

### ۵-۳-۲- وارد شدن داده‌ها از سایر منابع داده

از آنجاییکه هدف پژوهش، ارزیابی کیفیت ذاتی یک مجموعه داده پیش از انتشار بوده است، تمرکز رویکرد پیشنهادی بر ارزیابی یک مجموعه داده بصورت مجزا و بدون وابستگی به سایر منابع داده قرار گرفته است. بنابراین، وارد شدن<sup>۱</sup> داده‌ها از سایر منابع که بر نتیجه ارزیابی مؤثر است، یکی دیگر از موانع کار برای ارزیابی مجموعه‌های مورد آزمایش بوده است.

### ۵-۳-۳- محدود بودن افراد خبره در حوزه داده‌های پیوندی

همان‌طور که در استراتژی ارزیابی اشاره شد، لازمه ارزیابی تجربی مدل پیشنهادی، نظرسنجی از افراد خبره است. با توجه به نوظهور بودن وب معنایی و به‌خصوص داده‌های پیوندی، تعداد افراد خیره در این حوزه بسیار محدود است و این مسأله هم باعث تأخیر در اجرای فرایند ارزیابی شده و هم بر دقت نتایج کار تأثیر گذار بوده است.

### ۵-۴- کارهای آتی

گرچه نتایج ارزیابی رویکرد پیشنهادی رساله رضایت‌بخش و قابل قبول است، ولی برای رفع چالش‌های اشاره شده در بخش قبل، راه‌حلهایی برای توسعه و بهبود رویکرد ارائه شده وجود دارد. در این بخش، مهمترین گام‌هایی که در ادامه مسیر این پژوهش قابل انجام است، در شش بند ارائه خواهد شد.

### ۵-۴-۱- توسعه چارچوب پیاده‌سازی

دو دیدگاه برای توسعه و بهبود چارچوب پیشنهادی وجود دارد. نخست اینکه، چارچوب پیشنهادی به-نحوی توسعه یابد که بصورت آنلاین و با استفاده از Sparql بر روی هر مجموعه داده LOD قابل اعمال

---

<sup>1</sup> Import

باشد. در حال حاضر، چارچوب پیشنهادی، بصورت برون خط اجرا می شود و بر روی مجموعه داده‌هایی قابل اعمال است که امکان دریافت مجموعه داده بصورت کامل وجود داشته باشد. چنانچه این چارچوب به گونه‌ای بهبود یابد که قابلیت استفاده بصورت آنلاین را داشته باشد، می توان کیفیت ذاتی هر مجموعه داده ابر LOD را ارزیابی نمود.

دیدگاه دوم برای بهبود چارچوب پیشنهادی، درخصوص قابلیت اجرا روی سایر قالب‌های داده پیوندی است. در حال حاضر، چارچوب پیشنهادی برای مجموعه داده‌هایی قابل استفاده است که به قالب RDF توصیف شده باشند. یکی از مسیرهای آتی در راستای توسعه چارچوب، آن است که قابل استفاده برای سایر قالب‌های داده‌های وب معنایی باشد تا نیازی به تبدیل داده به قالب RDF نباشد.

همچنین لازم است تا یک نسخه عملیاتی برای منتشرکنندگان مجموعه داده چارچوب تهیه شود تا از طریق یک صفحه کاربری مناسب و با دراختیار داشتن یک دستورالعمل، بتوانند از آن استفاده نمایند. گرچه، در حال حاضر نسخه اولیه چارچوب پیشنهادی پیاده‌سازی شده و به صورت باز برای توسعه دهندگان قابل دسترس است، اما استفاده از این نسخه توسط منتشرکنندگان داده به سهولت امکان پذیر نمی باشد.

#### ۵-۴-۲- بهبود سنج‌های پیشنهادی

از آنجایی که سنج‌های پیشنهادی این رساله هر مجموعه داده را بر مبنای شمای موجود ارزیابی می کنند، این سنج‌ها در شرایطی که فقط داده‌ها بصورت باز در دسترس بوده و شمای داده منتشر نشده باشد، قابل استفاده نیستند. بنابراین، یک گام مهم در راستای بهبود سنج‌ها، اصلاح و طبقه‌بندی آنها به گونه‌ای است که سنج‌های وابسته به شما و مستقل از شما تفکیک شوند. در این صورت، می توان از سنج‌های پیشنهادی برای ارزیابی مجموعه داده‌های فاقد شما نیز استفاده نمود. مسیر دیگر در راستای بهبود

سنجه‌ها، توسعه سنجه‌های ساختاری برای ارزیابی سایر ابعاد کیفی است. به عبارت دیگر، بتوان مجموعه از سنجه‌های ساختاری تعریف نمود که قابلیت ارزیابی ابعاد کیفی دیگر نظیر بهنگامی و اعتماد را داشته باشند.

#### ۵-۴-۳- استخراج هستان‌شناسی از روی مجموعه داده

همان‌گونه در بخش ۳-۵ اشاره شد، یکی از چالش‌های اصلی رساله، پیدا کردن مجموعه داده مناسب است. منظور از مجموعه داده مناسب در این رساله، مجموعه‌ای است که داده به‌همراه شما در آن وجود داشته باشد. به‌منظور تعمیم سنجه‌های پیشنهادی برای انواع منابع موجود در ابر LOD، یک راه‌حل استخراج شمای داده و به عبارت بهتر استخراج هستان‌شناسی از روی داده‌هاست. چنانچه این امر محقق شود، سنجه‌های پیشنهادی برای طیف گسترده‌ای از مجموعه‌های داده ابر LOD قابل استفاده خواهند بود.

#### ۵-۴-۴- ارزیابی پیوندپذیری یک مجموعه داده RDF بر مبنای تئوری گراف

یکی از ابعاد کیفیت ذاتی که در این رساله، مورد ارزیابی قرار گرفته است، پیوندپذیری می‌باشد که تعدادی سنجه ساختاری براساس خصوصیات سه‌گانه‌های RDF برای ارزیابی پیوندپذیری ارائه شد. از آنجاییکه هر مجموعه داده، بصورت گراف RDF نیز قابل نمایش است، می‌توان بر مبنای تئوری گراف و با استفاده از خصوصیات ساختاری هر گراف، قابلیت برقراری پیوند یک گراف RDF را برای پیوستن به ابر LOD و برقراری پیوند با سایر مجموعه‌های داده پیش‌بینی نمود.

#### ۵-۴-۵- ارزیابی کامل بودن مجموعه داده براساس مدل مرجع

در این رساله، کامل بودن یک مجموعه داده بر مبنای میزان انطباق با هستان‌شناسی موجود و بدون توجه به دامنه موضوع و سایر منابع مرتبط مورد ارزیابی قرار گرفته است. یکی از گام‌های آتی، ارزیابی کامل بودن یک مجموعه داده براساس یک مدل مرجع در دامنه مورد مطالعه است. در این صورت کامل بودن یک

مجموعه داده، علاوه بر مقایسه با هستان‌شناسی مورد استفاده، بر مبنای میزان پوشش دانش دامنه نیز سنجیده می‌شود.

#### ۵-۴-۶- ارزیابی سازگاری مجموعه داده با سایر مجموعه‌های داده موجود در ابر LOD

همان‌گونه در تعریف سازگاری در فصل سوم اشاره شد، این بُعد کیفی از دو منظر قابل تعریف است: یکی سازگاری بین داده‌های موجود در یک مجموعه داده و دیگری سازگاری با مفاهیم و موجودیت‌های متناظر که در سایر منابع داده توصیف شده‌اند. از آنجایی که هدف این رساله، ارزیابی داده قبل از انتشار بوده، تمرکز بر رویکرد اول بوده است. بنابراین ارزیابی سازگاری مجموعه داده با سایر منابع منتشر شده می‌تواند زمینه خوبی برای ادامه کار باشد.