Archive of SID

Ontirandoc: Integrated Collaborative System for Developing Persian Ontology

Omid Milanifard
PhD Candidate in Information Technology Management; Iranian Research Institute for Information Science and Technology (IranDoc);
Corresponding Author milanifard@students.irandoc.ac.ir

Mohsen Kahani
Professor of Computer Engineering; Ferdowsi University of Mashhad; Visiting Professor; Iranian Research Institute for Information Science and Technology (IranDoc) kahani@um.ac.ir

Iranian Journal of Information Processing and Management

Received: 11. Dec. 2016 **Accepted: 23, Feb. 2017**

Iranian Research Institute for Science and Technology
ISSN 2251-8223
eISSN 2251-8231
Indexed by SCOPUS, ISC, & LISTA
Vol. 33 | No. 3 | pp. 453-486
Spring 2018

Abstract: While ontology development is beneficial, it is very costly and time consuming. In order to reduce this cost as well as to increase the accuracy and quality of ontology development, researchers have proposed different methodologies. The goal of these methodologies is to present a systematic manual or semi-automated development of ontologies, while each differs and has its strengths and weaknesses. In this paper, after reviewing current methodologies, we present a new integrated collaborative methodology for ontology development, and compare it with the existing ones. This new system, called Ontirandoc, has been used in two ontology development projects and its accuracy has been evaluated.

Keywords: Ontology, Ontology Development Methodology, Ontology Development System

www.SID.ir

Archive of SID

انتیرانداک: سامانهٔ یکپارچهٔ توسعهٔ مشارکتی هستان‌نگار فارسی

امید میلانی‌فرد
دانشجوی دکتری مدیریت فناوری اطلاعات؛
پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)؛
پدیدآور رابط milanifard@students.irandoc.ir

محسن کاهانی
دکتری مهندسی کامپیوتر؛ استاد؛
گروه کامپیوتر دانشگاه فردوسی مشهد؛
استاد مدعو پژوهشگاه علوم و فناوری اطلاعات ایران؛
kahani@irandoc.ac.ir

پژوهش‌نامهٔ پردازش و مدیریت اطلاعات

دریافت: ۱۳۹۵/۰۹/۲۱ | پذیرش: ۱۳۹۵/۱۱/۲۳ | مقاله برای اصلاح به مدت ۲ روز نزد پدیدآوران بوده است.

چکیده: ساخت هستان‌نگار با وجود داشتن مزایای بسیار، مستلزم صرف زمان و هزینهٔ زیادی است. محققان در راستای کاهش این هزینه‌ها و همچنین افزایش دقت و کیفیت در فرایند ساخت هستان‌نگار، روش‌های گوناگونی معرفی کرده‌اند. هدف اصلی این روش‌ها، ارائهٔ یک الگوی نظام‌مند دستی یا نیمه‌خودکار برای ساخت هستان‌نگار است. روش‌های موجود با وجود داشتن هدف مشترک، تفاوت‌های زیادی دارند و در هر یک از آن‌ها نقاط قوت و ضعفی وجود دارد. در این مقاله، پس از مروری بر هستان‌نگار و روش‌های ساخت آن، یک سامانهٔ یکپارچهٔ توسعهٔ مشارکتی هستان‌نگار معرفی و ابعاد مختلف آن با سایر روش‌ها مقایسه خواهد شد. از این سامانهٔ جدید با عنوان «انتیرانداک» در دو پروژهٔ جداگانهٔ توسعهٔ هستان‌نگار استفاده و صحت عملکرد آن آزمون شده است.

کلیدواژه‌ها: هستان‌نگار، روش توسعهٔ هستان‌نگار، سامانهٔ توسعهٔ هستان‌نگار

فصلنامه | علمی پژوهشی
پژوهشگاه علوم و فناوری اطلاعات ایران
شاپا (چاپی) ۸۲۲۳-۲۲۵۱
شاپا (الکترونیکی) ۸۲۳۱-۲۲۵۱
نمایه در LISTA، ISC، و SCOPUS
jipm.irandoc.ac.ir
دورهٔ ۳۳ | شمارهٔ ۳ | ص‌ص ۱۱۲۳-۱۱۵۲
بهار ۱۳۹۷

www.SID.ir

Archive of SID

۱. مقدمه

انسان از دیرباز به‌دنبال شناخت ماهیت و ساختار جهان پیرامون خود بوده است. واژهٔ «آنتولوژی»[1] را نخستین بار «ارسطو» برای تعریف علم «بودن چون بودن»[2] در متافیزیک خود به‌کار برده است. علم آنتولوژی یا هستان‌شناسی زیرشاخه‌ای از فلسفه به حساب می‌آید. با وجود آن‌که هستان‌شناسی یک علم کاملاً نظری است، کاربردهایی از آن وارد علم اطلاعات شده است. مقصود از آنتولوژی در علم اطلاعات، یک قلم اطلاعاتی است که دانش یک حوزهٔ خاص را مدل‌سازی می‌کند (Guarino, Oberle and Staab 2009). در این تعریف از آنتولوژی، وجود[3] چیزها و ارتباطات آن‌ها مورد مطالعه قرار می‌گیرد. اما در اینجا «وجود»، یک معنای عملی (و نه فلسفی) دارد: چیزی «وجود» دارد که بتوان آن را بازنمایی کرد (Gruber 1995). از آنجا که در عمل نمی‌توان یک حوزه را به‌صورت کامل شناخت و مدل اطلاعاتی تنها می‌تواند نگاشتی از برخی مفاهیم مهم یک حوزه و روابط بین آن‌ها را بازنمایی کند، استفاده از «نگاشتن» به جای «شناختن» مناسب‌تر به نظر می‌رسد. به همین دلیل، در این نوشتار از اصطلاح «هستان‌نگار» به‌عنوان معادل فارسی «آنتولوژی» در حوزهٔ علم اطلاعات استفاده شده است.

مشهورترین و پُرارجاع‌ترین تعریف از هستان‌نگار را «گروبر» توصیفی صریح از یک مفهوم‌سازی[4] معرفی می‌کند (Gruber 1993). «برُست» بخش‌های جدیدی به این تعریف اضافه کرد و هستان‌نگار را «توصیفی رسمی از یک مفهوم مشترک» خواند (Brost 1997). اشتراکی‌بودن مفهوم به این نکته اشاره دارد که در ساخت مفهوم باید اجماع نظر خبرگان حوزه وجود داشته باشد. همچنین، مفهوم باید به‌صورت *رسمی* بازنمایی شود؛ یعنی توسط ماشین قابل خواندن باشد. «ستودر، بنجامینز و فنسل» با ادغام دو تعریف فوق، هستان‌نگار را «توصیفی صریح و رسمی از یک مفهوم‌سازی مشترک»[5] تعریف کردند (Studer, Benjamins and Fensel 1998).

هدف اصلی ساخت هستان‌نگار، نمایش مدلی انتزاعی از واقعیت است که مورد تأیید

1. ontology
2. being qua being
3. existence
4. explicit specification of a conceptualization
5. formal, explicit specification of a shared conceptualization

www.SID.ir

Archive of SID

افراد ذی‌نفع باشد. واقعیت مدل‌سازی‌شده می‌تواند تنوع زیادی داشته باشد. از حوزه‌های دانشی بشر گرفته تا سیستم‌های نرم‌افزاری و مخازن اطلاعاتی می‌توانند به‌عنوان دامنه یا به‌عبارت دیگر، فضای گفتمان یک هستان‌نگار باشند. این موضوع باعث پدیدار شدن فضایی وسیع برای کاربرد هستان‌نگارها شده است. از نظر «آچولد و گرونینگر» سه کاربرد اساسی هستان‌نگار در بهبود ارتباطات[1]، افزایش قابلیت همکاری[2] و مهندسی سیستم‌هاست (Uschold and Gruninger 1996). کاربرد هستان‌نگار در حوزه‌های مختلف علم، فناوری و تجارت همچنان در حال توسعه است. به‌عنوان مثال، می‌توان از هستان‌نگار برای مدل‌سازی مفهومی داده (Gailly and Poels 2010) و یا در موتورهای جست‌وجو برای غنی‌سازی متن پرس‌وجوهای کاربران (با اضافه‌کردن مترادف‌ها و واژگان مرتبط) و در نتیجه، بهبود نتایج جست‌وجو استفاده کرد (Paolucci et al. 2002).

ساخت هستان‌نگار در عمل، کاری پرهزینه و دشوار است. محققان در راستای کاهش هزینه‌های ساخت و همچنین، افزایش دقت و کیفیت محصول نهایی، روش‌های گوناگونی معرفی کرده‌اند. در هر یک از این روش‌ها سعی شده است از دیدگاهی به فرایند ساخت هستان‌نگار توجه شده و الگویی نظام‌مند به‌صورت دستی یا نیمه‌خودکار برای ساخت ارائه داده شود. در این روش‌ها با وجود داشتن هدف مشترک، تفاوت‌های زیادی هست و در آن‌ها به برخی جنبه‌های فرایند ساخت توجه کافی نشده است. همچنین، اغلب آن‌ها دچار کمبود جزئیات در توصیف روش‌ها و فعالیت‌های به‌کاررفته هستند.

در ادامهٔ مقاله، ابتدا مختصری از پیشینهٔ نظری موضوع بیان شده و در آن برخی از پُراستنادترین روش‌های ساخت هستان‌نگار بررسی و در یک چارچوب با یکدیگر مقایسه شده‌اند. در بخش سوم مقاله، «انتیرانداک» به‌عنوان یک سامانهٔ یکپارچهٔ ساخت هستان‌نگار معرفی و در خصوص معماری و پیمانه‌های نرم‌افزاری[3] آن بحث شده است. در بخش چهارم، پیاده‌سازی و ارزیابی عملی این روش مورد بررسی قرار گرفته و در نهایت، بخش پنجم و انتهایی مقاله نتایج حاصل از به‌کارگیری «انتیرانداک» را بیان می‌کند.

1. communication
2. inter-operability
3. software modules

www.SID.ir

Archive of SID

پیشینهٔ نظری

برای ساخت هستان‌نگارها روش‌های متعددی پیشنهاد شده است. از نظر درجهٔ خودکاربودن می‌توان روش‌های ساخت را به دو دستهٔ *غیرخودکار (دستی)*[1] و *(نیمه) خودکار*[2] تقسیم کرد. در روش‌های دستی، دانش مفهومی توسط افراد کدگذاری می‌شود و در رویکرد نیمه‌خودکار سعی می‌شود با به‌کارگیری روش‌ها و الگوریتم‌های مختلف، مداخلهٔ انسان در مراحل مختلف تا حد امکان کمتر شده و در نتیجه، سرعت و هزینهٔ ساخت کاهش یابد.

در این بخش ابتدا پُراستنادترین روش‌های ساخت هستان‌نگار به‌صورت مختصر توضیح داده می‌شود و سپس، مقایسه‌ای بین آن‌ها انجام می‌گیرد.

مروری بر روش‌شناسی‌های موجود برای توسعهٔ هستان‌نگار

سایک[3]: در اواسط دههٔ ۸۰ میلادی «شرکت MMC»[4] شروع به ساخت «سایک» به‌عنوان یک پایگاه عظیم دانش عمومی بشر کرد. برای توسعهٔ آن فرایند زیر دنبال شده است (Lenat et al. 1990):

- ◇ کدگذاری دستی مقالات و قطعات دانش: این مرحله به‌صورت دستی انجام شد. چون سامانه‌های موجود پردازش زبان طبیعی، دانش عمومی کافی برای انجام این کار را نداشتند.
- ◇ کدگذاری دانش با کمک ابزار و با استفاده از دانشی که در مرحلهٔ قبل در پایگاه «سایک» قرار گرفته بود.
- ◇ کدگذاری دانش بر اساس ابزار و با استفاده از دانش موجود در پایگاه «سایک»: در این مرحله، بیشتر کار بر عهدهٔ ابزار گذاشته شده است و کاربران تنها منابع را به «سایک» پیشنهاد می‌دهند.

تاکنون از این روش تنها برای توسعهٔ پایگاه دانش «سایک» استفاده شده است.

---

1. manual
2. (semi) automatic
3. Cyc
4. Microelectronics and Computer Technology Corporation

www.SID.ir

Archive of SID

اِنترپرایز آنتولوژی[1]: این روش را «آچولد و کینگ» در سال ۱۹۹۵ ارائه دادند و در سال بعد بهینه‌سازی شد (Uschold and Gruninger 1996). بر اساس این روش برای ساخت هستان‌نگار باید فرایندهای زیر انجام شود:

۱. تعیین هدف: شامل تعریف هدف و قلمرو هستان‌نگار بوده و منظور از آن این است که چرا باید هستان‌نگار ساخته شود و کاربردهای مورد انتظار آن و واژگان مرتبط با آن در حوزهٔ مورد نظر چیست.
۲. ساخت هستان‌نگار: این فرایند شامل سه مرحلهٔ دریافت[2] هستان‌نگار (شناسایی مفاهیم کلیدی و ارتباط آن‌ها)، کدگذاری مفاهیم و روابط و یکپارچه‌سازی (استفادهٔ مجدد از هستان‌نگارهای موجود) است.
۳. ارزیابی: بررسی میزان انطباق هستان‌نگار ساخته‌شده با مشخصات هدف‌گذاری‌شدهٔ آن و قابلیت پاسخگویی هستان‌نگار به سؤالات شایستگی در این مرحله انجام می‌شود.
۴. مستندسازی: برای انجام این فرایند راهنماهایی متناسب با هدف و انواع مختلف هستان‌نگار ارائه شده است.

تاو[3]: این روش که بر اساس تجارب «گرونینگر و فاکس»[4] در اجرای پروژهٔ «تاو» در سال ۱۹۹۵ پیشنهاد شده است، بیشتر روی فرموله‌کردن و ارزیابی هستان‌نگار تمرکز دارد. مراحل پیشنهادی آن عبارت‌اند از:

۱. مشخص‌کردن مسئله و محدودهٔ هستان‌نگار: برای این کار مجموعه سناریوهایی که انگیزهٔ ساخت هستان‌نگار بوده و کاربردهای آن در آینده را مشخص می‌کنند، نوشته شده و مجموعه راه‌حل‌های ممکن برای هر سناریو مشخص می‌شود. این پاسخ‌ها ایدهٔ اولیهٔ معانی غیررسمی اشیاء و روابط آن‌هاست؛
۲. ایجاد سؤالات شایستگی غیررسمی بر اساس مرحلهٔ اول: سؤالات شایستگی که به زبان طبیعی نوشته می‌شوند، مشخص می‌کند که هستان‌نگار باید به چه سؤالاتی در خصوص سناریوهای نگارش‌شدهٔ مرحلهٔ قبل پاسخ دهد. در واقع،

1. Enterprise Ontology
2. capture
3. TOVE (Toronto Virtual Enterprise)
4. Gruninger and Fox

www.SID.ir

Archive of SID

نقش این سؤالات، مشخص کردن نیازمندی‌های هستان‌نگار است؛

۳. مشخص کردن مجموعهٔ اصطلاحات هستان‌نگار در زبان منطق مرتبهٔ اول؛

۴. رسمی‌سازی سؤالات شایستگی (با منطق مرتبهٔ اول)؛

۵. تعریف قواعد و تعاریف برای اصطلاحات (با منطق مرتبهٔ اول)؛

۶. آزمون راه‌حل‌ها از نظر کامل‌بودن در پاسخگویی به سؤالات شایستگی.

مت‌آنتولوژی[1]: این روش که در سال ۱۹۹۷ توسط «فرناندز، گومز-پرز و جوریستو» پیشنهاد شده، ریشه در فعالیت‌های اصلی چرخهٔ توسعهٔ نرم‌افزار دارد و از این نظر بالغ‌ترین روش موجود است (Fernández-López And Gómez-Pérez & Juristo 2002). این روش شامل فرایند توسعهٔ هستان‌نگار، یک چرخهٔ عمر مبتنی بر الگوهای تکاملی و تکنیک‌هایی برای انجام هر یک از فعالیت‌های مدیریتی، توسعه و پشتیبانی است. ابزارهای مختلفی چون ODE (Blázquez et al. 1998)، WebODE (Corcho et al. 2005)، «پروتژه»[2] (Noy, Fergerson and Musen 2000) و «آن‌تواِدیت»[3] (Sure et al. 2002) از این روش پشتیبانی می‌کنند.

فرایند توسعهٔ هستان‌نگار به فعالیت‌هایی اشاره دارد که باید برای ساخت هستان‌نگار انجام شود. توصیه شده است که برای جلوگیری از هرج‌ومرج در ساخت هستان‌نگار سه دسته از فعالیت‌ها انجام گیرد:

- فعالیت‌های مدیریتی: این فعالیت‌ها شامل برنامه‌ریزی[4]، کنترل، و تضمین کیفیت است. در فعالیت برنامه‌ریزی، کارهایی که باید انجام شود شناسایی شده، سپس، ترتیب انجام و زمان و منابع مورد نیاز برای هر کدام تعیین می‌شود. فعالیت کنترل، اجرای صحیح برنامه‌ریزی را تضمین می‌کند و فعالیت تضمین کیفیت نیز جهت کسب اطمینان از رضایت‌بخش بودن محصول نهایی انجام می‌شود.
- فعالیت‌های توسعه: این فعالیت‌ها به سه دستهٔ قبل از توسعه، توسعه و پس از توسعه تقسیم می‌شوند. در فعالیت‌های پیش از توسعه، مطالعهٔ محیطی برای شناخت بستر کاربردی هستان‌نگار صورت می‌گیرد. دستهٔ فعالیت‌های زمان توسعه، فعالیت تعریف مشخصات، دلیل ساخت، کاربردها و کاربران هستان‌نگار را مشخص می‌کند. فعالیت

1. METHONTOLOGY
2. Protégé
3. OntoEdit
4. scheduling

www.SID.ir

Archive of SID

مفهوم‌سازی، دانش حوزه را مدل‌سازی می‌کند. فعالیت رسمی‌سازی، مدل مفهومی ساخته‌شده را به مدلی رسمی یا نیمه‌رسمی تبدیل، و در نهایت، فعالیت پیاده‌سازی، مدل را به یک زبان هستان‌نگار پیاده‌سازی می‌کند. مرحلهٔ پس از توسعه، شامل فعالیت نگهداری برای به‌روزآوری و تصحیح‌های مورد نیاز بر روی هستان‌نگار است.

- ◇ فعالیت‌های پشتیبانی: شامل کسب دانش (استخراج دانش از خبرگان حوزه یا مستندات موجود توسط فرایندهای نیمه‌خودکار یادگیری هستان‌نگار)، ارزیابی (قضاوت فنی در مورد هستان‌نگار)، یکپارچه‌سازی (استفادهٔ مجدد از سایر هستان‌نگارهای موجود)، مستندسازی (ثبت دقیق جزئیات انجام هر مرحله و محصول آن) و مدیریت پیکره‌بندی (ثبت و ضبط تمام نسخه‌های مستندات و کد هستان‌نگار) است.

در فرایند توسعهٔ هستان‌نگار، مجموعه فعالیت‌های مورد نیاز جهت ساخت هستان‌نگار مشخص می‌شود، ولی در خصوص ترتیب انجام کارها صحبتی نمی‌شود. تعیین ترتیب انجام کارها بر عهدهٔ چرخهٔ زندگی هستان‌نگار است. این روش یک چرخهٔ عمر هستان‌نگار مبتنی بر الگوسازی تکاملی ارائهٔ می‌دهد. به این معنا که در هر بار الگوسازی، نسخهٔ جدیدی از هستان‌نگار با حذف، اضافه و تغییر واژگان الگوی قبلی ایجاد می‌شود. ساخت هر الگو با برنامه‌ریزی شروع می‌شود. در برنامه‌ریزی باید مشخص شود چه کارهایی و طبق چه زمان‌بندی باید انجام شوند و چه منابعی برای انجام هر کار باید تخصیص داده شوند. پس از برنامه‌ریزی، فعالیت توصیف مشخصات شروع شده و به‌طور همزمان کارهای مختلف مدیریتی (کنترل و تضمین کیفیت) و پشتیبانی (کسب دانش، یکپارچه‌سازی، مستندسازی و مدیریت پیکره‌بندی) آغاز می‌شوند. با ساخته‌شدن اولین الگو، مدل مفهومی درون فعالیت مفهوم‌سازی ساخته شده است. پس از آن پیاده‌سازی و رسمیت‌بخشی شروع می‌شود. اگر بعد از انجام این فعالیت‌ها نقصانی پیدا شد، می‌توان به عقب برگشت و اصلاحات لازم را انجام داد.

توسعهٔ هستان‌نگار ۱۰۱: «نوی و مک‌گینس» این روش را بر اساس این ایده که هیچ راه یگانه‌ای برای مدل‌سازی‌کردن یک حوزه وجود ندارد، ارائه دادند (Noy and McGuinness 2001). انتخاب بهترین الگو از بین گزینه‌های موجود، به کاربرد مورد انتظار بستگی دارد. برای توسعهٔ هستان‌نگاری که جهان واقع را به بهترین نحو ممکن منعکس کند، باید هستان‌نگار را در کاربرد مورد نظر ارزیابی کرد. این ارزیابی‌ها اغلب منجر به اعمال

www.SID.ir

Archive of SID

اصلاحات می‌شود. در نتیجه، می‌توان توسعهٔ هستان‌نگار را فرایندی بازگشت‌شونده[1] دانست. این روش ۷ مرحله را پیشنهاد می‌کند: تعیین حوزه و محدوده، بررسی هستان‌نگارهای موجود جهت استفادهٔ مجدد، شمارش اصطلاحات مهم در هستان‌نگار (برای این کار تجویز خاصی ندارد)، تعریف کلاس‌ها و سلسله‌مراتب در یک ابزار مدل‌سازی، تعریف خصوصیت‌های هر کلاس، تعریف ویژگی[2] هر خصوصیت[3]، و ساخت نمونه.

آن-تو-نالج[4]: این روش پیشنهاد می‌کند که برای ساخت هستان‌نگارها به کاربرد آن‌ها در آینده توجه شود (Sure and Studer 2002)، به همین دلیل هستان‌نگارهای توسعه‌یافته با این روش وابسته به کاربرد هستند. ویژگی دیگر این روش، تکیهٔ آن بر یادگیری هستان‌نگار برای کاهش زمان و انرژی لازم برای ساخت هستان‌نگار است. فرایندهای پیشنهادی این روش عبارت‌اند از:

۱. مطالعهٔ امکان‌سنجی: لازم است قبل از شروع و به‌عنوان اساس مرحلهٔ آغازین[5] انجام گیرد.
۲. مرحلهٔ آغازین: نتیجهٔ این فرایند سند شرح نیازها[6]ست که حوزه، هدف، راهنمای طراحی (به‌عنوان مثال، قواعد نامگذاری)، منابع دانشی موجود (کتاب، مجله، مصاحبه و مانند آن)، کاربران بالقوه و موارد کاربرد هستان‌نگار را مشخص می‌کند. از سؤالات شایستگی می‌توان در این سند استفاده کرد. این سند به سازندگان هستان‌نگار کمک می‌کند در مورد شامل‌شدن یا نشدن مفاهیم در هستان‌نگار و سلسله‌مراتب مفاهیم تصمیم‌گیری کنند.
۳. پالایش: در این فرایند، پیش‌نویس هستان‌نگار تولیدشده در مرحلهٔ قبل در تعامل با خبرگان حوزه کامل می‌شود. پس از تکمیل، هستان‌نگار باید با یک زبان هستان‌نگار پیاده‌سازی شود. ویرایشگرِ «آن‌تواِدیت»[7] برای این کار پیشنهاد شده است.

---

1. iterative
2. facet
3. slots
4. On-To-Knowledge
5. kickoff
6. Requirement Specification
7. OntoEdit

www.SID.ir

Archive of SID

۴. ارزیابی: در این فرایند با بررسی نیازمندی‌ها و سؤالات شایستگی مشخص می‌شود آیا هستان‌نگار می‌تواند نیازمندی‌های توصیف‌شده را برآورده سازد و به سؤالات شایستگی پاسخ گوید یا خیر. همچنین، هستان‌نگار در محیط کاربرد آن آزمون می‌شود. فرایند پالایش و ارزیابی به‌صورت یک چرخه تا زمان رسیدن به حد مطلوب اجرا خواهند شد.
۵. نگهداری: باید مشخص شود چه کسی مسئول به‌روزآوری و نگهداری است و چگونه این کار انجام می‌شود.

دیلیگنت[1]: «پینتو، تمپیچ و استاب» این روش را با رویکردی کاملاً متفاوت از سایر روش‌های توسعهٔ هستان‌نگار ارائهٔ دادند (Pinto, Tempich and Staab 2009). قابلیت همکاری بین تمام ذی‌نفعان، نقشی اساسی در این روش دارد و ایدهٔ اصلی، شروع از یک هستان‌نگار کوچک و قابل استفادهٔ به اشتراک گذاشته شده است. ذی‌نفعان بر اساس اهداف خود، نسخه‌ای را اقتباس کرده و به‌صورت محلی استفاده می‌کنند. یک شورای کنترلی از ذی‌نفعان تمام هستان‌نگارهای محلی را بررسی و نسخهٔ جدیدی از هستان‌نگار اشتراکی توسعه خواهد داد. این روش ۵ گام دارد: ۱. ساخت یک هستان‌نگار ساده از ابتدا؛ ۲. تطابق محلی[2]: هستان‌نگار توسط هر یک از ذی‌نفعان به‌صورت جداگانه استفاده و اختصاصی‌سازی می‌شود؛ ۳. تحلیل: شورای ذی‌نفعان بر اساس درخواست‌های رسیده و مشاهدهٔ تغییرات محلی تصمیم می‌گیرد چه تغییراتی بیشترین مزایا را برای کاربران دارد و باید در هستان‌نگار اشتراکی اضافه شود؛ ۴. بازبینی: شورا نسخهٔ بازبینی‌شده از هستان‌نگار را می‌سازد؛ ۵. به‌روزرسانی محلی: کاربران نسخهٔ جدید را در سیستم خود قرار می‌هند. بعد از این مرحله به مرحلهٔ ۲ برگشت می‌شود.

اِپان[3]: این روش را «نیکلا، میسیکف و نویگلی» با اقتباس از فرایند توسعهٔ نرم‌افزار ارائه داده است (De Nicola, Missikoff and Navigli 2009). فرایند توسعه در یک چرخهٔ جریان کاری[4] (نیازمندی، تحلیل، طراحی، پیاده‌سازی و آزمون) سازماندهی می‌شود. هر یک از مراحل این چرخه دارای تعدادی گام و دور تکرار است (آغاز، بسط جزئیات، ساخت و گذار) و از

1. DILIGENT (DIstributed Loosely-controlled evolvInG Engineering of oNTologies)
2. local adaptation
3. UPON (Unified Process for ONtology
4. work flow

www.SID.ir

Archive of SID

مشارکت خبرگان، به‌ویژه در مراحل اولیهٔ توسعه استفاده می‌شود.

نئون[1]: این روش را «سوآرز-فیگورا» در سال ۲۰۱۰ بر پایهٔ روش‌های «دیلیگنت»، «مت‌آنتولوژی» و «آن-تو-نالج» ارائهٔ داده است (Suarez-Figueroa 2010). این روش چهار پایهٔ اصلی دارد:

۱. واژه‌نامهٔ فرایندها و فعالیت‌ها: به‌دلیل ماهیت روبه‌تکامل مهندسی هستان‌نگار برخی اصطلاحات در این حوزه ممکن است دارای ابهام باشند. به همین منظور، فهرست واژگانی شامل ۵۹ فعالیت و فرایند به زبان انگلیسی ارائه شده که افراد درگیر پروژه در ارتباط با هم دچار ابهام نشوند.
۲. سناریوهای ساخت هستان‌نگار: بر اساس تجربیات حاصل از انجام چندین پروژه، سه وضعیت متفاوت ساخت هستان‌نگار (هستان‌نگار منفرد، مجموعهٔ هستان‌نگارهای منفرد در تعامل، شبکه‌های هستان‌نگار) تعریف می‌شود. بر همین اساس، ۹ سناریوی مختلف برای ساخت هستان‌نگار معرفی شده که تأکید ویژهٔ آن بر روی استفادهٔ مجدد از منابع دانش موجود و بازمهندسی است.
۳. دو مدل چرخهٔ زندگی هستان‌نگار
۴. مجموعه راهنماهایی برای انجام فرایندها و فعالیت‌ها

مقایسهٔ روش‌ها

چارچوب مقایسهٔ مورد استفاده در این پژوهش شامل میزان پوشش سه دستهٔ اصلی فعالیت مورد نیاز برای ساخت هستان‌نگار (چارچوب مقایسهٔ پیشنهادی Gómez-Pérez, Fernández-López and Corcho 2004)، پشتیبانی از زبان غیرانگلیسی، ابزار پشتیبانی فنی و میزان بیان جزئیات انجام کار است. جدول شماره ۱، نتیجهٔ این مقایسه را نشان می‌دهد.

همان‌گونه که در این جدول مشاهده می‌شود، تمام روش‌های مورد بررسی دارای خلأهایی در پشتیبانی از فرایندهای اصلی توسعهٔ هستان‌نگار هستند و تنها روش ترکیبی «نئون» این پشتیبانی را بر اساس سناریوهای نه‌گانهٔ خود انجام می‌دهد. با وجود این‌که در بسیاری از موارد، منابع دانشی قابل استفاده در فعالیت اکتساب دانش را متون غیرانگلیسی تشکیل می‌دهند، در این روش‌ها به چگونگی و روند به‌کارگیری و پشتیبانی از چنین

1. Neon

www.SID.ir

Archive of SID

منابعی اشاره نشده است.

ابزارهای پشتیبانی فنی از این روش‌ها بیشتر روی فرایند توسعه (به‌خصوص بخش پیاده‌سازی) تمرکز کرده‌اند و برای پشتیبانی فنی از سایر فعالیت‌ها توسعه‌دهندگان باید از ابزارهای دیگر در کنار ابزار اصلی استفاده کنند. به‌عنوان نمونه، هیچ‌یک از این نرم‌افزارها بخشی برای مدیریت پروژه (تعریف و مدیریت اعضای تیم، تخصیص و پیگیری وظایف و منابع)، مدیریت جلسات (برنامه‌ریزی و ثبت جلسات کسب دانش از خبرگان)، پرسشنامهٔ الکترونیکی و سامانهٔ مکاتبات (ارسال و دریافت و مدیریت مستندات تبادلی بین اعضای تیم توسعه‌دهنده و خبرگان) ندارند.

جدول ۱. چارچوب مقایسهٔ روش‌های توسعهٔ هستان‌نگار

روش‌شناسی	فعالیت‌های مدیریتی	فعالیت‌های پیش‌توسعه	فعالیت‌های توسعه	فعالیت‌های پس از توسعه	فعالیت‌های پشتیبانی	حمایت از زبان دیگر	ابزار	داشتن جزئیات
انترپرایز آنتولوژی	–	–	توصیف نیازمندی‌ها، پیاده‌سازی	–	کسب دانش، ارزیابی، یکپارچه‌ساز، مستندسازی	–	–	خیر
مت‌آنتولوژی	برنامه‌ریزی، کنترل، تضمین کیفی	–	توصیف نیازمندی‌ها، مفهوم‌سازی، رسمیت‌بخشی، پیاده‌سازی	نگهداری	کسب دانش، ارزیابی، یکپارچه‌سازی، مدیریت پیکره‌بندی، مستندسازی	–	پروتژه، آن‌توادیت	بلی
سایک	–	–	پیاده‌سازی	–	کسب دانش، مستندسازی	–	ابزار خاص	خیر
تاو	–	–	توصیف نیازمندی‌ها، مفهوم‌سازی، رسمیت‌بخشی، پیاده‌سازی	–	کسب دانش، ارزیابی، یکپارچه‌سازی، مستندسازی	–	–	خیر
توسعهٔ هستان‌نگار ۱۰۱	–	–	توصیف نیازمندی‌ها، مفهوم‌سازی، رسمیت‌بخشی، پیاده‌سازی	–	کسب دانش، یکپارچه‌سازی	–	–	خیر

www.SID.ir

Archive of SID

روش‌شناسی	فعالیت‌های مدیریتی	فعالیت‌های پیش‌توسعه	فعالیت‌های توسعه	فعالیت‌های پس از توسعه	فعالیت‌های پشتیبانی	حمایت از زبان دیگر	ابزار	داشتن جزئیات
دیلیگنت	برنامه‌ریزی، کنترل، تضمین کیفی	–	–	نگهداری	کسب دانش، ارزیابی، مدیریت پیکره‌بندی	–	–	خیر
اِپان	–	–	توصیف نیازمندی‌ها، مفهوم‌سازی، رسمیت‌بخشی، پیاده‌سازی	نگهداری	کسب دانش، ارزیابی	–	–	خیر
آن-تو-نالج	برنامه‌ریزی، کنترل، تضمین کیفی	مطالعهٔ محیطی، امکان‌سنجی	توصیف نیازمندی‌ها، مفهوم‌سازی، رسمیت‌بخشی، پیاده‌سازی	نگهداری، استفاده	کسب دانش، ارزیابی، یکپارچه‌سازی، پیکره‌بندی، مستندسازی	–	آن‌توادیت	خیر
نئون	برنامه‌ریزی، کنترل، تضمین کیفی	مطالعهٔ محیطی، امکان‌سنجی	توصیف نیازمندی‌ها، مفهوم‌سازی، رسمیت‌بخشی، پیاده‌سازی	نگهداری، استفاده	کسب دانش، ارزیابی، مدیریت پیکره‌بندی، مستندسازی، ادغام	بومی‌سازی	آن‌توادیت، پروتژه	بلی

با توجه به نقاط ضعف بیان‌شده در استفاده از روش‌های موجود، در ادامه، سامانهٔ یکپارچهٔ توسعهٔ هستان‌نگار (انتیرانداک) به‌عنوان یک روش جدید توسعهٔ هستان‌نگار معرفی و در خصوص آن بحث می‌شود.

معرفی انتیرانداک

سامانهٔ یکپارچهٔ توسعهٔ هستان‌نگار (انتیرانداک) با الهام از روش‌های موجود، به‌خصوص «آن-تو-نالج»، «مت‌آنتولوژی» و «نئون» و افزودن ویژگی‌های جدید به‌منظور پوشش نقاط ضعف سایر روش‌ها ایجاد شده است. «انتیرانداک» تنها یک ابزار توسعه نیست، بلکه یک روش کامل شامل توصیف فرایند، فعالیت‌ها و روش‌هایی است که به‌منظور توسعهٔ یکپارچهٔ هستان‌نگار استفاده می‌شوند.

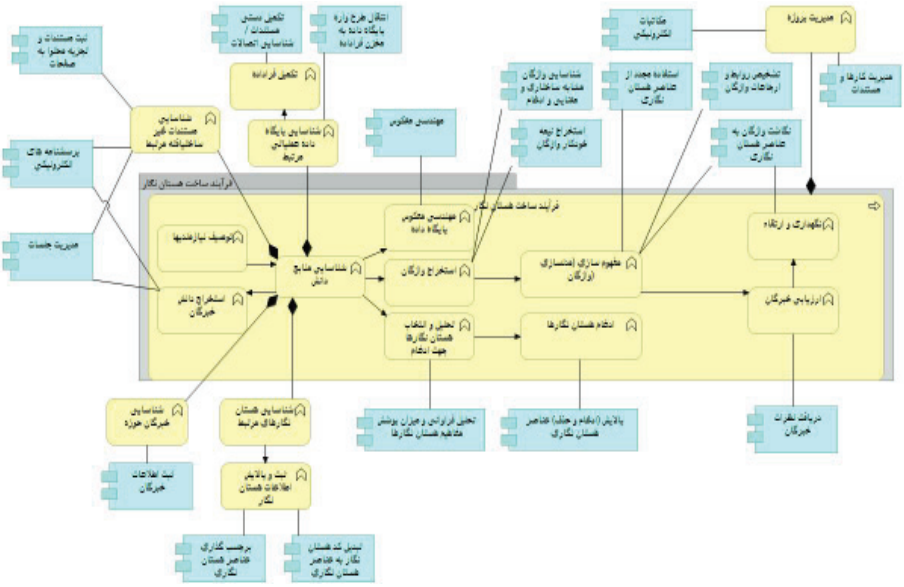
www.SID.ir

Archive of SID

معماری انتیرانداک

معماری «انتیرانداک» در شکل شماره ۱، نشان داده شده است[۱]. این معماری با توجه خاص بر پوشش فعالیت‌های مهم در فرایند ساخت هستان‌نگار و همچنین، امکان انجام مشارکتی این فرایند طراحی شده است.

فعالیت‌های مورد نیاز جهت انجام فرایند با آنچه که در سه روش «مت‌آنتولوژی، «نئون» و «آن-تو-نالج» معرفی شده، از نظر کلی مشابه هستند، ولی در جزئیات تفاوت‌هایی دارند که در ادامه، به تفصیل در خصوص آن‌ها بحث می‌شود. برای اجرای هر یک از فعالیت‌ها، پیمانه (ماژول) یا پیمانه‌های نرم‌افزاری در نظر گرفته شده است. طراحی پیمانه‌بندی ابزار در «انتیرانداک» این امکان را فراهم می‌آورد که هر یک از پیمانه‌ها با وجود یکپارچگی و هماهنگی با سایر پیمانه‌ها در سطح داده (شکل ۲)، به‌صورت مستقل قابلیت ارتقاء داشته و بتوان ویژگی‌های جدیدی به آن‌ها اضافه کرد.



شکل ۱. معماری «انتیرانداک»

فعالیت‌های شناسایی‌شده در این معماری به شرح زیر هستند:

---

۱. این شکل با زبان توصیف معماری «آرچیمیت» از زبان‌های توصیف معماری ذکرشده در استاندارد ISO/IEC/IEEE 42010 ترسیم شده است.

www.SID.ir

Archive of SID

۱. توصیف نیازمندی‌ها: تقریباً تمامی روش‌ها نیاز به وجود این فعالیت را برای انجام فرایند ساخت مورد تأیید قرار داده‌اند. در این فعالیت اهداف و نیازمندی‌های ساخت هستان‌نگار در قالب یک سند با عنوان مستند نیازها مشخص می‌شود. در «انتیراناداک» پیشنهاد می‌شود تیم توسعه‌دهنده این سند را به‌عنوان یکی از منابع دانشی به سامانه معرفی کند. مستند نیازها، راهنمای تیم توسعه برای جست‌وجوی منابع دانشی مرتبط و همچنین، مبنایی برای ارزیابی صحت و کامل‌بودن محصول نهایی خواهد بود. به‌دلیل تعبیهٔ زیرساخت ممیزی عملیات و مدیریت نسخه‌بندی، ارتباط بین مفاهیم ایجادشده در هستان‌نگار و مستند نیازها در مرحلهٔ ارزیابی هستان‌نگار کاملاً واضح خواهد بود و این موضوع کمک زیادی به انجام دقیق و سریع ارزیابی توسط خبرگان خواهد کرد.

۲. شناسایی منابع دانش: فعالیت‌های مطالعهٔ محیطی و امکان‌سنجی به‌عنوان فعالیت‌های پیش از توسعه، در دو روش «آن-تو-نالج» و «نئون» مورد اشاره قرار گرفته‌اند. در «انتیراناداک» این فعالیت‌ها در قالب فعالیت کلی شناسایی منابع دانشی مورد نیاز برای توسعهٔ هستان‌نگار مورد توجه هستند، با انجام این فعالیت مشخص می‌شود چه میزان دانش در قالب‌های مختلف برای ساخت هستان‌نگار در دسترس است و اصولاً نیازی به ساخت هستان‌نگار جدید وجود دارد یا خیر. برای انجام این فعالیت، اجرای موازی چند فعالیت پیشنهاد می‌شود:

۲-۱. شناسایی هستان‌نگارهای مرتبط موجود: استفادهٔ مجدد از هستان‌نگارهایی که توسط توسعه‌دهندگان دیگر برای حوزهٔ موضوعی تولید شده‌اند، می‌تواند به میزان قابل توجهی در سرعت ایجاد هستان‌نگار و دقت آن مؤثر باشد. این امکان نیز وجود دارد که هستان‌نگاری یافت شود که تا حد قابل قبولی هدف ایجاد هستان‌نگار را پوشش دهد و در نتیجه، نیاز به ادامهٔ کار وجود نداشته باشد. پس از انجام این فعالیت، انجام فعالیت ثبت و پالایش اطلاعات هستان‌نگارها پیشنهاد شده است که با استفاده از دو پیمانهٔ نرم‌افزاری تبدیل کد هستان‌نگار به عناصر هستان‌نگاری و برچسب‌گذاری عناصر هستان‌نگاری انجام می‌شود. پیمانهٔ تبدیل کد، فایل

www.SID.ir

Archive of SID

هستان‌نگار پیداشده را به زبان OWL[1] دریافت و پس از جداسازی عناصر اصلی هستان‌نگاری (کلاس‌ها و خصوصیات) آن‌ها را در مخزن عناصر هستان‌نگاری ذخیره می‌کند. توسعه‌دهندگان هستان‌نگار می‌توانند با استفاده از پیمانهٔ برچسب‌گذاری، به هر یک از عناصر هستان‌نگار یک برچسب معنایی به‌زبان فارسی نسبت دهند. این پیمانه که قابلیت شناسایی و پیشنهاد برچسب بر اساس مشابهت ساختاری و معنایی را دارد، معادلی برای فعالیت بومی‌سازی است که تنها در روش «نئون» به آن اشاره شده است. برچسب‌گذاری عناصر به زبان فارسی که لایهٔ معنایی جدیدی برای عناصر هستان‌نگاری فراهم می‌کند، یکی از مزایای قابل توجه «انتیرانداک» نسبت به سایر روش‌ها برای استفاده در ساخت هستان‌نگارهای بومی است.

۲-۲. شناسایی مستندات غیرساخت‌یافتهٔ مرتبط با حوزه: مستندات غیرساخت‌یافته (متنی) معمولاً به‌عنوان منبع اصلی دانش حوزه شناخته می‌شود. تنوع اسناد قابل استفاده ممکن است بسیار گسترده باشد. در این فعالیت لازم است اسناد از نظر میزان دانش و اعتبار آن‌ها اولویت‌بندی شوند (به‌عنوان مثال، اصطلاح‌نامه‌های حوزه و همچنین، اسناد بالادستی سازمانی مانند سند برنامهٔ راهبردی و شاخص‌های ارزیابی حوزهٔ کسب‌وکار از معتبرترین و مهم‌ترین اسناد دانشی هستند). برای انجام این فعالیت استفاده از سه پیمانهٔ نرم‌افزاری پیشنهاد می‌شود. توسعه‌دهندگان پس از یافتن هر مستند با استفاده از پیمانهٔ ثبت مستندات آن را در یک مخزن ذخیره‌سازی می‌کنند. این پیمانه علاوه بر ثبت فایل، امکان ثبت محتوای هر صفحه را به‌صورت مجزا فراهم می‌کند. استخراج محتوای مستند و ذخیره‌سازی آن در صفحات مجزا امکان اعمال روش‌های استخراج نیمه‌خودکار واژگان از متن با روش‌های آماری و زبان‌شناختی (در پیمانهٔ استخراج نیمه‌خودکار واژگان) و انجام تحلیل هم‌رخدادی برای تشخیص روابط معنایی بین واژگان استخراج‌شده (در پیمانهٔ نرم‌افزاری تشخیص روابط و ارجاعات واژگان) و همچنین، نمایش منبع هر عنصر مدل‌سازی‌شده در هستان‌نگار نهایی را (در پیمانهٔ دریافت

1. Web Ontology Language (WOL)

www.SID.ir

Archive of SID

نظـرات خبـرگان) بـرای افزایـش دقـت و صحـت عمـل ارزیابـی هسـتان‌نگار فراهـم مـی‌آورد.

۲-۳. شناسـایی پایـگاه داده‌هـای عملیاتـی حـوزه: سـاختار و محتـوای پایـگاه داده‌هـای عملیاتـی موجـود می‌توانـد منبـع خوبـی بـرای اسـتخراج مفاهیـم و روابـط میـان آن‌هـا در حـوزهٔ مـورد نظـر باشـد. در ایـن فعالیـت بایـد تـلاش شـود سـامانه‌های عملیاتـی کـه حداکثـر پوشـش خودکارسـازی عملیـات را در حـوزهٔ مـورد نظـر دارنـد (به‌عنـوان مثـال، سـامانه‌های برنامه‌ریـزی منابـع سـازمان[1] در حـوزهٔ مـورد نظـر) شناسـایی، و دسـترس‌پذیری بـه پایـگاه داده‌هـای آن‌هـا بـرای انجـام مهندسـی معکـوس مشـخص شـود. طرحـوارهٔ[2] پایـگاه داده‌هـای مـورد نظـر بـا اسـتفاده از پیمانـهٔ انتقـال طرحـوارهٔ پایـگاه داده بـه مخـزن فـراداده به‌عنـوان نسـخهٔ اولیـه از سـاختار پایـگاه داده ذخیره‌سـازی می‌شـود. از آنجـا کـه معمـولاً طرحـوارهٔ پایـگاه داده شـامل فـرادادهٔ کامـل بـرای مهندسـی معکـوس نیسـت، فعالیـت بعـدی بـا عنـوان تکمیـل دسـتی مسـتندات بخشـی از روابـط بیـن عناصـر داده در طرحـواره بـه شـکل کلیدهـای خارجـی و اصلـی وجـود نداشـته و پـس از ایـن کار لازم اسـت فعالیـت تکمیـل فـراداده اجـرا شـود. در ایـن فعالیـت اطلاعات تکمیلـی در مـورد سـاختار پایـگاه داده توسـط پیمانـهٔ نرم‌افـزاری تکمیـل دسـتی مسـتندات (شـرح فیلدهـا و جـداول اطلاعاتـی)/ شناسـایی اتصـالات (شناسـایی ارتباطـات بیـن جـداول کـه در طرحـواره به‌شـکل کلیـد خارجـی وجـود ندارنـد) ثبـت می‌شـود.

۲-۴. شناسـایی خبـرگان حـوزه: خبـرگان، یـک منبـع دانشـی ارزشـمند بـرای شـناخت حـوزه هسـتند و اسـتفاده از نظـرات آن‌هـا جهـت ارزیابـی و اصـلاح هسـتان‌نگار باعـث اعتباربخشـی بـه محصـول نهایـی خواهـد شـد. در ایـن فعالیـت بایـد فهرسـتی از خبـرگان بـه همـراه اولویت‌بنـدی آن‌هـا از نظـر میـزان دانـش و اعتبـار و همچنیـن، میـزان دسـترس‌پذیری بـه آن‌هـا مشـخص شـود. چنانچـه ایـن فهرسـت اعضـای کمـی داشـته باشـد، توسـعه و ارزیابـی هسـتان‌نگار دشـواری

1. ERP (Enterprise Resource Planning)
2. schema

www.SID.ir

Archive of SID

بیشتری خواهد داشت و حتی ممکن است مقرون‌به‌صرفه تشخیص داده نشود. اطلاعات خبرگان با کمک پیمانهٔ نرم‌افزاری ثبت اطلاعات خبرگان ثبت و ضبط می‌شود. خبرگان ثبت‌شده در سامانه می‌توانند از طریق پیمانهٔ نرم‌افزاری دریافت نظرات خبرگان محصول نهایی را مشاهده و نظرات اصلاحی خود را ثبت کنند.

۳. مهندسی معکوس پایگاه داده: این فعالیت به‌همراه فعالیت‌های استخراج واژگان از متن و ادغام هستان‌نگارهای موجود، سه فعالیتی هستند که در راستای استخراج دانش از منابع شناسایی‌شده در فعالیت قبلی انجام می‌شوند. فعالیت مهندسی معکوس پایگاه داده، زمانی قابل انجام است که پایگاه دادهٔ مناسبی در حوزهٔ موضوعی پیدا شده و فرادادهٔ آن تکمیل شده باشد. خروجی این فعالیت که به کمک پیمانهٔ نرم‌افزاری مهندسی معکوس انجام می‌شود، هستان‌نگاری است که عناصر آن به مخزن عناصر هستان‌نگاری موجود در حوزهٔ موضوعی اضافه می‌شود. از این عناصر در فعالیت مفهوم‌سازی استفاده خواهد شد. روش پیشنهادی برای انجام مهندسی معکوس بر پایهٔ وجود یک فراداده تکمیل شده است که شامل شرح جداول و فیلدها و مشخص‌کردن ارتباطات ساختاری (کلیدهای خارجی در سطح پایگاه داده) و معنایی (ارتباط جداول در سطح برنامهٔ کاربردی) میان جداول است.

۴. استخراج واژگان: همان‌طور که اشاره شد، منابع غیرساخت‌یافته مهم‌ترین منابع دانشی حوزه هستند. در این فعالیت، واژگان مورد استفاده در مستندات غیرساخت‌یافته، که به‌صورت صفحه‌های مجزا جداسازی شده‌اند، توسط اعضای تیم توسعه و با کمک پیمانهٔ نرم‌افزاری استخراج نیمه‌خودکار واژگان به‌دست آمده و پس از پالایش و ادغام (با کمک پیمانهٔ نرم‌افزاری شناسایی واژگان ساختاری و معنایی) در مخزن واژگان ثبت و ضبط می‌شوند.

۵. تحلیل و انتخاب هستان‌نگارهای موجود برای ادغام: در این فعالیت، هستان‌نگارهای یافت‌شده در مرحلهٔ قبل با کمک پیمانهٔ نرم‌افزاری تحلیل فراوانی و میزان همپوشانی مفاهیم هستان‌نگارها بررسی شده و بر اساس اطلاعات آماری حاصل و همچنین میزان تناسب مفاهیم آن‌ها با حوزهٔ مورد نظر، در خصوص انتخاب هستان‌نگارهای مناسب برای ادغام، تصمیم‌گیری می‌شود.

www.SID.ir

Archive of SID

۶. ادغـام هسـتان‌نگارها: ایـن فعالیـت پـس از انجـام تحلیـل و انتخـاب هسـتان‌نگارهای موجـود بـرای ادغـام قابـل انجـام اسـت. در ایـن فعالیـت مجموعـهٔ هسـتان‌نگارهای منتخـب بـا اسـتفاده از پیمانـهٔ نرم‌افـزاری پالایـش عناصـر هسـتان‌نگاری بـا هـم ادغـام می‌شـوند و عناصـر هسـتان‌نگاری حاصـل در قالـب یـک هسـتان‌نگار جدیـد بـه مخـزن افـزوده می‌شـود.

۷. مفهوم‌سـازی واژگان: در ایـن فعالیـت توسـعه‌دهندگان بـا بررسـی فهرسـت واژگان مسـتخرج بـا کمـک پیمانه‌هـای نرم‌افـزاری اسـتفادهٔ مجـدد از عناصـر هسـتان‌نگاری و تشـخیص روابـط و ارجاعـات واژگان در خصـوص مدل‌سـازی‌کردن هـر واژه بـه یـک عنصـر هسـتان‌نگار، تصمیم‌گیـری کـرده و عنصـر مدل‌سازی‌شـده را بـا اسـتفاده از پیمانـهٔ نرم‌افـزاری نگاشـت واژگان بـه عناصـر هسـتان‌نگاری در مخـزن ثبـت می‌کننـد.

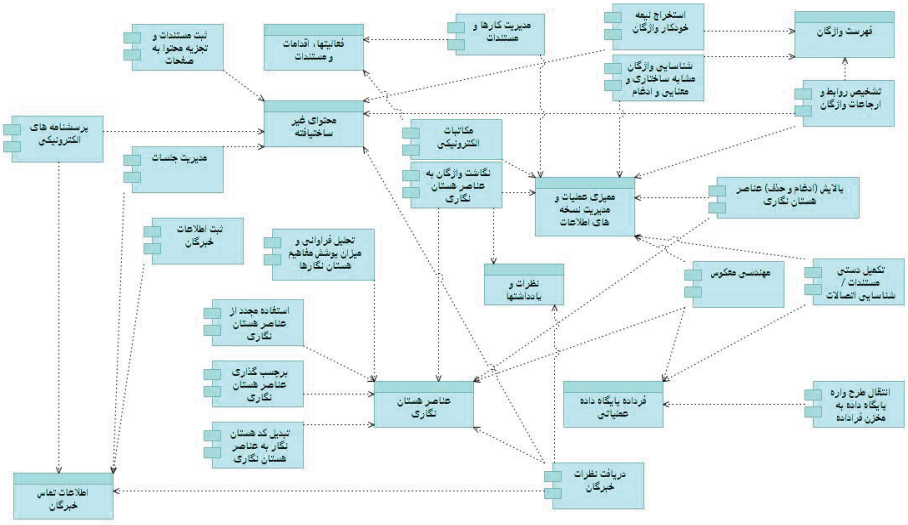
۸. ارزیابـی خبـرگان: پـس از انجـام فعالیت‌هـای توسـعه، هسـتان‌نگار هـدف شـکل گرفتـه اسـت و بایـد صحـت و دقـت آن مـورد بررسـی قـرار گیـرد. در ایـن فعالیـت، خبـرگان بـا اسـتفاده از پیمانـهٔ نرم‌افـزاری دریافـت نظـرات خبـرگان، در خصـوص تمامـی اجـزای هسـتان‌نگار تولیدشـده اظهـار نظـر می‌کننـد. ایـن نظـرات در اختیـار تیـم توسـعه قـرار می‌گیـرد تـا طـی فعالیـت بعـدی بـرای تصحیـح و بهبـود هسـتان‌نگار مـورد اسـتفاده قـرار گیـرد.

۹. نگهـداری و ارتقـاء: پـس از سـاخت نسـخهٔ اولیـهٔ هسـتان‌نگار لازم اسـت بـر اسـاس اطلاعـات حاصـل از ارزیابـی خبـرگان، تیـم توسـعه‌دهنده تغییـرات لازم را در هسـتان‌نگار اعمـال کنـد. ایـن کار ممکـن اسـت چندین‌بـار و به‌صـورت متنـاوب تـا رسـیدن بـه نتیجـهٔ دلخـواه تکـرار شـود. در نهایـت، بـا توجـه بـه پویایـی محیـط، هسـتان‌نگار نیـز بایـد به‌صـورت مـداوم تحـت بازبینـی و ارتقـاء باشـد تـا مـدل قابـل قبولـی از واقعیـت باقـی بمانـد.

۱۰. مدیریـت پـروژه: توسـعهٔ یـک هسـتان‌نگار بـا توجـه بـه میـزان زمـان و انـرژی لازم، اغلب نیازمنـد مشـارکت چندین‌نفـر در قالـب یـک تیـم توسـعه‌دهنده اسـت. «انتیرانـداک» پیشـنهاد می‌کنـد کـه فعالیـت مدیریـت پـروژه به‌عنـوان یـک فعالیـت مـوازی و جانبـی در کنـار سـایر فعالیت‌هـای ذکرشـده انجـام شـود. پیمانه‌هـای نرم‌افـزاری مکاتبـات الکترونیکـی و مدیریـت کارهـا و مسـتندات، بـه تیـم مشـارکت‌کننده در انجـام کار کمـک خواهـد کـرد.

www.SID.ir

Archive of SID



شکل ۲. ارتباط و یکپارچگی پیمانه‌های نرم‌افزاری در سطح داده

برای انجام هر یک از فعالیت‌های فوق یک یا چند پیمانهٔ نرم‌افزاری پیشنهاد شده است. هر پیمانهٔ نرم‌افزاری داده‌های مرتبط با خود را در یک یا چند مخزن داده‌ای که خود شامل جداول پایگاه داده رابطه‌ای هستند، ثبت و ضبط می‌کند. مخازن داده‌ای بین پیمانه‌ها به اشتراک گذاشته شده‌اند و به این ترتیب، هر پیمانه می‌تواند از داده‌های سایر پیمانه‌ها استفاده کرده و به‌صورت یکپارچه با بقیه عمل کند (شکل ۲).

پیمانه‌های نرم‌افزاری انتیرانداک

«انتیرانداک» ابزار خاص خود را برای حمایت از توسعهٔ هستان‌نگار پیشنهاد می‌دهد. این ابزار مجموعه‌ای از پیمانه‌های نرم‌افزاری است که در قالب زیرسامانه های مختلف طراحی شده‌اند. در ادامه، زیرسامانه‌ها، پیمانه‌های مرتبط و امکانات پیش‌بینی‌شده در آن‌ها به‌صورت مختصر معرفی شده است:

۱. مدیریت پروژه: این پیمانه امکانات مختلفی برای انجام فعالیت مدیریت پروژه دارد. در گام نخست توسعهٔ هستان‌نگار باید یک پروژه تعریف شود. تعریف پروژه شامل عنوان، تاریخ شروع و تاریخ پایان تخمینی، اعضای تیم و همچنین مراحل

www.SID.ir

Archive of SID

مهم[1] است. هر یک از اعضای تیم، نام کاربری و کلمهٔ عبوری دریافت خواهند کرد که از طریق آن می‌توانند با ورود به سایت سامانه کارتابل کار خود را مشاهده کرده و شرح دقیق کاری را که باید انجام دهند، مشاهده کنند. همچنین، می‌توانند نتایج کاری خود را تحت عنوان اقدامات به هر کار پیوست کنند. زمان‌های انجام کار به‌صورت تجمیع‌شده در گزارشات مدیریتی به مدیر پروژه نمایش داده خواهد شد و با توجه به ثبت قراردادهای کاری افراد در سامانه، امکان ثبت و ضبط امور مالی و پرداختی به اعضای تیم را فراهم می‌آورد. همچنین، کلیهٔ فعالیت‌هایی که افراد در سامانه انجام می‌دهند، در مخزن ممیزی عملیات ذخیره شده و در نتیجه، مستندسازی فعالیت‌ها در سطح ریزترین جزء به‌صورت خودکار انجام خواهد شد.

۲. مکاتبات الکترونیکی: این پیمانه امکان تبادل الکترونیکی متون و مستندات را بین اعضای پروژه فراهم می‌آورد. از سوابق پیام‌ها و فایل‌های رد و بدل شده در این پیمانه می توان در مستندسازی کار استفاده کرد.

۳. مدیریت جلسات: این پیمانه امکان تعریف و درخواست حضور در جلسات، تعیین دستور کار، ثبت و ضبط صورت‌جلسه‌ها و اقدامات و یادداشت‌های مرتبط با هر جلسه را فراهم می‌آورد. جلسات برگزارشده می‌توانند مربوط به برگزاری مصاحبه با خبرگان برای استخراج دانش باشد که در این صورت، از نتایج جلسه (فایل مستندات متنی) به‌عنوان منبع دانشی غیرساخت‌یافته استفاده خواهد شد و ممکن است جلسه مربوط به هماهنگی و امور اجرایی تیم توسعه‌دهنده باشد که در این صورت، نتیجهٔ جلسه به مستندسازی پروژه کمک خواهد کرد.

۴. پرسشنامهٔ الکترونیکی: با استفاده از این پیمانه اعضای تیم می‌توانند پرسشنامه طراحی کرده و سپس، برای جامعهٔ آماری مورد نظر خود حساب‌های کاربری موقت برای پرکردن پرسشنامه ایجاد کنند. پرسش‌شوندگان می‌توانند از طریق وب، اطلاعات پرسشنامه را تکمیل کنند. از این پیمانه می‌توان برای استخراج دانش از خبرگانی که دسترسی حضوری به آن‌ها مقدور نیست، استفاده کرد.

۵. ثبت اطلاعات خبرگان: این پیمانه امکان مدیریت فهرست خبرگان در دسترس را

1. Milestone (تاریخ‌هایی که باید خروجی مشخصی از پروژه گرفته شود)

www.SID.ir

Archive of SID

فراهم می‌آورد.

۶. مدیریت هستان‌نگارها: هستان‌نگارهایی که در طی فعالیت جست‌وجوی منابع دانشی پیدا شده‌اند و یا توسط فعالیت‌های ادغام هستان‌نگار، مفهوم‌سازی واژگان و یا مهندسی معکوس ایجاد می‌شوند، توسط این زیرسامانه ثبت و مدیریت می‌شوند. پیمانه‌های مختلفی که در این زیرسامانه استفاده می‌شوند، عبارت‌اند از:

- ◇ تبدیل کد هستان‌نگار به عناصر هستان‌نگاری: این پیمانه با دریافت کد هستان‌نگارهای موجود به زبان OWL آن‌ها را به عناصر هستان‌نگاری، تجزیه و در مخزنی ثبت می‌کند. مدیریت فراداده مربوط به هستان‌نگارها (عنوان، مسیر فایل در وب و ...) از دیگر امکانات این پیمانه است.
- ◇ برچسب‌گذاری عناصر هستان‌نگاری: این پیمانه امکاناتی فراهم می‌کند که کاربران بتوانند به هر یک از عناصر هستان‌نگارها، یک برچسب فارسی نسبت دهند. در این پیمانه نیز از الگوریتم‌های مشابهت‌یابی ساختار و معنایی برای پیداکردن برچسب‌های مشابه و پیشنهاد ادغام آن‌ها به کاربر استفاده می‌شود. همچنین، این پیمانه بر اساس اطلاعاتی که از پیش موجود است، به کاربر برچسب‌هایی برای عناصر هستان‌نگاری پیشنهاد می‌دهد.
- ◇ تحلیل فراوانی و میزان پوشش هستان‌نگارها: این پیمانه با استفاده از لایهٔ معنایی جدید حاصل از برچسب‌گذاری فارسی عناصر هستان‌نگارها، آن‌ها را از نظر میزان همپوشانی مفاهیم و خصوصیات با یکدیگر مقایسه می‌کند. بر اساس نتایج حاصل از این بررسی، تیم توسعه‌دهنده هستان‌نگارهایی را برای ادغام، انتخاب خواهد کرد تا از عناصر آن‌ها در هستان‌نگار نهایی استفاده شود.
- ◇ پالایش (ادغام و حذف) عناصر هستان‌نگاری: این پیمانه امکان تعریف پروژهٔ ادغام هستان‌نگارها را فراهم می‌آورد. در یک پروژهٔ ادغام، چند هستان‌نگار منتخب با استفاده از الگوریتم ادغام نیمه‌خودکار «پرامپت»[1] (Noy and Musen 2003) با هم ادغام می‌شوند. همچنین، این پیمانه امکان ادغام عناصر موجود

1. PROMPT

www.SID.ir

Archive of SID

در یک هستان‌نگار را هم فراهم می‌سازد. پیشنهادات ادغام ارائه‌شده توسط این پیمانه بر اساس مشابهت‌های ساختاری و معنایی نام اصلی عناصر (به زبان انگلیسی) و برچسب‌های معنایی ثبت‌شده (به زبان فارسی) است. استفاده از دو زبان در فرایند ادغام از ویژگی‌های منحصربه‌فرد «انتیرانداک» است.

- ◇ دریافت نظرات خبرگان (ارزیابی هستان‌نگار): برای ارزیابی، پیمانهٔ دریافت نظرات خبرگان امکان منحصربه‌فردی را جهت تبدیل هستان‌نگار به مجموعه گزاره‌های فارسی جهت پرسش از خبرگان به‌صورت الکترونیکی فراهم می‌کند. برای انجام این کار باید از قواعد نام‌گذاری مشخصی در زمان برچسب‌گذاری فارسی مفاهیم هستان‌نگار پیروی کرد: کلیهٔ مفاهیم با نام مفرد و روابط معنایی بین مفاهیم (خصوصیات شیء) به شکل مصدر فعل برچسب‌گذاری شوند. پس از پایان عملیات مفهوم‌سازی، مدیر پروژه خبرگانی را برای ارزیابی به هستان‌نگار منتسب می‌کند. این خبرگان می‌توانند از طریق آدرسی که در اختیار آن‌ها قرار گرفته با ثبت یک کد رمز وارد سامانه شده و به‌ترتیب، برای تمام مفاهیم، خصوصیات و روابط مدل‌سازی‌شده اظهار نظر کنند. اظهار نظر آن‌ها به شکل پذیرش، رد یا نداشتن نظر روی گزاره‌هایی خواهد بود که به زبان فارسی به آن‌ها نشان داده می‌شود. همچنین، خبرگان می‌توانند نظرات خود را در خصوص هر یک از عناصر هستان‌نگاری به‌صورت متنی در سامانه ثبت کنند. مدیر پروژه با مشاهدهٔ تجمیع نظرات، آن‌ها را برای اعمال در هستان‌نگار به افراد ذی‌ربط منتسب خواهد کرد.

۷. مدیریت منابع دانشی غیرساخت‌یافته: این زیرسامانه برای کمک به توسعه‌دهندگان برای ثبت منابع غیرساخت‌یافته و انجام عملیات مختلف استخراج واژگان و مدل‌سازی آن‌هاست. پیمانه‌های نرم‌افزاری این زیرسامانه عبارت‌اند از:

- ◇ ثبت مستندات و تجزیهٔ محتوا به صفحات: با استفاده از این پیمانه می‌توان فایل و اطلاعات تکمیلی منابع غیرساخت‌یافته و محتوای آن‌ها را به‌صورت صفحه‌بندی‌شده ذخیره‌سازی و پایش کرد. همچنین، این پیمانه با فراهم‌آوردن گزارش آماری از میزان واژگان استخراج‌شده از هر صفحه، امکان نظارت و کنترل بر فرایند استخراج واژگان را فراهم می‌آورد.

www.SID.ir

Archive of SID

- ◇ استخراج نیمه‌خودکار واژگان از متن: این پیمانه امکان شناسایی واژگان مهم هر صفحه را با استفاده از روش‌های زبان‌شناختی فراهم می‌آورد. فهرست واژگان پیشنهادی به کاربر نمایش داده می‌شود تا از بین آن‌ها موارد صحیح و مرتبط را انتخاب و به مخزن اضافه کند. در زمان استخراج واژگان، توسعه‌دهندگان باید شمارهٔ صفحه و پاراگرافی را که واژه در آن مورد ارجاع قرار گرفته است، ثبت کنند. با این کار در مرحلهٔ مفهوم‌سازی، امکان بررسی واژگان هم‌رخداد با تحلیل هم‌رخدادی امکان‌پذیر خواهد شد.
- ◇ شناسایی واژگان مشابه ساختاری و معنایی و ادغام آن‌ها: این پیمانه با کمک روش مشابهت‌یابی ساختاری (فاصله لونشتین)[1] و معنایی (استفاده از شبکهٔ واژگان فارسی)، به کاربران کمک می‌کند که واژگان مستخرج از متون مشابه را با هم ادغام کنند. هرگونه ادغام یا حذف واژگان در سوابق ثبت می‌شود.
- ◇ مفهوم‌سازی واژگان: با استفاده از این پیمانه، توسعه‌دهندگان واژگانی را که از متون استخراج شده، به‌صورت یک فهرست مرتبط بر اساس میزان اهمیت (با استفاده از روش وزن‌دهی فراوانی واژه و وزن‌دهی فراوانی واژه- معکوس فراوانی سند[2]) مشاهده کرده و با انتخاب هر واژه می‌توانند آن را به یکی از عناصر هستان‌نگاری مدل‌سازی کنند. این پیمانه برای مفهوم‌سازی، امکانات گوناگونی را فراهم می‌آورد. از جملهٔ آن‌ها می‌توان به موارد زیر اشاره کرد:
  - □ تحلیل معنایی واژه با استفاده از شبکهٔ واژگان فارسی «فردوس‌نت» و نمایش واژگان هم‌رده، عام‌تر و خاص‌تر (برای تشخیص و ساخت سلسله‌مراتب مفاهیم)
  - □ تحلیل ساختار نحوی (نمایش واژگان مشابه از نظر فاصلهٔ «لونشتین»، واژگان مشابه با پسوند یا پیشوند و همچنین، واژگان حاوی این واژه) برای تشخیص ارتباطات معنایی (خصوصیات شیء در هستان‌نگار)
  - □ نمایش عناصر مشابه موجود در مخزن هستان‌نگار (حاصل از انجام

1. Levenshtein distance
2. TF/IDF (Term Frequency/ Inverse Document Frequency)

www.SID.ir

Archive of SID

پروژه‌های ادغام) بر اساس برچسب‌گذاری فارسی انجام‌شده (توسعه‌دهنده از این طریق می‌تواند با دیدگاه سایر مدل‌سازان نسبت به آن واژه در هستان‌نگارهای موجود آشنا شده و انتخاب بهتری برای مفهوم‌سازی انجام دهد)

- مشاهدهٔ واژگان هم‌رخداد و فراوانی هم‌رخدادی (بر این اساس می‌توان روابط معنایی بین مفاهیم را مدل‌سازی کرد).

۸. مدیریت فرادادهٔ پایگاه داده و مهندسی معکوس: این زیرسامانه برای استخراج عناصر هستان‌نگاری از پایگاه داده به تیم توسعه کمک می‌کند و شامل پیمانه‌های نرم‌افزاری زیر است:

- ◇ انتقال طرحوارهٔ پایگاه داده به مخزن فرادادهٔ: این پیمانه با اتصال به موتور پایگاه داده، طرحوارهٔ آن را دریافت و به مخزن فرادادهٔ «انتیرانداک» اضافه می‌کند. چنانچه عناصر طرحواره (جداول و فیلدها) دارای شرح و توضیح باشند، به‌صورت پیش‌فرض به‌عنوان برچسب آن عنصر در فرادادهٔ ثبت می‌شوند.
- ◇ پیمانهٔ تکمیل دستی مستندات/ شناسایی اتصالات: روش پیشنهادی «انتیرانداک» برای مهندسی معکوس پایگاه داده بر پایهٔ ایجاد یک فرادادهٔ مناسب پایگاه داده عمل می‌کند. فرادادهٔ‌ای مناسب است که اجزای پایگاه داده (شامل جداول، فیلدها، شرح مقادیر برای فیلدهای شمارشی و اتصالات جداول خارج از طرحوارهٔ پایگاه داده) به‌صورت دستی یا نیمه‌خودکار مستند شده باشند.
- ◇ مهندسی معکوس: این پیمانه بر اساس قوانین نگاشت عناصر پایگاه داده رابطه‌ای به عناصر هستان‌نگاری، یک هستان‌نگار از روی فرادادهٔ ایجاد کرده و عناصر آن را به مخزن عناصر هستان‌نگاری اضافه می‌کند. الگوریتم مورد استفاده اقتباسی از الگوریتم (Zarembo 2015) است.

پیاده‌سازی و ارزیابی عملی:

از «انتیرانداک» برای اجرای دو پروژهٔ ساخت هستان‌نگار بومی (هستان‌نگار مرجع مؤسسات آموزشی و پژوهشی زیرمجموعهٔ «وزارت علوم، تحقیقات و فناوری»، به سفارش

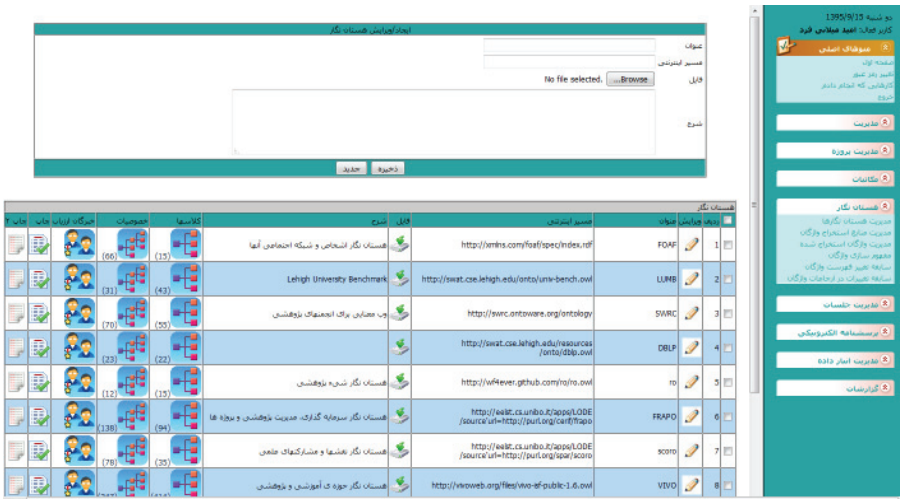
www.SID.ir

Archive of SID

«دفتر فناوری اطلاعات وزارت عتف»[1]، و هستان‌نگار پژوهش و فناوری استان «خراسان رضوی» به سفارش «منطقه ویژهٔ علم و فناوری رضوی»)، استفاده شده و بر اساس تجربیات حاصل بهبود یافته است. در ادامه، شرح مختصری از ایجاد هستان‌نگار پژوهش و فناوری استان «خراسان رضوی» به‌عنوان کاربرد عملی این روش آمده است. در شروع کار، تیم توسعه توسط مدیریت پروژه انتخاب گردید و ثبت و دسترسی‌های آن‌ها در پیمانهٔ مدیریت پروژه تنظیم شد.

مدیر پروژه در طول فرایند توسعه، با استفاده از پیمانهٔ مدیریت پروژه، عملیات مختلف مدیریتی از جمله تخصیص کار به اعضا، مدیریت مستندات، نظارت بر انجام کارها و اقدامات و اخذ گزارشات مختلف زمانی و مالی را انجام می‌داد. بخش مکاتبات الکترونیکی و مدیریت جلسات نیز به‌عنوان ابزار حمایتی از روند کلی انجام کار توسط اعضا مورد استفاده قرار می‌گرفتند.

پس از انجام فعالیت‌های شناسایی منابع دانش، هستان‌نگارهای یافت‌شده (۱۰ مورد) در بخش مدیریت هستان‌نگارها (شکل ۳) و مستندات مرتبط (۲۸ مورد) نیز در بخش مدیریت منابع استخراج واژگان وارد شدند.



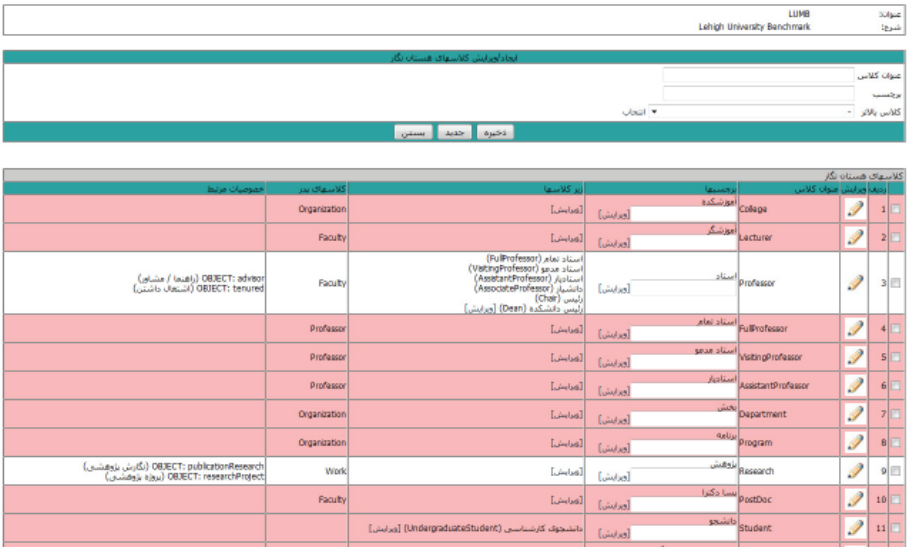
شکل ۳. مدیریت هستان‌نگارهای موجود

۱. علوم، تحقیقات و فناوری

www.SID.ir

Archive of SID

عناصر هستان‌نگارهای یافت‌شده (کلاس‌ها و خصوصیات)، با استفاده از امکانات شرح داده‌شده در ۷-۲ برچسب‌گذاری فارسی و ادغام شدند (شکل ۴). بر اساس عناوین عناصر و برچسب‌های آن‌ها از نظر میزان همپوشانی مفاهیم و روابط تحلیل آماری انجام شده و در نهایت، ۵ هستان‌نگار برای ادغام، انتخاب و با کمک ابزار ادغام شدند که حاصل آن هستان‌نگاری با ۱۰۲ کلاس و ۲۷۹ خصوصیت داده و شیء بود.

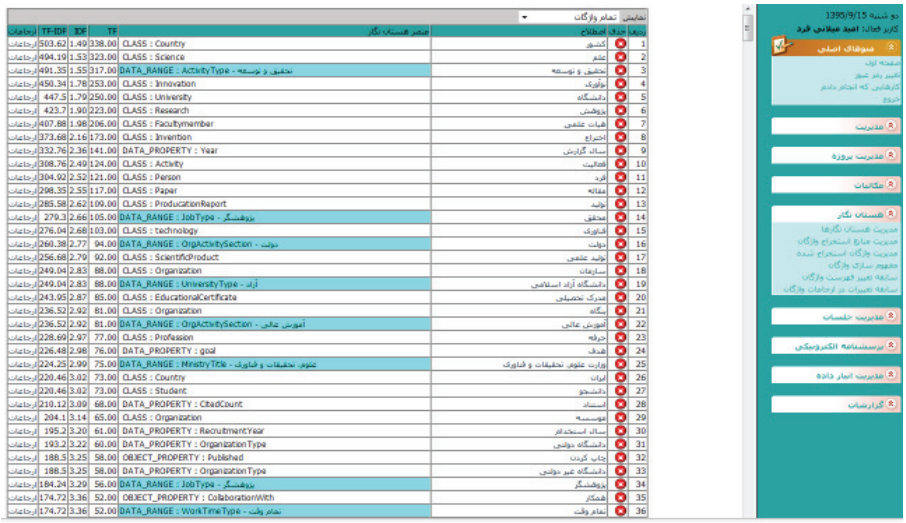


شکل ٤. برچسب‌گذاری عناصر هستان‌نگار

به‌صورت موازی، کار استخراج واژگان از منابع غیرساخت‌یافته توسط اعضای تیم صورت گرفت و در نتیجه، ۶۶۲ واژه به‌دست آمد. این ابزار در فرایند استخراج واژگان، حمایت هایی چون پیشنهاد واژه‌های مشابه ساختاری و معنایی، امکان ادغام واژگان مشابه در حین کار و نمایش عنصر مدل‌سازی‌شده در هستان‌نگار برای اعضای تیم فراهم می‌کرد. پس از استخراج واژگان، عملیات نگاشت هر واژه به عنصر هستان‌نگاری معادل، با کمک ابزاری که تحلیل‌های مختلف مشابهت‌یابی ساختاری و معنایی، استفاده از عناصر هستان‌نگاری موجود و همچنین، تحلیل هم‌رخدادی انجام می‌داد، صورت گرفت (شکل ۵).

www.SID.ir

Archive of SID



شکل ۵. ایجاد یا نگاشت عنصر هستان‌نگاری معادل هر واژه

پس از تکمیل عملیات ایجاد هستان‌نگار، خبرگانی برای ارزیابی تعیین شدند. این خبرگان جهت انجام عملیات ارزیابی حساب کاربری دریافت کردند و با ورود به سامانهٔ ارزیابی الکترونیکی، نظرات خود را در خصوص هر یک از عناصر هستان‌نگار ثبت کردند. صفحهٔ ارزیابی هر عنصر هستان‌نگار شامل توضیح کاملی از هر مفهوم، خصوصیات و روابط آن با سایر مفاهیم و همچنین، خصوصیات و روابط غیرمستقیم (به‌واسطهٔ روابط سلسله‌مراتبی) است. بر اساس نظرات دریافتی از خبرگان، تیم توسعه اصلاحات مورد نیاز را در مفاهیم و روابط انجام داد. عمل ارزیابی مجدداً تکرار شد و اصلاحات نهایی صورت گرفت. هستان‌نگار حاصل، ۱۱۱ کلاس و ۲۱۹ خصوصیت دارد و در ایجاد انبارهٔ داده منطقهٔ ویژهٔ علم و فناوری «خراسان رضوی» مورد استفاده قرار گرفته است.

به‌دلیل پیدا نشدن پایگاه دادهٔ عملیاتی مناسب در این حوزه، از بخش مهندسی معکوس در پروژهٔ ساخت هستان‌نگار منطقهٔ ویژهٔ علم و فناوری «خراسان رضوی» استفاده نشده است. در پروژهٔ ساخت هستان‌نگار مرجع مؤسسات آموزشی و پژوهشی زیرمجموعهٔ وزارت «عتف»، پایگاه داده عملیاتی «سدف»[1] به‌عنوان منبع دانشی ساخت‌یافته مورد استفاده قرار گرفته است و این بخش از روش و ابزار «انتیرانداک» آزمون شده است. هستان‌نگار مرجع

۱. سامانه‌های دانشگاه فردوسی مشهد (سدف) که به‌صورت متن باز طراحی و پیاده‌سازی شده‌اند، در بیش از ۳۰ دانشگاه و مرکز آموزش عالی در سراسر کشور نصب شده و در حال بهره‌برداری است.

www.SID.ir

Archive of SID

یادشده در مرحلهٔ ارزیابی نهایی است.

نتیجه‌گیری

روش «انتیراناداک»، فرایند توسعهٔ ساخت هستان‌نگار را به‌صورت مجموعه‌ای از فعالیت‌هایی که باید انجام شود، به همراه شرح جزئیات انجام کار و همچنین، ابزاری یکپارچه برای حمایت از اجرای فعالیت‌ها ارائه کرده است. ارتباط بین فعالیت‌های شناسایی‌شده در چارچوب مقایسهٔ پیشنهادی (جدول ۱) و فعالیت‌های شناسایی‌شدهٔ «انتیراناداک» به‌صورت خلاصه به‌شرح زیر است:

- ◇ فعالیت‌های مدیریتی: فعالیت مدیریت پروژه؛
- ◇ فعالیت‌های پیش از توسعه: شناسایی منابع دانش؛
- ◇ فعالیت‌های توسعه: فعالیت‌های توصیف نیازمندی‌ها، استخراج دانش خبرگان، استخراج واژگان، مهندسی معکوس، ادغام هستان‌نگارها، مفهوم‌سازی؛
- ◇ فعالیت‌های پس از توسعه: ارتقا و نگهداری؛
- ◇ فعالیت‌های پشتیبانی: شناسایی خبرگان، شناسایی هستان‌نگارهای مرتبط، شناسایی پایگاه داده‌های مرتبط، شناسایی مستندات غیرساخت‌یافته، ارزیابی خبرگان، مدیریت پروژه، ادغام هستان‌نگارها.

در خصوص معیار حمایت از زبان دیگر، «انتیراناداک» با برچسب‌گذاری فارسی عناصر هستان‌نگار، امکان تحلیل‌های معنایی و ساختاری برای ادغام عناصر هستان‌نگاری را فراهم می‌کند. همچنین، در استخراج واژگان، ابزاری برای پالایش و ایجاد سازگاری در فهرست واژگان بر اساس مشابهت‌یابی ساختاری و معنایی فارسی دارد. مزیت قابل توجه دیگر «انتیراناداک»، تولید فرم‌های ارزیابی از خبرگان بر اساس گزاره‌های فارسی مبتنی بر برچسب‌های فارسی عناصر است.

در نهایت، روش «انتیراناداک» بر اساس تجربیات عملی ساخت هستان‌نگار، جزئیات کافی برای انجام فعالیت‌ها و چگونگی عملکرد پیمانه‌های نرم‌افزاری را دارد و ابزار جامعی نیز برای انجام فرایند ساخت ارائه می‌کند.

«انتیراناداک» به‌عنوان یک روش جدید توسعهٔ هستان‌نگار با بهره‌گیری از ایده‌های روش‌های موجود و افزودن ویژگی‌های جدید (مانند پشتیبانی از زبان فارسی، پیمان‌های

www.SID.ir

Archive of SID

بودن معماری، توجه به توسعهٔ مشارکتی و یکپارچگی داده‌ها)، طراحی شده و بر اساس تجربیات حاصل از اجرای دو پروژهٔ ساخت هستان‌نگار بومی توسعه یافته و در عمل کارایی خود را نشان داده است. پیمانه‌های نرم‌افزاری «انتیرانداک» با استفاده از زبان php و پایگاه داده MySQL به‌صورت تحت وب پیاده‌سازی شده است. این سامانه به‌صورت متن باز است و طراحی پیمان‌های آن این قابلیت را فراهم آورده است که سایر پژوهشگران و علاقه‌مندان، بخش‌های مورد نظر خود را در آن اضافه کرده و الگوریتم ها و ویژگی‌های آن را شخصی‌سازی کنند.

فهرست منابع

Blázquez, J., M. Fernández, J. M. García-Pinar, and A. Gómez-Pérez. 1998. Building ontologies at the knowledge level using the ontology design environment. *Proceedings of the 11th Banff Knowledge Acquisition for Knowledge-Based Systems.* Banff, Alberta, Canadá (pp. 4.1-4.15)

Brost, W. N. 1997. *Construction of engineering ontologies*. University of Twenty, Center for Telematica and Information Technology: Enschede

Corcho, O., M. Fernández-López, A. Gómez-Péreza and A. López-Cima. 2005. *Building legal ontologies with METHONTOLOGY and WebODE: 142-157*. Berlin, Heidelberg: Springer.

De Nicola, A., M. Missikoff and R. Navigli. 2009. A software engineering approach to ontology building. *Information systems* 34 (2): 258-275.

Fernández, M, A Gómez-Pérez and N. Juristo. 1997. *METHONTOLOGY: From Ontological Art towards Ontological Engineering*. AAAI Spring Symp. Series. Menlo Park, Calif: AAAI Press: 33-40.

Fernández-López, M. And A. Gómez-Pérez. 2002. Overview and analysis of methodologies for building ontologies. *The Knowledge Engineering Review* 17 (02): 129-156.

Gailly, F .and G. Poels. 2010. Conceptual modeling using domain ontologies: Improving the domain-specific quality of conceptual schemas. *Presented at the 10th Workshop on Domain-Specific Modeling*. DSM '10. New York, USA: ACM.

Gómez-Pérez, A., M. Fernández-López and O. Corcho. 2004. Methodologies and methods for building ontologies. Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web: 107-197. Berlin, Heidelberg: Springer

Gruber, T. R. 1993. A translation approach to portable ontology specifications. *Knowledge acquisition* 5.199-220 :(2)

1995 ._____. Toward principles for the design of ontologies used for knowledge sharing. *International journal of human-computer studies* 43 (5): 907-928.

Grüninger, M. And M. S. Fox. 1995. Methodology for the Design and Evaluation of Ontologies. *Presented at the Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95*. Montreal.

Guarino, N., D. Oberle and S. Staab. 2009. What is an Ontology? In *Handbook on ontologies*. Springer: Berlin Heidelberg. 1-17.

Lenat, D. B., R. V. Guha, K. Pittman, D. Pratt and M. Shepherd. 1990. Cyc: toward programs with common sense. *Communications of the ACM* 33 (8): 30-49.

Noy, N. F and D. L. McGuinness. 2001. *Ontology development 101: A guide to creating your first ontology*. Palo Alto: Stanford Medical Informatics.

www.SID.ir

Archive of SID

Noy, N. F and M. A. Musen. 2003. The PROMPT suite: interactive tools for ontology merging and mapping. *International Journal of Human-Computer Studies* 59 (6): 983-1024.

Noy, N. F., R. W. Fergerson and M. A. Musen. 2000. The knowledge model of Protege-2000: Combining interoperability and flexibility. In *International Conference on Knowledge Engineering and Knowledge Management*. Berlin Heidelberg: Springer. (pp 17-32).

Paolucci, M., T. Kawamura, T. R. Payne and K. Sycara. 2002. Semantic matching of web services capabilities. In *The Semantic Web—ISWC 2002*. Berlin Heidelberg: Springer. (pp 333-347).

Pinto, H. S., C. Tempich and S. Staab. 2009. Ontology engineering and evolution in a distributed world using DILIGENT. Handbook on ontologies: 153-176. Berlin Heidelberg: Springer.

Studer, R., V. R. Benjamins and D. Fensel. 1998. Knowledge engineering: principles and methods. *Data & knowledge engineering* 25 (1): 161-197.

Suarez-Figueroa, M. C. 2010. NeOn methodology for building ontology networks: Specification, scheduling and reuse (Doctoral dissertation). Madrid: Univ. Politécnica.

Sure, Y. and R. Studer. 2002. On-To-Knowledge methodology. Technical Report. Karlsruhe, Germany: Institute AIFB, University of Karlsruhe.

Sure, Y., M. Erdmann, J. Angele, S. Staab, R. Studer and D. Wenke. 2002. OntoEdit: Collaborative ontology development for the semantic web. In *International Semantic Web Conference*. Berlin Heidelberg: Springer. (pp 221-235).

Uschold, M. and M. Gruninger. 1996. Ontologies: Principles, methods and applications. *The knowledge Engineering Review* 11 (02): 93-136.

Uschold, M. and M. King. 1995. towards a methodology for building ontologies. *Workshop Held in Conjunction with IJCAI on Basic Ontological Issues in Knowledge Sharing*, Montreal.

Zarembo, I. 2015. Automatic Transformation of Relational Database Schema into OWL Ontologies. In *Environment. Technology. Resources. Proceedings of the International Scientific and Practical Conference* (Vol. 3) (pp 217-222), Rezekne, Latvia.

امید میلانی‌فرد

متولد سال ۱۳۵۶، دارای مدرک کارشناسی مهندسی نرم‌افزار، کارشناسی ارشد مدیریت اجرایی و در حال گذراندن دورهٔ دکتری مدیریت فناوری اطلاعات است. وی در حال حاضر کارشناس ارشد فناوری اطلاعات «دانشگاه فردوسی مشهد» است. سامانهٔ اطلاعات مدیریت، انبارهٔ داده، مدل‌سازی و مهندسی داده علاقه‌مندی‌های پژوهشی وی است.

محسن کاهانی

متولد سال ۱۳۴۶، دارای مدرک دکتری مهندسی کامپیوتر از «دانشگاه ولونگونگ استرالیا»ست. وی در حال حاضر استاد تمام گروه مهندسی کامپیوتر «دانشگاه فردوسی مشهد» و مدیر آزمایشگاه‌های تخصصی «آپا» و فناوری وب است. وب معنایی و فرایندکاوی از علاقه‌مندی‌های پژوهشی وی است.

www.SID.ir