

بازیابی اطلاعات چندزبانه از پژوهش تا عمل

کارول پیترز
مارتین بارشلر
پل کلاف

ترجمه:
بهمن ریل
امیرضا اصنافی



۱۳۹۶

سرشناسه	: پیترز، کارول Carol, Peters
عنوان و نام پدیدآور	: بازیابی اطلاعات چندزبانه از پژوهش تا عمل / کارول پیترز، مارتین بارشلر، پل کلاف؛ ترجمه بهروز یل، امیررضا اصنافی.
مشخصات نشر	: تهران: اساطیر پارسی: چاپار، ۱۳۹۶.
مشخصات ظاهری	: ۳۶۲ص.
شابک	: 978-600-97299-2-0
وضعیت فهرست نویسی	: فیبا
یادداشت	: عنوان اصلی:
یادداشت	: کتابنامه.
موضوع	: بازیابی اطلاعات بین‌زبانی
موضوع	: Cross-language information retrieval
شناسه افزوده	: براشلر، مارتین Martin, Brashler
شناسه افزوده	: کلاق، پاول (Paul) D Clough, Paul (Paul)
شناسه افزوده	: بهروز، ۱۳۶۲ - مترجم
شناسه افزوده	: اصنافی، امیررضا، ۱۳۵۹ - مترجم
رده بندی کنگره	: Z5/667/... ۲۰۲۶
رده بندی دیویی	: ۰۰۲۵
شماره کتابشناسی ملی	: ۴۶۴/۵۰۲



تهران: ۸۸۸۹۹۶۸۰ - ۰۱۱۱۹۰۰۲۱۴

تهران- خیابان ولی عصر (عج) - بالاتر از رشت
کوچه نوربخش - پلاک ۴۰ - نشر چاپار

بازیابی اطلاعات چندزبانه از پژوهش تا عمل

کارول پیترز، مارتین بارشلر، پل کلاف

ترجمه: بهروز یل، امیررضا اصنافی

صفحه‌آرا و طراح جلد: رحیم کبیرصابر

ناظر فنی چاپ: مریم تفضلی

چاپ اول ۱۳۹۶

شمارگان: ۵۰۰ نسخه

شابک: ۹۷۸-۶۰۰-۹۷۲۹۹-۲-۰ ISBN: 978-600-97299-2-0

حق هرگونه چاپ و انتشار محفوظ است.

قیمت: ۲۶۰۰۰ تومان

پیشگفتار

این کتاب توسط کارول پیترز^۱، مارتین براسکلر^۲ و پل کلاف^۳ طی تقریباً یک دهه پژوهش عمیق در مورد دسترسی به اطلاعات چندزبانه نوشته شد، همچنین این کتاب ترکیب مناسب مجموعه‌ای از دستاوردها است و به دستورالعمل‌های جدید برای پژوهش در سازه‌های آتی اشاره می‌کند. همان‌طور که مشخص است، ما در جهان چندزبانه زندگی می‌سیم و این جنبه‌ای است که باید به‌عنوان یک اصل، هم اینک و در آینده باقی بماند. حتی اگر صد سال از زبان طبیعی در طول دهه‌های بعد ناپدید شوند و منجر به این شود که مجموعه که از زبان‌ها توسط تعداد زیادی از مردم صحبت شوند، بر اساس گزارش وب‌سایت اتنولوگ^۴ ۸۵ زبان توسط حدود ۳۹ درصد از انسان‌ها استفاده می‌شود، در حالی که ۸۵ زبان توسط ۷۸ درصد از انسان‌ها استفاده می‌شود. در حال حاضر رایج‌ترین زبان‌ها، زبان چینی، اسپانیایی و انگلیسی هستند، اما حتی اگر دومین زبان رایج را «زبانی انگلیسی» قلمداد کنیم، گونه‌های زبانی زیادی که در منهن یا در حومه شهر نیویورک و همچنین در روزنامه‌ها مانند تایمز یا در پست‌های وبلاگ پیدا می‌شود. چارچوب گونه‌های زبانی پیش‌فرض در کاربرد کاملاً مهم است و بنابراین باید طراحان سامانه رایانه‌ای، تغییراتی برای دسترسی به گونه‌های زبانی به عهده بگیرند.

این کتاب همچنین درباره پوشش‌های ارزیابی است که توسط سخنگاه ارزیابی برحسب زبان متقابل با تمرکز روی بازایی چندزبانه در کشورهای اروپایی

پایه‌سازی شده‌اند. مشکلات و چالش‌های خاص مربوط به دسترسی به اطلاعات چندزبانه اهمیت ویژه‌ای برای قاره کهن (اروپا) دارند، اروپایی‌ها مجبورند که اطلاعات تهیه‌شده در انواع زبان‌ها را در هنگام پایه‌سازی وظایفی چون جستجو برای آمارهای ملی که به زبان خارجی نوشته‌شده، رزرو هتل یا حتی هزینه تعطیلات در کنار دریا را کنترل کنند. برعکس در کشورهای شرق دور، مشکلات ارتباطات معمولاً به زبان انگلیسی و زبان‌های دیگر (به‌عنوان مثال زبان چینی، ژاپنی یا کره‌ای) محدود می‌شود. و اروپایی‌ها در یک محیط چندزبانه زندگی می‌کنند هنوز اکثر اروپایی‌ها دانش سطحی از یک یا دو زبان دیگر را دارند. در این بافتار، در طول ۱۰ سال اول، پویش‌های سخگه ارزیابی چندزبانه متفاوت، مجموعه‌ای از پیکره مرتبط و ابزارهایی که برای دستکاری برخی زبان‌ها استفاده می‌شوند را ایجاد کرده است. از این تلاش بین‌المللی سپاسگاری می‌نمایم، اکثر ابزارهای باکیفیت که در حال حاضر قابل دسترس هستند، برای دستیابی به اطلاعات در گروه‌های زبانی متعدد از جمله گروه‌های هم‌خانواده آلمانی (ژرمنی) (زبان آلمانی، انگلیسی، هلندی و سوئدی)، رومیایی (زبان‌های اسپانیایی، پرتغالی، فرانسوی و ایتالیایی)، اسلاوی (زبان‌های روسی، بلغاری و چک) و همچنین اعضاء خانواده اوستایی (فنلاندی، مجارستانی) نوشته شده‌اند.

همان‌طور که در این کتاب توضیح داده شده است، تنوع زبان، انواع گسترده‌ای از چالش‌هایی را که توسط دانشمندان علم رایانه حل شده بود را ایجاد می‌کند. در زبان آلمانی، برای مثال، ساختارهای مرکب خیلی متعدد هستند اما نگرانی واقعی این است که مفهوم مشابه ممکن است توسط دو (یا بیشتر) فرمول‌بندی بیان شود و در نتیجه ارائه آن را برای یافتن انطباق مفید بین کلیدواژه‌های جستجو و مقوله‌های هدف مشکل‌تر می‌سازد. موارد دستوری متعدد، در دستور زبان فنلاندی پیدا شد که مثال دیگر را ارائه می‌دهد. در اینجا مشکل واقعی بی‌نظمی‌های اعمال‌شده توسط هارمونی‌های مصوت را در برمی‌گیرد که به نوبت ملزومات پردازش ریخت‌شناسی

پیچیده را ایجاد می‌کند. پژوهش بیشتر برای ترویج پردازش خودکار بهتر مانند ساختار زبانی نیاز است.

علاوه بر طراحی روش‌های پردازش مؤثر و کارآمد برای زبان‌های طبیعی مختلف، دسترسی به اطلاعات چندزبانه باید استفاده از بعضی قالب‌های ترجمه خودکار را ایجاد کند. برای عبور از این موانع زبانی، راهبردهای جستجوی مختلف، پیشنهاد، پیاده‌سازی و ارزیابی شده است. هنگام آماده‌سازی زبان‌های مرتبط، به‌عنوان مثال مانند زبان فرانسه و انگلیسی، ما تنها یک زبان را به‌عنوان قالب غلط نوشته‌ایده توسط دیگران در نظر می‌گیریم، به‌طوری‌که ترجمه به‌سادگی به‌عنوان مشکل تصحیح املا تلقی می‌شود. روش‌های دیگر و مؤثرتر دوره آموزشی توسعه یافته است و اغلب زبانی فرهنگ‌های لغت دوزبانه ماشین‌خوان^۱، سامانه‌های ترجمه ماشینی خودکار^۲ و همچنین ترجمه‌های آماری مبتنی بر پیکره موازی و قابل مقایسه هستند. در این نه‌هوا، وجود این، اکثر مشکلات هنوز حل نشده است، به‌خصوص آن‌هایی که زبان‌های گفتاری زیرمتداول را در برمی‌گیرند.

دسترسی به اطلاعات چندزبانه نه فقط به اسناد متنی محدود نمی‌شود بلکه همچنین توسط رسانه‌های دیگر شامل صوت (مانند مصاحبه‌ها)، موسیقی، تصاویر، عکس‌ها، ویدئوها، فیلم‌ها و غیره قابل اجرا هستند. هر یک از این کتاب نشان می‌دهد، میراث فرهنگی از منابع اطلاعاتی مفید تشکیل می‌شود. راه‌نانی که دسترسی به اطلاعات چندزبانه را جستجو می‌کنند. برای مثال آثار هنری مشهور معمولاً تحت نام‌های مختلف بسته به زبان اصلی شناسایی می‌شوند (Mona Lisa, Monna Lisa, Mona Liza, La Joconde or La Gioconda).

از اعتبار راه‌حل‌های فناورانه پیشرفته کاسته خواهد شد اگر کاربران آن‌ها را نپذیرند. در این دیدگاه، این کتاب مطالعات مناسب و نتایج قابل اجرا برای طراحی رابط‌های کاربری چندزبانه برای رسانه‌های مختلف را پوشش می‌دهد. چنین

جنبه‌هایی در هنگام طراحی نرم‌افزارهای چندزبانه برای وب‌سایت‌ها، تلفن‌های همراه و سایر دستگاه‌های فناورانه، مهم هستند. به‌طور خلاصه، این کتاب جذاب، دامنه‌ای از مشکلات ارتباطی انسان مربوط به رسانه و فناوری را پوشش می‌دهد، در تلاش‌هایشان، تسهیل شناخت و ارتباط بین انسان‌ها را به عهده می‌گیرند، گفتگوی بهتر بین فرهنگ‌ها را تشویق می‌کنند و پل‌هایی بین چالش‌های میراث فرهنگی گذشته، حال و آینده ایجاد می‌کنند.

ژاک ساوی^۱

دانشکده علم رایانه

دانشگاه نیوچاتل^۲

سونیس، ژوئن ۲۰۱۱

مقدمه

این کتاب، توصیف جامعی از فناوری درگیر در طراحی و توسعه سامانه‌ها و کارآمدی دسترس و بازیابی اطلاعات چندزبانه را ارائه می‌دهد. هدف ما این است که خواننده شناخت کاملی را از مسائل مختلف مرتبط با ایجاد سامانه‌ها برای در دسترس قرار دادن مقوله‌های اطلاعاتی ذخیره‌سازی شده دیجیتالی - صرف‌نظر از زبانی که با آن نوشته شده‌اند- پیدا کند. کتاب معطوف به دانشجویان فارغ‌التحصیل و پزشکان با شناخت اصلی از روش‌های بارابی متن کلاسیک می‌شود. میزان فزاینده اطلاعات غیر از زبان انگلیسی که در سطح جهان قابل دسترس است و عرضه جهان‌گسترده فزاینده [محصولات] شرکت‌ها در سطح جهان، تردیدی باقی نمی‌گذارد که این گروه‌های هدف باید دانش موجودشان را، را جمع، روش‌های بازیابی اطلاعات با محیط‌های چندزبانه جدید سازگار کنند. هدف ما این است که شکاف بین منابع تحت پوشش درسامه‌های بازیابی اطلاعات که برای متعدد و حقایق عینی عملکردی جدید را از بین ببریم.

کتاب به شش فصل تقسیم می‌شود. هدف این است که خواننده را گام‌به‌گام از طریق مرحله‌های مختلف درگیر در ایجاد، به‌کارگیری و ارزیابی سامانه‌های بازیابی اطلاعات چندزبانه^۱ همراه کند و با مثال‌هایی از نرم‌افزارهای اخیر فناوری بازیابی اطلاعات به پایان می‌رسد. برخی از روش‌هایی که ما توصیف می‌کنیم به‌تازگی در سامانه‌های جستجوی تجاری پدیدار شده‌اند، درحالی که روش‌های دیگر این قابلیت را دارند که بخشی از تجسم‌های بعدی باشند.

هدف فصل اول این است که دلایلی را ارائه کند که چرا دسترسی مؤثر و کارآمد به اطلاعات در سراسر مرزهای زبانی در جامعه جهانی دیجیتال ما، به تعیین‌کننده‌ترین عامل تبدیل شده است. مفاهیم اصلی بیان و مورد بحث قرار می‌گیرد و اصطلاحات اصلی مورد استفاده در کتاب، تعریف می‌شوند. مقدمه، همچنین پیشینه مختصری از فعالیت علمی و تجاری در این بخش را ارائه می‌دهد و چالش‌های اصلی که امروز با آن‌ها مواجه هستند را بیان می‌کند.

همان‌طور که در مقدمه بیان شد، توسعه سامانه بازیابی اطلاعات چندزبانه، درباره مدیریت دسترسی به اطلاعات، کشف در زبان‌های متعدد هم تک‌زبانه و هم در سراسر زبان‌ها است، در حالی که بازیابی اطلاعات برحسب زبان‌های متقابل^۱ به‌طور خاص به آن فناوری‌های نو اشاره می‌کند که راجع به پرسوجوی مجموعه چندزبانه در یک زبان برای بازیابی اسناد، مرتبط در زبان‌های دیگر بودند. بازیابی اطلاعات تک‌زبانه^۲ کارآمد یکی از پیش‌شرط‌های یاده‌سازی سامانه‌های بازیابی اطلاعات چندزبانه/بازیابی اطلاعات برحسب زبان‌ها، متقابل در نظر گرفته شده است. بنابراین فصل ۲ سازوکارهای ضروری برای ایجاد سامانه‌های بازیابی اطلاعات تک‌زبانه را ارائه می‌دهد و به عبارت دیگر راهبردهای نمایه‌سازی و طبقه‌بندی و جزئیاتی را در مورد انطباق این سازوکارها به زبان‌های مختلف را ارائه می‌کند.

لذا فصل ۳ راهبردهای مختلفی را که برای حمایت بازیابی اطلاعات برحسب زبان‌های متقابل استفاده می‌شود را ارائه می‌دهد و ادغام انواع مؤلفه‌های ترجمه را در سامانه بازیابی اطلاعات توصیف می‌کند. مشکلات پدید آمده توسط ابهام پرسش و/یا ابهام ترجمه، دامنه واژگانی ناکافی و کیفیت منابع قابل دسترسی برای زبان‌های مختلف، بحث می‌شوند و پیشنهادهایی برای تمرین‌های بهتر ارائه می‌شوند.

در سامانه‌های بازیابی اطلاعات چندزبانه/بازیابی اطلاعات برحسب زبان‌های متقابل، طراحی رابط کاربری جستجوی مؤثر و تهیه کارآمدی برای کمک به کاربر در

انجام جستجو، ضروری هستند، زیرا کاربران از مرزهای زبانی عبور می‌کنند و با منابع در زبان‌های ناآشنا ارتباط برقرار می‌کنند. اثرات مهارت‌های زبانی کاربران و پس‌زمینه فرهنگی بر طراحی رابط کاربری تأثیر می‌گذارند. به این دلایل، در فصل ۴ تمرکز روی رفتار جستجوی اطلاعات کاربر است، بافتارهایی که در آن کاربران اطلاعات و نیازها و توانایی‌های شناختی‌شان را جستجو می‌کنند. دستورالعمل‌هایی تهیه می‌شوند در مورد اینکه چگونه رابط‌های کاربری جستجوی چندزبانه را طراحی کنند؟ تعالی کاربر سامانه را به عنوان مثال در فرمول‌بندی و ترجمه پرسش، در انتخاب و ارزیابی سند، در جستجو و مصورسازی پشتیبانی کند.

فصل ارزیابی سامانه‌ها را برای بازایی اطلاعات چندزبانه از دیدگاه سامانه‌مدار و کاربر، ارزیابی می‌کند و برای خوانندگان اطلاعاتی را درباره اینکه چگونه ارزیابی فعالیت خودشان را انجام دهند، فراهم می‌کند و نکاتی که باید قبل از شروع مورد توجه قرار دهند را فهرست می‌کند. این فصل همچنین در مورد اهمیت نقش ارزیابی در پیشبرد بررسی و توسعه وجود و تهیه سامانه‌هایی که نتایج بهتری را ارائه می‌دهند، بحث می‌کند؛ این اقدام نه فقط برحسب عملکرد در بازایی اطلاعات مرتبط بلکه همچنین از لحاظ کسب رصداً نظارات کاربر موفق بوده‌اند، انجام می‌شود.

بازایی اطلاعات از این به بعد تنها درباره بازایی سند و سند نیست؛ امروزه محتوا به طرز چشمگیری چندرسانه‌ای است و علایق و نیازهای کاربران به سرعت در حال تغییر هستند. به همین دلیل، در فصل آخر، ما در مورد روش‌هایی بحث می‌کنیم که در آن روش‌ها فناوری‌های پیشرفته برای پیاده‌سازی سامانه‌ها جهت بازایی اطلاعات چندزبانه و برحسب زبان‌های متقابل در حوزه‌هایی اتخاذ شده‌اند که جستجوی سند متنی مانند بازایی چندرسانه‌ای و استخراج اطلاعات از مرزهای زبانی یا فراتر گذاشته‌اند. همچنین ما دامنه‌ای از حوزه‌های نرم‌افزارهای عملی را که

این فناوری‌ها را استفاده کرده‌اند، توصیف می‌کنیم و بعضی از چالش‌های برآمده از تحول پیش‌نمون پژوهش سامانه‌های عملیاتی دنیای واقعی را ارائه می‌کنیم.

موضوع مورد پوشش، گسترده و پیچیده است و سعی ما بر این است که بررسی جامعی را تهیه کنیم. هدف ما تهیه دستورالعمل‌ها و اطلاعاتی در مورد همه جنبه‌هایی است که باید در هنگام ایجاد سامانه‌های بازیابی اطلاعات چندزبانه/بازیابی اطلاعات برحسب زبان متقابل در نظر گرفته شود، درحالی‌که بیش از حد از 'جزئیات عملی' که به سرعت کم می‌شوند، همانند نرم‌افزاری که پیشرفت دائمی دارد، اجتناب می‌کنیم.

هر فصل با فهرستی از منابع مرجعی که با دقت گردآوری شده، مجهز می‌شود و فصل‌های ۲-۵، جزئیات اصلی در مورد توسعه و ارزیابی سامانه‌های بازیابی اطلاعات چندزبانه/بازیابی اطلاعات برحسب زبان‌های متقابل را تهیه می‌کند و همچنین هر فصل با سیاه‌پشتی از منابع خواندنی پیشنهادی به پایان می‌رسد تا در صورت نیاز اطلاعات دقیق‌تر را پیدا کنند. امیدواریم که این کتاب مطالب خواندنی جذابی را ارزیابی کند و منبع اطلاعاتی مفیدی برای دانشجویان، دانشمندان، توسعه‌دهندگان سامانه و همه کسانی که به این روش علاقه دارند، تهیه کند تا اینکه نیازهای آن‌ها برای یادگیری و اشاعه اطلاعات را برآورده کند، با این حال، هر جا و در هر زبانی، ذخیره شد.

نویسنده: پائول
معاونین: پائول
پل کلاف

فهرست مطالب

۲۳	فصل ۱: مقدمه.....
۲۳	۱.۱. اهداف و اهداف.....
۲۴	۱.۱.۱. شناسایی نیازها.....
۲۹	۲.۱. مرحله اول: اطلاعات.....
۳۰	۳.۱. پیشینه مقاله.....
۳۱	۱.۳.۱. فناوری‌ها و استانداردهای فعالسازی.....
۳۲	۲.۳.۱. پروژه‌های پژوهش با بودجه عمومی.....
۳۶	۳.۳.۱. کنفرانس‌ها و پویس‌های رزبان.....
۴۰	۴.۳.۱. محصولات تجاری.....
۴۴	۴.۱. چالش‌های پژوهش فعلی.....
۴۹	فصل ۲: بازیابی اطلاعات تک‌زبانی.....
۵۰	۱.۲. مقدمه.....
۵۴	۲.۲. مشکلات بازیابی اطلاعات و پیامدهای آن.....
۵۷	۳.۲. پیاده‌سازی سامانه بازیابی اطلاعات تک‌زبانه.....
۶۰	۴.۲. مرحله نمایه‌سازی.....
۶۱	۱.۴.۲. پیش‌پردازش (مرحله ۱).....
۶۴	۲.۴.۲. شناسایی زبان (مرحله ۲).....
۶۵	۳.۴.۲. شکلگیری سند (مرحله ۳).....
۶۶	۴.۴.۲. قطعه‌بندی، نشانه‌گذاری، تجزیه (مرحله ۴).....
۷۲	۵.۴.۲. نرمال‌سازی مشخصه (مرحله ۵).....
۷۹	۶.۴.۲. گسترش (مرحله ۶).....

۸۰	 ۵.۲ مرحله انطباق
۸۱	 ۱.۵.۲ مدل 'بسته‌ای از کلمات'
۸۳	 ۲.۵.۲ نمایه معکوس
۸۶	 ۳.۵.۲ الگوریتم انطباق اصلی
۸۸	 ۴.۵.۲ مدل فضای برداری
۹۱	 ۵.۵.۲ طرح وزن دهی کسینوس فراوانی اصطلاح- فراوانی سند معکوس
۹۶	 ۶.۵.۲ بازخورد ربط
۹۹	 ۷.۵.۲ ارجح‌های وزن دهی احتمالاتی
۱۰۲	 ۸.۲ رتبه‌بندی با استفاده از مدل‌های زبانی
۱۰۳	 ۹.۵.۲ اطلاعات بر مبنای صفحه‌ای: رتبه‌بندی
۱۰۵	 ۶.۲ خلاصه و جهت‌گیری‌های آینده
۱۰۶	 ۷.۲ منابع پیشنهادی
۱۱۱	 فصل ۳: بازیابی اطلاعات برحسب زبان متقابل
۱۱۲	 ۱.۳ مقدمه
۱۱۳	 ۲.۳ پیاده‌سازی بازیابی اطلاعات برحسب زبان متقابل
۱۱۳	 ۱.۲.۳ ترجمه پرسش و ترجمه سند
۱۱۵	 ۲.۲.۳ بدون ترجمه
۱۱۶	 ۳.۲.۳ انواع مختلف منابع ترجمه
۱۱۹	 ۴.۲.۳ چندمعنایی اصطلاح
۱۲۰	 ۳.۳ رویکردهای ترجمه بازیابی اطلاعات برحسب زبان متقابل
۱۲۰	 ۱.۳.۳ واژه‌نامه‌های ماشین خوان
۱۲۴	 ۲.۳.۳ رویکردهای آماری
۱۳۱	 ۳.۳.۳ بسط پرسش بیش‌تر ترجمه‌ای و پس‌ترجمه‌ای
۱۳۵	 ۴.۳.۳ ترجمه ماشینی
۱۳۶	 ۵.۳.۳ ترکیب رویکردها
۱۳۹	 ۴.۳ مدیریت زبان‌های متعدد
۱۳۹	 ۱.۴.۳ گردش کار بازیابی اطلاعات برحسب زبان متقابل
۱۴۴	 ۲.۴.۳ ادغام‌سازی در سراسر زبان‌ها

۱۴۷	۳.۴.۳ ترجمه سند
۱۴۷	۴.۴.۳ ترجمه غیرمستقیم
۱۴۸	۵.۳ خلاصه و جهت‌گیری‌های آینده
۱۵۲	۶.۳ منابع پیشنهادی
۱۵۷	فصل ۴: تعامل و رابطه‌ای کاربری
۱۵۸	۱.۴ جستجوی اطلاعات و تعامل کاربر
۱۶۵	۲.۴ نیازهای اطلاعاتی کاربران و وظایف جستجو
۱۶۸	۳.۴ مهارت‌های زبانی کاربران و تفاوت‌های فرهنگی
۱۷۲	۴.۴ پشتیبانی تعامل چندزبانه کاربر
۱۷۴	۱.۴.۱ فرایندی و ترجمه پرسش
۱۸۵	۲.۴.۱ گزینش و ارزیابی سند
۱۹۲	۳.۴.۴ فرمول‌دهی و مدل پرسش
۱۹۶	۴.۴.۴ تورق و سری و سری
۲۰۲	۵.۴ طراحی رابط‌های کاربری جستجو چندزبانه
۲۰۶	۱.۵.۴ طراحی کاربرمدار
۲۰۹	۲.۵.۴ بین‌المللی‌سازی و بومی‌سازی
۲۱۴	۳.۵.۴ مطالعه موردی: بازیابی اطلاعات در زبان متقابل در جستجو وبی گوگل
۲۱۷	۶.۴ خلاصه و جهت‌گیری‌های آینده
۲۱۸	۷.۴ منابع پیشنهادی
۲۲۷	فصل ۵: ارزیابی سامانه‌های بازیابی اطلاعات چندزبانه
۲۲۸	۱.۵ مقدمه
۲۲۹	۲.۵ ارزیابی سامانه‌محور
۲۳۰	۱.۲.۵ آزمون کرانفیلد
۲۳۳	۲.۲.۵ پوشش‌های ارزیابی
۲۳۴	۳.۲.۵ ایجاد مجموعه آزمون
۲۴۸	۴.۲.۵ ارتقاء پژوهش درباره توسعه سامانه چندزبانه و چندرسانه‌ای از طریق ارزیابی
۲۵۳	۵.۲.۵ روش‌شناسی‌های جایگزین برای ایجاد مجموعه آزمون

۶.۲.۵	سنجه‌های عملکرد	۲۵۶
۷.۲.۵	آزمون معناداری آماری	۲۷۰
۸.۲.۵	کارآیی سامانه و رضایت کاربر	۲۷۲
۳.۵	ارزیابی کاربر مدار	۲۷۳
۱.۳.۵	طرح آزمایشی	۲۷۵
۲.۳.۵	ارزیابی سامانه‌های بازیابی اطلاعات برحسب زبان متقابل در سخنگاه ارزیابی برحسب	
	زبان متقابل	۲۷۸
۳.۵.۸	سنجه‌های عملکرد جایگزین	۲۸۱
۴.۵	رزیابی سامانه خودتان	۲۸۳
۵.۵	خلاصه و چندی‌های آینده	۲۸۸
۶.۵	منابع پیش‌هادی	۲۹۰
فصل ۶:	نرم‌افزارهای دسترسی به اطلاعات چندزبانه	۲۹۷
۱.۶	مقدمه	۲۹۸
۲.۶	فراتر از بازیابی سند متنی چندزبانه	۲۹۹
۱.۲.۶	بازیابی تصویر	۳۰۰
۲.۲.۶	بازیابی گفتار	۳۰۸
۳.۲.۶	بازیابی ویدئو	۳۱۳
۴.۲.۶	سامانه پرسش و پاسخ	۳۱۹
۳.۶	دسترسی به اطلاعات چندزبانه در عمل	۳۲۹
۱.۳.۶	جستجوی وبی	۳۲۹
۲.۳.۶	کتابخانه‌های دیجیتال و میراث فرهنگی	۳۳۲
۳.۳.۶	پزشکی و بهداشت	۳۳۸
۴.۳.۶	دولت و حقوق	۳۴۱
۵.۳.۶	کسب و کار و تجارت	۳۴۵
۴.۶	جمع‌بندی	۳۴۹
واژه‌نامه	نشانه‌های اختصاری	۳۵۹