[آسان داک](http://www.asandoc.com/) (www.Asandoc.com)

بکارگیری فناوری های داده های مرتبط برای داده های دولت باز یونان

خلاصه :

داده های دولت باز منبع ارزشمندی از اطلاعات است که مورد توجه برخی از دریافت کنندگان مهم قرار گرفته است. با اینحال، این اطلاعات همواره در فرمت خام ارائه نمی شوند، برای مثال پیرو خط مشی های خاص ارائه می شوند و یا همچنان کشف نشده باقی می مانند. هدف از فناوری های داده های مرتبط از سوی دیگر، انتشار داده های وب سایت ها در فرمت های قابل خواندن توسط ماشین هاست (معمولا RDF که از URIS استفاده می کند) تا بتوان آنها را به سایر مجموعه دادههای بیرونی مرتبط ساخت. سایت های دولت یونان توجه کمتری به این جهت گیری داشته است. نمونه موردی اطلاعاتی است که در پرتال های یونان ERMIS و برای ادارات دولتی ارائه شده است که شامل خدماتی است که بر اساس Directive 123/2006/EC ارائه می شوند. در این مقاله ما نمونه موردی را روی کاربرد فناوری های داده های مرتبط در داده های دولت باز یونان بررسی می کنیم که در پرتال دولت یونان ERMIS برای ادارات دولتی قرار دارند. به طور خاص، ما آزمون می کنیم که چگونه این اطلاعات می توانند به داده های مرتبط انتقال داده شوند و به طور مناسبی مرتبط با اطلاعاتی باشند که در سایت های کشورهای اروپی است، ما راهکارهایی نیز ارائه می کنیم.

1. مقدمه

میزان گسترده ای از اطلاعات روی وب سایت ها ، مشکل کسب اطلاعات را زمان بر می کند. اگرچه تکنیک های مختلفی برای استفاده از موتورهای جستجو وجود دارد که استفاده از آنها را کارامد می سازد، بری مثال ارائه صفحات مشابه، باز هم این وظایف مشکلاتی به دنبال دارند. وب معنایی به دنبال حل این مشکل است که اطلاعات مرتبط با وب سایت های مختلف را به یک روش ترکیب می کند. برای رسیدن به این امر به فناوری های ماهواره ای نیاز است. شناخته شده ترین فناوری پاراداریم داده های مرتبط LDP است.

پارادایم داده های مرتبط شامل اقداماتی برای انتشار، به اشتراک گذاری و مرتبط ساختن داده ها روی وب به یکدیگر است و راه جدیدی از یکپارچگی داده ها و عملیاتی کردن آنها را ارائه می دهد. به طور خلاصه، داده های مرتبط به استفاده از وب در لینک های میان داده ها از منابع مختلف می پردازند. محرک اصلی در اینجا ارائه فضاهای داده های مرتبط در فناوری RDF است. اصول اولیه این پارادایم داده های مرتبط عبارتنداز : 1) استفاده از مدل داده های RDF برای انتشار داده های ساختارمند روی وب و 2) استفاده از لینک های RDF برای داده هایی که میان منابع داده ای مختلف ارتباط میانی دارند. هدف از فناوری های داده های مرتبط افزایش استفاده از وب داده هاست که روی جریان های فعلی وب باز بنا شده اند. حرکت های داده های باز مهم است که حرکت های نوظهوری در این صنعت است.

داده های بخش دولتی باز ، داده های باز مرتبط با اطلاعات مفید برای شهروندان و سازمان ها هستند که با بخش دولتی مبادلاتی دارند. به طور مشخص آنها هرچیزی را که به هرکسی ارسال شده است، تحلیل کرده و دوباره از آن استفاده می کنند. با این حال، اگرچه اطلاعات بخش دولتی باز است، سازمان دهی شده و به شیوه آشفته منتشر می شود. این بدان معنی است که همان اطلاعات نه تنها در وب سایتهای مختلف پیدا می شود، بلکه می توان ارتباطی میان وب سایت هایی برقرار کرد که در ظاهر دیده نمی شوند. حتی بدتر ارتباط با اطلاعات مرتبط در سایت هایی که دیگر وجود ندارد. این شرایط موجب می شود تا نیاز باشد که داده های بخش دولتی به داده های باز مرتبط LOD ارتباط داشته باشند تا بتوان حداقل الزامات ارتباط و استفاده مجدد از داده ها را برآورده ساخت.

در این مطالعه، ما از نمونه موردی استفاده می کنیم که LOD را با تمرکز بر داده های پرتال دولت یونان ERMISبرای اداره عمومی ایجاد کرده است و به طور خاص روی خدماتی تمرکز دارد که مبتنی بر Directive 123/2006/EC است. این بخشنامه به همه کشورهای عضو تحمیل شده است که ارائه دهنده اطلاعات مفیدی برای خدمات به دو زبان است ، به زبان اعضای کشورهای و به انگلیسی. با این حال هیچ ارتباطی میان اطلاعات مرتبط در وب سایت ها وجود ندارد، اگر که ارتباط های میان وب سایت های هر کشور عضو را مد نظر قرار ندهیم. این اطلاعات به طور بالقوه به کاربردان نهایی داده می شود تا بتوانند مقایسه ای میان اسناد موجود داشته باشند و کارشان را انجام دهند و یا فعالیت های خدماتی را در سایر کشورها دنبال کنند. به همین شکل، این کاربر می تواند کشوری را انتخاب کند که در آن چنین موضوعی ساده تر بحث می شود.

مقاله ما به شرح زیر سازماندهی می شود: بخش 2 نشان دهنده مطالعات مرتبط با این حوزه است که روی مطالعاتی در وب سایت های بخش دولتی یونان تمرکز دارد. بخش 3 روی مسائل به طور تفصیلی تمرکز می کند، بخش 4 مروری دارد بر گام های ضروری برای انتقال داده ها که در صفحت وب HTML ارائه می شود که مرتبط با LOD است. بخش 5 ارائه دهنده توصیفی از متدولوژی مفروض است و بخش 6 توصیف کننده مطالعات آتی است.

1. مطالعات مرتبط

داده های باز مرتبط به نسبت حوزه مطالعاتی جدیدی است که قابلیت بالایی دارد و بنابراین بسیار مورد توجه قرار گرفته است. بعلاوه وقتی در خصوص داده های باز یونان صحبت می کنیم، مطالعات کمی در ادبیات وجود دارد که مرتبط با انتقال به داده های لینک شده باشد. با دانشی که ما داریم ، سه تلاش انجام گرفته است. اولین مطالعات 3، بیان کننده طرح طبقه بندی برای داده های دولتی باز است و مشوق های OGD را بر اساس آن تحلیل و طبقه بندی می کند. بعلاوه، این امر موجب نمونه موردی از داده های ارتباطی میان سه وب سایت بخش دولتی مرتبط می شود: وب سایت مدرسه ، وب سایت دومین همایش آموزشی بومی آتی و وب سایت وزارت آموزش. با این حال، این مطلعات موردی حوزه محدودی دارد و صرفا برخی از وب سایت ها را در بر می گیرد.

دومین مطالعه که بیشتر به مساله ما نزدیک است در 4 ارائه شده است. در این مطالعه، دسترسی های وب به طور دولتی مد نظر قرار می گیرند که هدف از آن ارتقا شفافیت و بهبود آگاهی شهروندان مرتبط با مخارج دولتی در یونان است که به سادگی از طریق نمودارهای دیداری نشان داده می شود. اطلاعات ارائه شده بر اساس فرایند های معنایی از داده های باز واقعی توسط دولت یونان از طریق برنامه Diavgeia و سیستم اطلاعات مالیاتی یونان انجام شده است. صفحات وب از این سایت ها دانلود شده اند و اعتبار آنها بررسی شده و به طور معنایی این مفاهیم با استفاده از شبکه های ساختارمند بررسی شده است. هدف نهایی این مطالعه ارتباط درونی داده ها به طرح های داده ای موجود است که مشابه با آن چیزی است که در سراسر دنیا با نام های دیگر در حال پیگیری است.

سومین تلاش با هدف ارتباط میان نهادهای یونان و LOD انجام شده است که در مطالعه 6 بدان اشاره شده است. این مطالعه گام اولیه ای برای رسیدن به ابر LOD است که به طور اولیه ترکیبی از مچموعه داده های بهم مرتبط و قطب های اصلی Greek DBpedia است. در خصوص این پروژه در عین حال که این ابر شکل می گیرد و تقویت می شود، محققان به مسائل گسترده تری از ویژگی های زبان غیر لاتین اشاره می کنند که در منابع و سری های SPARQL بیان شده اند. آنها روشی برای حل این مساله دارند که قابل کاربرد در همه زبان هایی است که الفبای غیر لاتین دارند.

1. مساله

مساله ای که دراینجا مطرح می شود در بردارنده داده های دولتی است که در پرتال یونان ERMIS وجود دارد و برای اداره های دولتی استفاده می شود و بسیار مرتبط با بخش نامه 123/2006/EC [7 است. این بخشنامه بر روی ساده سازی رویه های حرفه ای برای شهروندان اروپایی در کشور های عضو تمرکز دارد. اجرای این بخش نامه هر کشور عضو را وادار می سازد که پرتال های مشخصی با PSC ارائه دهند. این پرتال باید شامل اطلاعاتی باشد که همه اسناد حمایتی مورد نیاز برای فعالیت های خدماتی را در بر می گیرد، و نیز در بردارنده اطلاعات مرتبط در دو زبان، زبان کشور عضو و انگلیسی. پرتال مرکزی ارائه دهنده ویژگی های خاصی است که توسط هر کشور عضو پشتیبانی می شود.

هر فعالیت خدماتی به یک یا چند طبقه تقسیم می شود که به دنبال NACE است. طبقه بندی اروپایی از فعالیتهای اقتصادی به دنبال یک چهارچوب مرجع اروپایی برای تولید و انتشار آمار است که مرتبط با فعالیت های اقتصادی است. NACE یک ابزار مهم برای مقایسه داده های آماری مرتبط با فعالیتهای اقتصادی در سطح دنیاست. جدول 1 ارائه دهنده لیستی از فعالیتای خدماتی است که در PSC یونان بیان شده است و براساس طبقه بندی NACE است:

هر فعالیت خدماتی شامل اطلاعات زیر است:

* نام فعالیت خدماتی
* توصیف فعالیت خدماتی
* مسوولیت بخش دولتی برای چهارچوب قانونی فعالیت خدماتی
* مسوولیت بخش دولتی برای دریافت و پرداخت حقوق فعالیت های حدماتی
* نوع فعالیت خدمای از جمله کسب اجازه، مجوز ، گواهینامه و غیره
* رویداد زندگی، همچون شروع یک کسب وکار، بیمه شدن ، درس خواندن و غیره
* قواعد قانونی همچون همه قانون های مرتبط
* همه پیش فرض هایی که ممکن است برخی از گروههای متقاضی را محدود سازد
* هزیه فعالیت های خدماتی
* زمان تحویل برای مثال زمان مورد نیاز برای ارائه تصمیم اختیاری
* نظرات که ممکن است شامل هر گونه اطلاعاتی باشد که می تواند برای تایید متقاضی مفید باشد همچون تلفن و یا ایمیل برای کمک و یا شکایت
* لیستی از کلید واژه های مرتبط
* محیط سرمایه گذاری برای مثال، اینکه این فعالیت خدماتی توسط برنامه های سرمایه گذاری حمایت می شود یا خیر
* اینکه حضور فیزیکی برای ثبت تقاضا مورد نیاز است یا خیر
* اینکه حضور فیزیکی برای دریافت نتایج تقاضا مورد نیاز است یا خیر
* توصیف از چگونگی کاربردی که شامل بخش های دولتی میشود و توجه خاصی را به عواملی ایجاد می کند که منجر به رد شدن تقاضاها میشود
* همه اسناد حمایتی موردنیاز
* حساب های پشتیبان که در جایی است که هزینه فعالیت های خدماتی مورد نیاز در بانگ پرداخت میشود.

قوانین NACE : برای مثال طبقه شماتیکی که در آن فعالیت خدمات می تواند طبقه بندی شود

محتوایی که در بالا بدان اشاره شد برای هر فعالیت خدماتی توصیف می شود که بر اساس آنتولوژی e-GIF خاصی است این آنتولوژی بخشی از چهارچوب عملیات میانی در دولت الکترونیک یونان است، پیمایشی که مرتبط با لیستی از قواعد برای ارائه به خدمات دولت الکترونیک به نهادهای دولتی، کسب وکار ها شهروندان به شیوه ای متحد است که همه اداره های دولی یونان را در بر می گیرد. هدف از چهارچوب دولت الکترونیک پشتیبانی از دولت الکترونیک کارامد در سطح مرکزی، منطقه ای و بومی است که سهمی در کسب عملیات های درونی سیستم ها ، رویه ها و داده های اطلاعاتی دارد.

با این حال، راهی که اطلاعات فعالیت خدماتی در Greek PSC سازماندهی می شود، هیچ ارتباط درونی میان فعالیتهای خدماتی ایجاد نمی کند. برای مثال ، هیچ ارتباطی میان دو فعالیت خدماتی ایجاد نمی شود وقتی که هر دو یک پیش نیاز دارند. بعلاوه ارتباطی میان فعالیت خدماتی و تبدیل آن به انگلیسی وجود ندارد. این بدلیل این حقیقت است که هیچ ارتباط معنایی میان فعالیتهای خدماتی مرتبط وجود ندارد. بنابراین نیاز است تا OGD به LOD مرتبط شود.

* 1. داده های دولت باز

مشوق های OGD فقط اخیرا مورد بحث قرا گرفته اند که فاقد طرح های طبقه بندی برای تحلیل هستند. با اینحال، تعداد زیادی از خط مشی های عملی توسط ذینفعان مختلفی بیان شده اند. گروه دولت الکترونیکی کنسرسیوم وب گسترده جهانی سه مرحله را برای اداره های دولتی بیان می کند که بتوانند داده هایشان را به صورت باز ارائه کرده و به اشتراک گذارند:

* در ابتدا انتشار داده های خام در شکل هایی از فایل های شناخته شده و یا فرمت های غیر اختصاصی همچون CSV و XML
* دوم، ایجاد کاتالوگ های آنلاین از داده های خام
* و در آخر ایجاد دادهها در فرمت های قابل خواندن در ماشین
  1. LOD داده های باز مرتبط

داده های مرتبط به نظر می رسند که نقش مهمی در آیده مشوق های OGD داشته باشند. واژه داده های مرتبط اشاره دارد به داده هایی که در وب به گونه ای منتشر می شوند که قابل خواندن در ماشین هستند. این بدان معنی است که میتواند به طور صریح تعریف شود و مرتط با سایر مجموعه داده های بیرونی باشد و می تواند با مجموعه داده های بیرونی ارتباط برقرار کند. مشوق های داده های مرتبط بر اساس این فلسفه و یا استفاده از فناوری های وب معنایی است. با اینحال، در مقایسه با دیدگاه وب معنایی کامل، این امر به معنی انتشار داده های ساختارمند در RDF با ساتفاده از URI است و نه تمرکز بر سطح واسط و یا ماشینی. این امر موجب ایجاد شبکه ای از داده ها می شود، همچون داده هایی که منابع غیر متمرکز و یا چندگانه دارند که می توانند از طریق لینک های مختلفی بهم متصل شوند. شبکه داده ها به دنبال جایگزینی داده های بسته با مجموعه داده هایی است که در بستر وب می توانند به طور گسترده ای توزیع شوند.

رویکرد داده های مرتبط نیاز به شناسایی منابع با رفرنس های URI دارد که می تواند روی پروتکل های HTTP و یا در دادههای RDF تعریف شود که توصیف کننده منابع شناسایی شده است. بعلاوه، داده های مرتبط شامل ایجاد لینک هایی میان مراجع URI هستند ، بنابراین می توانند داده های بیشتری را کشف کنند. به طور خاص تر، چهار اصل داده های مرتبط به شرح زیر می باشند:

* همه آیتم ها باید با استفاده از URI ها شناسایی شوند، به جای رشته های خالی که نمی توانند لینک های RDF را ایجاد کنند چون محدود به مستنداتی نیستند که بتوان آنها را ارائه داد
* همه URI ها باید دوباره ارجاع دهی شوند با استفاده از HTTP URI ها که این امکان را می دهد که آیتم ها را از طریق URI را شناسایی کرد.
* وقتی به دنبال یک URI هستیم این امر به اصل راه را بگیر و راست برو اشاره دارد.
* ارتباط با سایر URI ها به منظور آن است که بتوان داده های بیشتری راکشف کرد

داده های مرتبط میان منابع اطلاعاتی و غیر اطلاعاتی تمایزی قائل می شوند. منابع اطلاعاتی اشاره به همه منابعی دارد که در اسناد سنتی همچون اسناد، تصاویر و غیره آنها را پیدا می کنیم و اسناد غیر اطلاعاتی اشاره به چیز هایی از دنیای واقعی دارد همچون افراد، مدارس، قوانین ، نهادهای دولتی و غیره. پذیرش شناسایی کنندگان امکان درک منحصر به فرد از منابع اطلاعاتی را در دنیای وب می دهد، ولی نه آنچیزی که منابع اطلاعاتی به دنیای واقعی ارجاع می دهند. بعلاوه، مسائل اصلی در این شبکه از داده ها به دنبال یافتن شناسایی کننده هایی است که به دنیای واقعی اشاره دارند. این شناسایی کننده ها با نام URI شناخته می شوند. روش مفیدی برای هدایت انتشار داده های مرتبط روی وب .

پس از شناسایی منابع اصلی در منابع داده ای و برای ارتباط داده ها با هم، از برنامه هایی استفاده باید کرد که متکی بر استفاه مجدد از URL مشترک باشند. با اینحال استدلال مورد نیاز برای رسیدن به ویزگی های عملیاتی معکوس در وب سایت ها می تواند هزینه بر باشد و از هرگونه شاخص بندی کارامدی جلوگیری کند. برای بیان این چالش ها در کشف داده های وب معنایی غیر متمرکز ، نیاز به عناصر معماری است که نیازمند ایجاد لینک های ازدست رفته است که این امکان را به کلاینت های وب معنایی می دهد که بتوانند گراف های غیر مستقل از هر ارتباط را به وب سایتی به گراف های بزرگتر مجازی ارتباط دهند

1. متدولوژی پیشنهادی

با در نظر گرفتن تحلیل های این مساله که در بخش 2 بیان شد، می توانیم عناصر زیر را ارائه دهیم که هر کدام درزیر بخش های جداگانه ای توضیح داده می شوند؟

* 1. بازیابی و ذخیره صفحات وب

همانطور که در بخش 2 آمده است، صفحات وبی که در سایت پرتال دولت یونان ERMIS برای ادارات دولتی آمده اند دراینجا مد نظر قرار می گیرد. اگرچه اطلاعات موجود دراین پرتال در دسترس دولتی هاست، هیچ نوع مداخله ای صورت نمیگیرد، برای مثال افزودن تگ های RDF امکان پذیر است. این امر منجر به این ضرورت می شود که بتوان کپی از هر کدام از این صفحات دانلود کرد و نیز مکانیزمی داشت که به طور دوره ای بتوان تغییرات روی محتوای هر صفحه را بررسی کرد. مجموعه داده ها همچنین می تواند میزبان همه اطلاعات مرتبط با هر فعالیت خدماتی باشد. این اطلاعات در سایت های مشخصی قرار دارد که توسط موسسه آموزشی فناوری آتن ارائه می شود

چون هدف ما بهبود اطلاعاتی است که توصیف کننده هر فعالیت خدماتی در اطلاعات معنایی در شکل های ارتباطات RDF است، یک آنتولوژی نیاز است تا بتوان مفاهیم اصلی را با شاخص ها نشان داد و نیز روابط میان آنها را به تصویر کشید. در بخش 2، ما بیان می کنیم که آنتولوژی e-GIF که توسط ERMIS استفاده شده است پایه و اساسی برای سازماندهی اطلاعات است. با این حال، آزمون تفصیلی از این آنتولوژی ما را به این نتیجه می رساند که باید برخی از اطلاعات حذف شوند. در میان آنها روابط نیز وجود دارند که هیچ ارجاعی بدانها داده نشده است و فاقد هر گونه ویژگی های مفهومی هستند. این موارد موجب می شود تا همه مفاهیم ضروری ، مالکیت ها و روابط در کنار هم قرار گیرد تا بتوان تصویری ازوضعیت فعلی روابط میانی معنایی بیان کرد. آنتولوژی به روز شده به ما کمک می کند تا بتوان لینکی میان همه RDF ها برقرا کرد. مثال هایی از چنین لینک هایی به شرح زیر است: لینک های میان فعالیت خدماتی در یونان و تبدیل آنها به انگلیسی، لینک های مرتبط با مسوولیت بخش دولتی برای چهارچوب قانونی فعالیت خدمات، مسوولیت بخش دولتی برای دریافت و تسریع اختیار فعالیت خدمات ، نوع فعالیت خدماتی ، هر نوعی از پشتیبانی از اسناد و نیز قواعد NACE. برای مثال، طبقه شماتیکی که در آن فعالیت خدماتی می تواند در آن طبقه بندی شود. وقتی از RDF در هر صفحه وب استفاده می شود، ارتباطاتی برقرار می شود، برای مثال کاربر قادر خواهد بود تا اطلاعات مرتبط را هدایت کند.

گام بعدی شامل بهبود معنایی عناصر بالای صفحات وب است که همراه با مفهومی از e-GIF به روز شده است. این امر می تواند از ابزارهایی همچون چهارچوب زنا استفاده کند. فاز بعد، شامل بیان شواهد از سه گانه های مرتبط با RDF باشد. (برای مثال داده های واقعی) که توسط اجرای ابزارهای مناسب بدست می آید و یا با استفاده از زبان مسیر دهی D2RQ نشان داده می شود که بین طرح های پایگاه داده و RDFS/OWL ارتبطی برقرار می کند. سه گانه های RDF در عنصر های سرور LOD ذخیره می شود که می توانند نوع خاصی از پایگه داده را شکل دهند و امکان ذخیره سازی، ارجاع دهی و برقراری ارتبط میان داده های RDF و OWL را می دهند. به منظور انتشار محتوای پایگاه داده ما روی وب معنایی می توانیم از ابزارهای شناخته شده ای همچون Open Link Virtuoso [17], or D2R [18]. The D2R SPARQL استافده کنیم و یا متناظر با آن از Open Link Virtuoso Environment که کاربران را قادر می سازد تا کوئری های SPARQL را برای بازیابی داده ها استفاده کنند. توضیحات گرافیکی از متدولوژی ما دز شکل 1 آمده است

43. آزمون ارتباطات میانی بالقوه با سایر مجموعه های داده ای

وقتی اطمینان حاصل کنیم که ارتباطات درونی برقرار است می توان به همه صفحات وب از طریق پرتال های دولت یوتان برای ادارات دولتی دست پیدا کرد که تحت بخش نامه 123/2006/EC [7]‌ف ما اینگونه برنامه ریزی می کنیم که بتوانیم ارتباطات درونی آنها را با لطاعات بدست آمده از سایت های مشخص از کشورهای دیگر ارتباط دهیم . به یاد داشته باشید که این امر به سادگی بدست می آید چرا که هر کشوری باید اطلاعات فعالیت های خدماتی اش را به انگلیسی نیز ارائه دهد؛ بنابراین ارتباط میان سایت های مختلف از طریق صفحات وب امکان پذیر است. وقتی این امر رخ می دهد، شهروند اروپایی می تواند به مجموعه فعالیتهای خدماتی از کشور انتخابی دسترسی داشته باشد که نیازمند حداقل پوشش اسنادی است که موجب هدایت ساده وب سایت های یک کشور به دیگری می شود تا بتوان کشور هدف را به درستی درک کرد

وظیفه بعدی نیازمند آنتولوژی است که به مثابه یک پل عمل می کند. آنتولوژی مورد نظر DBpedia است. بر اساس این توصیف DBpedia یک اقدام جامعه برای استخراج اطلاعات ساختارمند از ویکیپدیا است تا بتواند اطلاعات را در دسترس وب قرار دهد. DBpedia به شما این امکان را می دهد تا در ویکی پدیا به دنبال کوئری های پیچیده بود و ارتباطی میان مجموعه داده های دیگر روی وب با داده های ویکیپدیا برقرار کند. ارائه گرافیکی از آنتولوژی DBpedia را می توانید در 20 دنبال کنید. همانطور که دیده می شود، آنتولوژی DBpedia نمی تواند شامل مفاهیم و روابطی باید که بخش های دولتی، فعالیتهای خدماتی و یا اسناد پشتیبانی را ارائه می دهد. با این حال مزیت های DBpedia از این حقیقت به کاربر این امکان را می دهد تا مفاهیم و طبقه ها را بهبود داد. بنابراین ، در آزمون عمقی از همه وب سایت های نمونه از همه کشورهایی که انتخاب شده اند، در ابتدا باید تحلیلی روی محتوای آنان داشت. طبقه بندی از مفاهیم و روابط باید ایجاد شود تا بتوان ترکیبی از همه اطلاعات معنایی از وب سایت ها ارائه کرد. این طبقه بندی می تواند به DBpedia ارائه شد که عنصری است که نه تنها اهداف را مد نظر قرار می دهد بلکه از نقاط شروع نیز برای سایر اهداف استفاده می کند. ارتباط بالقوه دیگر نیز روی DBpedia یونان با اجرای مستمر این موضوع دنبال می شود.

زبان D2RQ و یا سایر ابزار

مثال های سه گانه RDF

منابع و ویژگی های URI

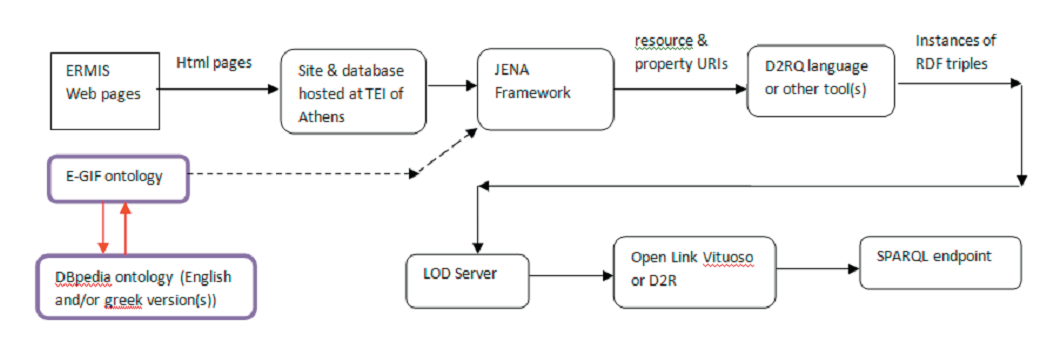
چهارچوب

سایت ها و پایگاه های داده تحت پوشش TEI آتن

آنتولوژی

صفحات

صفحات وب



نقطه پایانی

ارتباطات باز

سرور LOD

آنتولوژی(ورزن های انگلیسی و یونانی)

شکل 1. طرح متدولوژی مفروض

همه مطالعاتی که بیان شدند در بردارنده شناسایی لینک های میانی بین مجموعه داده هاست که از طریق جستجوهای دستی بدست می آید. چهارچوب سیلک 22 میتواند به طور خودکار شناسایی شده و چنین لینک هایی را ارائه و مورد آزمون قرار دهد.

4.4. نتایج آماری

RDF ناشی از اقداماتی است که در بخش 4.2 بیان شده است و همچنین برای سایر اهداف استفاده می شود. در میان آنها، می توان نتایج آماری دیگری را نیز بدست آورد: 1) چه تعداد فعالیتهای خدماتی از طریق بخش دولتی انجام می شوند 2) در کدام فعالیتهای خدماتی سند پشتیبانی خاص 1 مورد نیاز است 3) در چه تعداد فعالیت خدماتی به هر دو سند پشتیبانی نیاز است. 4) کدام فعالیت خدماتی نیازمند حداقل/حداکثر اسناد پشتیبانی است. 5) برای چه تعدادی از فعالیتهای خدماتی، فعالیت خدماتی 1 پیش فرض است. برای آن دسته از نتایج آماری که ورودی به مراجع دولتی است باید مروری بر دشواری فعالیتهای خدمات خاص اجرایی داشت بنابراین این امر راهی برای برآوردن نیاز انجام اقدامات ساده سازی اداری است. مراجع دولتی همچنین ممکن است از اطلاعات بدست آمده از وب سایت ها و از سایر کشورها نیز بهره مند شوند که مطلوب اجرای اقداماتی است که در بخش 4.3 آمده است. بدون شک، اقدامات ساده سازی اداری مزیتی برای کشورماست که در امور کسب وکار تقویت شوند.

1. نتیجه گیری- مطالعات آتی

در این مقاله ما نمونه موردی از تبادل اطلاعات در دسترس دولتی بری فعالیت خدماتی بر اساس بخش نامه 123/2006/EC [7 بیان کردیم که بیانگر داده های باز مرتبط است. مساله مورد بحث منجر به پیش بینی برخی از مسائل و نیازها می شود1) دانلود صفحات وب در یک نسخه بومی و 2) تقویت آنتولوژی e-GIF به منظور پوشش دادن همه انواع مفاهیم و روابطی که به دقت صفحات وب را توصیف می کنند 3) تبدیل صفحات وب به RDF با استفاده از آنتولوژی e-GIF . 4) انتخاب ابزارهای مناسب برای اجرای همه اقدامات کارامد برای رسیدن به داده های باز مرتبط

دیدگاه ما شامل دو مرحله می شود. مرحله اول موجب ایجاد لینک های مطلوب می شود، برای مثال LOD میان صفحات که متعلق به مجموعه داده های اصلی است که باید برخی از صفحات را در ورژن ترجمه شده به انگلیسی در صفحات وب قرار داد. مرحله دوم شامل ارتباط میانی مجموعه داده های با با معادل های آنان است که سایت های کشورهای اروپایی است، برای مثال ایجاد ارتباط میانی بین سایت هایی که بر محتوای انگلیسی هر سایت تمرکز دارند. به منظور انجام این امر، آنتولوژی DBpedia به عنوان آنتولوژی مورد استفاده در نظر گرفته شده است. برای این امر ، بهبود پتانسیل آنتولوژی DBpedia موجب طبقه بندی ها و روابط مناسبی می شود که باید مورد بررسی قرار گیرند.

جدول 1. لیستی از فعالیتهای خدماتی در PSC یونان

|  |  |  |
| --- | --- | --- |
| طبقه بندی | فعالیتهای خدماتی به انگلیسی | فعالیتهای خدماتی در یونان |
| کشاورزی، جنگل و ماهیگیری  معدن و حفاری  تولید  تامین آب، مدیریت ضایعات و فعالیتهای تجدیدی  ساخت و عمران  تجارت عمده فروش و خرده فروش؛ تعمیر وسایل موتوری و موتورسیکلت  حمل و نقل و ذخیره سازی  اسکان و فعالیتهای خدمات غذایی  اطلاعات و ارتباطات  فعالیتهای مالی و بیمه  فعالیتهای دارایی واقعی  حرفه ای، علمی و فعالیتهای فن ی  فعالیتهای خدمات اداری و پشتیبانی  آموزش  هنرها، سرگرمی و تفریح  سایر فعالیتهای خدماتی |  |  |