افزاری می‌توان به سه دسته پیش از ادغام، حین ادغام و بعد از ادغام دسته‌بندی نمود. مسائل و مشکلات پیش از ادغام در حوزه سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای مشتمل بر شناسایی اصطلاحات (رسم‌الخط، روابط و مفاهیم)، حروف ربط و اضافی در شناسایی تشابهات، شناسایی حروف در زبان عربی یا فارسی به‌دلیل تفاوت زبان (نظیر یای عربی و فارسی که یکی محسوب نمی‌شود)، شناسایی ترجمه اصطلاح در تشخیص تشابه بین دو زبان مختلف اصطلاحات، شناسایی فاصله در بین عناوین اصطلاحات، و عدم در نظر گرفتن سوابق اصطلاحات بایگانی‌شده است. زبان ثبت اصطلاحات، فارسی است در حالی که اصطلاحات زیادی به زبان‌های عربی و انگلیسی نیز وجود دارد که باید در این دامنه ثبت شوند. در حین ادغام نیز چالش‌هایی نظیر نبودِ قوانین مشخص جهت کنترل اصطلاحات، به‌هم‌ریختگی در ساختار اصطلاحنامه با توجه به میزان مشارکت اصطلاح در سایر اصطلاحات، ادغام بیش از دو رکورد اطلاعاتی مشابه، و احتمال خلط میان اصطلاحات متشابه در دو حوزه مختلف نیز وجود دارد که باید برای حل آن‌ها راه‌حل‌های عملیاتی ذکر گردد. بعد از ادغام نیز مسائل و مشکلاتی در حوزه این سیستم‌ها وجود دارد. مدیریت تغییرات در سایر بخش‌های مورد استفاده از اصطلاحنامه، خطا در بازگردانی اطلاعات ادغام و اعمال تغییرات در سایر بخش‌ها، ایجاد شبکه اضافی در ساختار کلی علم، و اختلال در بازیابی اطلاعات از چالش‌های بعد از ادغام است. به‌طور کلی، باید در هر مدل ادغام اطلاعات در سیستم‌های مدیریت اطلاعات اصطلاحنامه‌ای، فرایندهای کلی تشابه‌یابی، بررسی تشابهات، تجمیع پارامترها و ادغام اطلاعات و مدیریت تأثیرپذیری زیرسیستم‌ها مورد توجه قرار گیرند.

چنان‌که بیان شد، تنها مدل ادغام SemFus، مبتنی بر درک مفاهیم است. با توجه به حرکت بسیاری از سیستم‌های اطلاعاتی به‌سوی وب معنایی و هوشمندسازی، لزوم توجه به این نوع مدل‌ها ضروری است. فرایند ادغام سیستمی در سیستم‌های مدیریت اطلاعات رابطه‌مدار با توجه به مفاهیم و درک آنها بسیار حایز اهمیت است. ادغام اطلاعات به‌معنای ادغام دانش و اطلاعات با یکدیگر در مورد موضوعات مختلف است که بر اساس معنا و مفهوم انجام می‌شود. ادغام اطلاعات معنایی به‌عنوان یک فرایند هوشمندانه و مبتنی بر دانش، با ترکیب دانش‌ها، اصطلاحات و مفاهیم مختلف، به ایجاد دانشی جامع در مورد یک موضوع خاص منجر می‌شود. بر خلاف روش‌های سنتی ادغام اطلاعات که بیشتر بر ترکیب داده‌ها تمرکز دارند، ادغام اطلاعات معنایی به‌جای ترکیب داده‌ها و اطلاعات، بر ترکیب

دانش‌هـای مرتبـط و مفاهیـم ذخیره‌شـده در اصطلاحنامه‌هـا تأکیـد دارد. به‌عبـارت دیگـر، در ادغـام اطلاعـات معنایی، هـدف نهایی ترکیـب داده‌هـا نیسـت، بلکـه ترکیـب مفاهیـم و دانش‌هـای مرتبـط بـا یکدیگـر به‌منظـور ایجـاد یـک فهـم جامـع و همگـن از مفهـوم مـورد نظـر اسـت. از ایـن‌رو، توجـه بـه رویکردهـای معنایـی در فراینـد ترکیـب و ادغـام مدل‌هـای جدیـد پیشـنهاد می‌گـردد.

همچنیـن پیشـنهاد می‌شـود کـه از روش‌هـای جدیـدی کـه در مبحـث ادغـام اطلاعـات مطـرح اسـت، در زمینـه توسـعه و ایجـاد سیسـتم‌های مدیریـت اطلاعـات اصطلاحنامـه‌ای اسـتفاده شـود. ایـن روش‌هـا عبارت‌انـد از: Deep Learning, Transfer LearningT Multi-Modal data Fusion, Graph Convolutional Networks, Ensemble Learning و Bayesian Methods. به‌طـور کلـی، روش‌هـای بیـزی، نوعـی اسـتنتاج آمـاری اسـت کـه از قضیـه بیـز بـرای به‌روزرسـانی احتمـالات بـر اسـاس شـواهد جدیـد اسـتفاده می‌کننـد. در زمینـه ادغـام، روش‌هـای بیـزی را می‌تـوان بـرای ترکیـب اطلاعـات از منابـع متعـدد و ایجـاد تخمین‌هـای دقیـق و قابـل اعتمادتـر اسـتفاده کـرد. ادغـام داده‌هـای بیـزی شـامل اسـتفاده از مدل‌هـای احتمالـی بـرای نشـان‌دادن عـدم قطعیـت و وابسـتگی بیـن داده‌هـا از منابـع متعـدد اسـت. ایـن مدل‌هـا را می‌تـوان بـا در دسـترس قـرار گرفتـن داده‌هـای جدیـد به‌روزرسـانی کـرد و بـه سیسـتم ادغـام اجـازه می‌دهـد تـا بـا شـرایط متغیـر سـازگار شـود. یکـی از رویکردهـای رایـج در ترکیـب داده‌هـای بیـزی اسـتفاده از شـبکه‌های بیـزی اسـت کـه از نـوع مدل‌هـای گرافیکـی هسـتند. در ایـن مدل‌هـا روابـط بیـن متغیرهـا بـا اسـتفاده از گره‌هـا و یال‌هـا نشـان داده می‌شـود. هـر گـره نشـانگر یـک متغیـر و لبه‌هـا نشـانگر وابسـتگی بیـن متغیرهاسـت. بـا توجـه بـه شـواهد موجـود اسـتفاده از شـبکه‌های بیـزی بـرای مدل‌سـازی سیسـتم‌های پیچیـده و اسـتنباط توزیـع احتمـال مشـترک متغیرهـا وجـود دارد.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

فهرست منابع

آبادیس. (بی‌تا). ادغام. برگرفته از ادغام/https://abadis.ir/fatofa (دسترسی در ۱۳۹۹/۰۲/۰۴)

پوراسـداللهی‌نژاد، مهسـان. ۱۳۸۸. بررسـی وضعیـت نرم‌افزارهـای مدیریـت و ارائـه اصطلاحنامه‌هـای فارسـی. *تحقیقـات اطلاع‌رسـانی و کتابخانه‌هـای عمومـی* ۱۵ (۳): ۱۰۹-۱۲۹.

حسـن‌زاده، محمـد، آرش محمدخانـی، و آزاد پاک‌نـژاد. ۱۳۹۰. ارزیابـی نرم‌افزارهـای اصطلاحنامـه در ایـران. *نظام‌هـا و خدمـات اطلاعاتـی* ۱ (۱): ۱-۱۱.

رشیدی، علی جبار. ۱۳۹۵. *مفاهیم، نظریه‌ها و کاربردهای ادغام اطلاعات*. تهران: دانشگاه صنعتی مالک اشتر.

طاهـری، سـید مهـدی. ۱۳۹۹. *روش تدویـن مقـالات مفهومـی (Conceptual paper)*. کارگاه ارائه‌شـده در کتابخانـه حضـرت آیـت‌الله العظمـی بروجـردی. قـم.

علیزاده، حوا، محسن کاهانی، بهشید به‌کمال، و علیرضا شکیبا. ۱۳۹۰. چارچوب داده‌آمیزی معنایی مبتنی بر مدل JDL در پنجمین کنفرانس ملی فرماندهی و کنترل ایران. دانشگاه تهران. تهران.

لامعی، ابوالفتح. ۱۳۸۶. تحلیل مفهوم ادغام. *طب و تزکیه* ۳-۴: ۲۲-۳۰.

میرزابیگی، مهدیه، و عباس حری. ۱۳۸۵. بررسی تأثیرپذیری اصطلاحنامه از ورود اصطلاحات جدید با رویکردی بر اصطلاحنامه اریک. *تحقیقات کتابداری و اطلاع‌رسانی دانشگاهی* ۴۰ (۴۵): ۱۲۵-۱۴۶.

ولی‌نژادی، علی، فریدون آزاده، عباس حری، محمدرضا شمس‌اردکانی، و مازیار امیرحسینی. ۱۳۸۷. طرح ادغام سرشاخه خوشه طب سنتی ایران در ساختار ابَراصطلاحنامه «نظام زبان واحد پزشکی (UMLS). *پیاورد سلامت* ۲ (۳): ۶۷-۷۴.

References

Alizadeh Noughabi, H., M. Kahani, & B. Behshid. 2013. SemFus: Semantic fusion framework based on JDL. In Lecture Notes in *Electrical Engineering*, K. Elleithy and T. Sobh (Ed.s), 583-594. New York: Springer; Business Media.

Beddar-Wiesing, S. & M. Bieshaar. 2020. Multi-sensor data and knowledge fusion: A proposal for a terminology definition. https://arxiv.org/pdf/2001.04171.pdf (accessed March 1, 2023)

Bedworth, M., & J. O'Brien. 2000. The Omnibus model: A new model of data fusion? *IEEE Aerospace and Electronic Systems Magazine* 15 (4): 30-36.

Boyd, J. R. 2018. *A discourse on winning and losing*. Edited and compiled by Grant T. Hammond. Alabama: Air University Library.

Castanedo, F. 2013. A review of data fusion techniques. The Scientific World Journal, 2013. https://doi.org/10.1155/2013/704504.

Dasarathy B. 1997. Comparison between C3I and embedded fusion applications and illustrative applications. In *Proceedings of the IEEE . IEEE* 4 (1): 24-28.

Devlin K. 2006. *Situation theory and situation semantics*. In Handbook of the History of Logic, 601-664. Amsterdam: North-Holland.

El Faouzi, N., H. Leung, & A. Kurian. 2011. Data fusion in intelligent transportation systems: Progress and challenges – A survey. *Information Fusion* 12 (1): 4-10.

Elmenreich, W. 2002. *An Introduction to Sensor Fusion*. Austria: Institut fur Technische Informatik.

Endsley, M. 1995. Toward a theory of situation awareness in dynamic systems. *Human Factors* 37 (1): 32-64.

Ganzmarm, J. 1990. Criteria for the evaluation of thesaurus software. *International Cllasification* 17: 148-157.

Grossmann P. 1998. Multisensor data fusion. *The GEC journal of Technology* 15 (1): 27-37.

Habitzel, K. & R. Gregor. 1999. Analysis and General Design of Indexing Systems. https://www.yumpu.com/en/document/view/48641559/analysis-and-general-design-of-indexing-systems-natural- (accessed March 27, 2023)

Hall, D. L. & J. Llinas. 1997. An introduction to multisensor data fusion. In Proceedings of the IEEE 85 (1): 6 - 23. Monterey, USA: IEEE.

Harris, C. J., A. Bailey, & T. J. Dodd. 1998. Multi-sensor data fusion in defence and aerospace. *The Aeronautical Journal* 102 (1015): 229-244.

Jitendra, R. 2009. *Concepts and Theory of Data Fusion. In Multi-sensor data fusion with MATLAB*, 11-61. New York; CRC Press.

Krishnamurthi, R., A. Kumar, D. Gopinathan, A. Nayyar, & B. Qureshi. 2020. An overview of IoT sensor data processing, fusion, and analysis techniques. *Sensors (Basel)* 20 (21): 6076. DOI:10.3390/s20216076.

Llinas, J., C. Bowman, G. Rogova, A. Steinberg, E. Waltz, & W. Frank. 2004. *Revisiting the JDL Data Fusion Model II*. https://apps.dtic.mil/sti/pdfs/ADA525721.pdf (accessed March 14, 2023)

Markin, M., C. Harris, M. Bernhardt, J. Austin, M. Bedworth, P. Greenway, R. Johnston, A. Little, & D. Lowe. 1997. *Technology foresight on data fusion and data processing*. LOndon: The Royal Aeronautical Society.

Microsft. 2023. How to merge Word documents. Retrieved from https://learn.microsoft.com/en-us/office/troubleshoot/word/merge-word-documents (accessed April 05, 2023)

Mokhtari, R. & M. Akhoondzadeh. 2021. Data Fusion and Machine Learning Algorithms for Drought Forecasting Using Satellite Data. *Journal of the Earth and Space Physics* 46 (4): 231-246.

Nakamura, E. F., A. A. F. Loureiro, & A. C. Frery. 2007. Information fusion for wireless sensor networks: Methods, models, and classifications. *ACM Computing Surveys* 39 (3). DOI:10.1145/1267070.1267073.

Pai, F. P., L. J. Yang, & Y. C. Chung. 2017. Multi-layer ontology-based information fusion for situation awareness. *Applied Inteligence* 46: 285-307.

Qi, J., P. Yang, L. Newcombe, X. Peng, Y. Yang, & Zh. Zhao. 2020. Examining data fusion techniques for internet of things enabled physical activity recognition and measure: A systematic survey. *Information Fusion* 55: 269-280.

Raol Jitendra R. 2009. Concepts and Theory of Data Fusion. In Multi-sensor data fusion with matlab, 11-61. New York: CRC Press.

Salemo, J. 2002. Information fusion: A high-level architecture overview. In *Proceedings of the 5th International Conference on Information Fusion*, 680-686. Annapolis, IEEE.

Schuk, T. D., B. Hunter, & D. D. Wilson. 2009. Developing information fusion methods for combat identification. In. *Hnadbook of multisensor Data fusion: Theory and peactice*. Edited by Martin E. Liggins, David L. Hall and James Llinas, 773-812. New York: CRC Press.

Soergel, D. 2001. Data models for an integrated thesaurus database. https://www.dsoergel.com/cv/B54.pdf (accessed March 10, 2023)

Steinberg, A. N., C. Bowman, & F. White. 1999. Revisions to the JDL data fusion. In *Proceedings SPIE 3719, Sensor Fusion: Architectures, Algorithms, and Applications III*. Orlando, United States. DOI:10.1117/12.341367.

Tsanousa, A., E. Bektsis, C. Kyriakopoulos, A. G. González, U. Leturiondo, I. Gialampoukidis, A. Karakostas, S. Vrochidis, & I. Kompatsiaris. 2022. A review of multisensor data fusion solutions in *smart manufacturing: Systems and trends*. Sensors, *22* (5), 1734. DOI:10.3390/s22051734.

Will, L. D. 2013. *Software for building and editing thesauri*. http://www.taxobank.org/content/thesauri-and-vocabulary-control-thesaurus-software (assessed March 24, 2023)

مرتضی محمدی استانی

دارای مدرک دکتری در رشته علم اطلاعات و دانش‌شناسی است. ایشان هم‌اکنون استادیار گروه علم اطلاعات و دانش‌شناسی دانشگاه اصفهان است.

سازماندهی اطلاعات و دانش، هستی‌شناسی، وب معنایی، فرادادە، مدیریت دانش و کتابخانه‌های عمومی از جمله علایق پژوهشی وی است.

محمد کریمی

دارای مدرک تحصیلی کارشناسی ارشد در رشته مدیریت سیستم‌های اطلاعاتی از مؤسسه آموزش عالی غیر انتفاعی تعالی است. ایشان هم‌اکنون کارشناس فنی سامانه‌های پژوهشی اداره فناوری اطلاعات است.

علم داده، تحول سازمانی، هوش تجاری، مدیریت فرایند و بهبود سامانه‌های پژوهشی از جمله علایق پژوهشی وی است.

علی میرعرب

دارای مدرک دکتری مدیریت فناوری اطلاعات است. ایشان هم‌اکنون استادیار پژوهشکده مدیریت اطلاعات و مدارک اسلامی پژوهشگاه علوم و فرهنگ اسلامی است.

تجزیه‌وتحلیل کلان‌داده، یادگیری عمیق و توسعه موتورهای جست‌وجوی معنایی از جمله علایق پژوهشی وی است.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی