



جهاد دانشگاهی استان کرمان

روش های پیشرفته بازیابی اطلاعات

تألیف:

دکتر عادل سلیمانی نژاد

عضو هیات علمی دانشگاه شهید باهنر کرمان

آمنه سلیمانی

سرشناسه	: سلیمانی نژاد ، عادل
عنوان و نام پدیدآور	: روش های پیشرفته بازیابی اطلاعات /عادل سلیمانی نژاد ، آمنه سلیمانی.
مشخصات نشر	: کرمان : جهاددانشگاهی ، واحد کرمان، ۱۳۹۳.
مشخصات ظاهری	: ۲۵۸ ص. : نمودار .
وضعیت فهرست نویسی	: فیا
یادداشت	: کتابنامه
موضوع	: ذخیره و بازیابی اطلاعات
موضوع	: تکنولوژی اطلاعات
شناسه افزوده	: سلیمانی ، آمنه ۱۳۵۹-
شناسه افزوده	: جهاددانشگاهی، سازمان انتشارات. واحد کرمان
رده بندی کنگره	: ۱۳۹۳ ۸۹/س ۶۹۹/ Z
رده بندی دیویی	: ۴۴/ ۲۵
شماره کتابشناسی ملی	: ۲۵۳۱۴۴۰

نام کتاب	: روش های پیشرفته بازیابی اطلاعات
مؤلفان	: عادل سلیمانی نژاد - آمنه سلیمانی
ناشر	: جهاد دانشگاهی استان کرمان
تاریخ انتشار	: ۱۳۹۳
تیراژ	: ۵۰۰ جلد
نوبت چاپ	: اول
شابک	: ۹۷۸-۶۰۰-۶۷۸۹-۱۹-۴
قیمت	: ۱۶۱۰۰ تومان

فصل اول: کلیات	۱
تعاریف و اصطلاحات	۱
بازیابی اطلاعات	۱
داده	۱
اطلاعات	۲
دانش	۲
تعریف یک نظام بازیابی اطلاعات	۲
تاریخچه نظامهای بازیابی اطلاعات	۳
عناصر تشکیل دهنده نظامهای بازیابی اطلاعات	۷
اهداف سیستم بازیابی اطلاعات	۸
وظایف یک سیستم بازیابی اطلاعات	۸
اجزا یک سیستم بازیابی اطلاعات (چهارچوب کلی یک نظام بازیابی اطلاعات)	۹
فصل دوم- پایگاه داده	۱۳
تعاریف	۱۳
پایگاه داده	۱۳
تفاوت بانک داده با پایگاه داده	۱۴
کاراکتر	۱۴
فیلد	۱۵
رکورد	۱۵
فایل	۱۵
استقلال داده	۱۶
جامعیت داده	۱۶
وحدت	۱۶
افزودگی	۱۶
موجودیت	۱۷
صفات خاصه	۱۷
ارزش	۱۷
ارتباط	۱۷

۱۹.....	تاریخچه پایگاه داده.....
۲۱.....	نسلهای پایگاه داده:.....
۲۱.....	نسل اول: فایل ترتیبی.....
۲۲.....	نسل دوم: شیوههای دسترسی.....
۲۳.....	نسل سوم: سیستم مدیریت داده.....
۲۴.....	نسل چهارم: سیستم مدیریت پایگاه داده.....
۲۵.....	نسل پنجم: پایگاه معرفت.....
۲۵.....	نسل ششم.....
۲۶.....	انواع پایگاه داده.....
۲۶.....	انواع پایگاه داده (از دیدگاه سنتی).....
۲۷.....	انواع پایگاه داده (از دیدگاه طراحی).....
۳۲.....	عناصر تشکیل دهنده پایگاه داده.....
۳۳.....	سخت افزارهای ذخیره سازی.....
۳۳.....	سخت افزارهای پردازشگر.....
۳۳.....	سخت افزارهای ارتباطی.....
۳۴.....	نرم افزارهای سیستمی.....
۳۴.....	نرم افزارهای کاربردی.....
۳۴.....	طراحان.....
۳۴.....	مدیر پایگاه داده.....
۳۴.....	کاربر نهانی.....
۳۵.....	داده های ورودی.....
۳۵.....	دادههای عملیاتی.....
۳۵.....	دادههای خروجی.....
۳۵.....	ویژگیهای داده در پایگاه داده.....
۳۵.....	اشتراکی شدن.....
۳۵.....	ماندگاری.....
۳۵.....	امنیت.....
۳۶.....	طراحی پایگاه داده.....
۳۶.....	سطح خارجی.....

۳۶	سطح ادراکی
۳۶	سطح داخلی
۳۶	تبدیل سطح خارجی به ادراکی
۳۷	تبدیل سطح ادراکی به خارجی
۳۷	میزبان
۳۹	فصل سوم - استراتژیهای جستجو
۳۹	استراتژی جستجو
۴۰	فرآیند جستجو از نظر کلوردون
۴۱	انواع جستجو
۴۱	جستجوی ساده و جستجوی پیشرفته
۴۲	جستجوی عبارتی
۴۲	جستجوی مجاورتی (همجواری)
۴۳	جستجوی میدانی
۴۴	کوتاه سازی
۴۵	جستجوی زبان طبیعی
۴۵	کنترل واژگان
۴۶	حساسیت به حروف کوچک و بزرگ
۴۶	مرور نتایج
۴۷	رتبه بندی نتایج
۴۷	عملگرها
۴۷	انواع عملگرها
۴۹	چند نمونه عملگر
۵۲	نمایه سازی
۵۳	تعریف نمایه
۵۳	هدف از تهیه نمایه
۵۳	انواع نمایه
۵۶	اجزای یک نمایه
۵۶	زیانهای نمایه سازی
۵۷	مراحل نمایه سازی

۵۸	 عمق نمایه سازی
۵۹	 فصل چهارم - مدل سازی
۵۹	 مقدمه
۶۰	 شیوه های بازیابی اطلاعات
۶۱	 تورق
۶۱	 بازیابی
۶۱	 مدل های بازیابی اطلاعات کلاسیک
۶۴	 توصیف کلی از مدل های بازیابی اطلاعات
۶۴	 بازیابی اطلاعات کلاسیک
۶۴	 مفاهیم اولیه
۶۵	 مدل بولی
۶۹	 مدل فضای برداری
۷۶	 مدل احتمالی
۷۹	 مقایسه مختصر مدل های کلاسیک
۸۰	 مدل فازی
۸۷	 مدل بولی گسترش یافته
۸۹	 فصل پنجم - پروتکلها
۸۹	 تعریف پروتکل
۸۹	 لزوم استفاده از سیستم های کد گذاری
۹۰	 انواع سیستم های کد گذاری
۹۰	 ASCII
۹۱	 UNICODE
۹۲	 UTF-8
۹۳	 Z39.50
۹۳	 تاریخچه
۹۴	 اهداف
۹۴	 کاربرد در کتابخانه ها
۹۵	 MARC
۹۶	 تاریخچه

۹۶		ساختمان مارک
۹۷		مارک ایران
۹۷		بلوکهای اصلی مارک ایران
۹۸		UNIMARC
۹۹		اِبَر داده
۱۰۰		تعریف
۱۰۰		کارکردهای اصلی اِبَر داده
۱۰۱		طرحهای اِبَر داده ای
۱۰۲		استاندارد اِبَر داده ای دو بِلِن کور
۱۰۴		FRBR استاندارد اِبَر داده ای
۱۰۵		موجودیت های FRBR
۱۰۷		زبانهای نشانه گذاری
۱۰۸		زبان نشانه گذاری SGML
۱۰۸		زبان نشانه گذاری HTML
۱۰۹		زبان نشانه گذاری XML
۱۱۲		زبان نشانه گذاری XHTML
۱۱۳		فصل ششم - بازیابی اطلاعات چند رسانه ای
۱۱۳		منابع اطلاعاتی چند رسانه ای
۱۱۴		بازیابی اطلاعات چند رسانه ای
۱۱۵		بازیابی صوت
۱۱۶		بازیابی موسیقی
۱۱۷		بازیابی تصویر
۱۱۷		حوزه های کاربرد
۱۱۸		سیستم های اولیه بازیابی تصاویر
۱۱۹		بازیابی محتوایی تصویر
۱۲۰		بازیابی اطلاعات ویدئویی
۱۲۱		موتورهای کاوش چند رسانه ای
۱۲۴		رویکردهای نوین بازیابی چند رسانه ای

۱۲۵.....	کتابخانه دیجیتالی.....
۱۲۶.....	تعریف کتابخانه دیجیتالی.....
۱۲۷.....	منابع اطلاعاتی کتابخانه دیجیتالی.....
۱۲۸.....	اهداف کتابخانه دیجیتالی.....
۱۲۸.....	ویژگیهای کتابخانه دیجیتالی.....
۱۲۹.....	تجهیزات مورد نیاز کتابخانه دیجیتالی.....
۱۳۰.....	فرآیندهای ایجاد یک کتابخانه دیجیتالی.....
۱۳۱.....	بازایی اطلاعات در کتابخانه دیجیتالی.....
۱۳۱.....	انواع سیستمهای بازایی اطلاعات در کتابخانه دیجیتالی.....
۱۳۴.....	قابلیتهای رایج بازایی اطلاعات در کتابخانه دیجیتالی.....
۱۳۷.....	فصل هفتم - اینترنت و وب.....
۱۳۷.....	مقدمه.....
۱۴۱.....	روشهای دستیابی به اطلاعات.....
۱۴۲.....	ابزارهای کاوش اینترنت.....
۱۴۸.....	موتورهای کاوش.....
۱۵۰.....	ابرموتورهای کاوش.....
۱۵۰.....	راهنماهای موضوعی.....
۱۵۱.....	معیارهای انتخاب ابزارهای کاوش.....
۱۵۲.....	وب پنهان.....
۱۵۳.....	دلایل ایجاد وب پنهان.....
۱۵۴.....	بخشهای وب پنهان.....
۱۵۶.....	اهمیت وب پنهان.....
۱۵۷.....	شیوه های اطلاع یابی در وب نامرئی.....
۱۶۱.....	رابط کاربر.....
۱۶۴.....	طراحی رابط کاربر.....
۱۶۶.....	جستجوی اطلاعات و رابط کاربر.....
۱۶۶.....	رابط های کاربر صفحه نمایش نتایج.....
۱۶۷.....	فصل هشتم-بازایی اطلاعات و وب معنایی.....
۱۶۷.....	وب ۲.....

تاریخچه وب ۲.....	۱۶۹
عوامل پیدایش وب ۲.....	۱۶۹
تفاوت وب ۱ و وب ۲.....	۱۷۰
مزایای وب ۲.....	۱۷۰
فناوریهای وب ۲.....	۱۷۱
وب معنایی.....	۱۷۴
تاریخچه وب معنایی.....	۱۷۵
تعریف وب معنایی.....	۱۷۵
مقایسه وب معنایی و وب کنونی.....	۱۷۷
معماری وب معنایی.....	۱۷۸
اجزای اصلی وب معنایی.....	۱۸۲
رابطه بین رده بندی، هستی شناسی، پایگاه دانش.....	۱۸۵
طیف آنتالوژی.....	۱۸۶
Wrapper.....	۱۹۰
بازیابی اطلاعات مبتنی بر وب معنایی.....	۱۹۲
موتورهای کاوش وب معنایی.....	۱۹۳
فصل نهم - بازیابی هوشمند اطلاعات	۱۹۵
مقدمه.....	۱۹۵
هوش مصنوعی.....	۱۹۶
تاریخچه هوش مصنوعی.....	۱۹۶
ویژگیهای یک سیستم هوشمند.....	۱۹۷
تواناییهای یک سیستم هوشمند.....	۱۹۷
بخشهای اصلی یک سیستم هوشمند.....	۱۹۸
کاربردهای هوش مصنوعی.....	۱۹۸
سیستمهای خبره.....	۱۹۸
تاریخچه سیستمهای خبره.....	۱۹۹
انواع سیستمهای خبره.....	۲۰۰
ویژگیهای یک سیستم خبره.....	۲۰۰
اجزاء سیستمهای خبره.....	۲۰۱

۲۰۲		کاربردهای سیستمهای خبره
۲۰۲		مزایای سیستمهای خبره
۲۰۳		محدودیتهای سیستمهای خبره
۲۰۳		سیستم های خبره در اطلاع رسانی
۲۰۷		فصل دهم- ارزیابی نظامهای بازیابی
۲۰۷		مقدمه
۲۰۷		لزوم ارزیابی نظامهای بازیابی
۲۰۸		معیارهای ارزیابی
۲۱۰		باز یافت یا جامعیت
۲۱۱		دقت یا مانعیت
۲۱۱		رابطه بین جامعیت و مانعیت
۲۱۴		یک سیستم ایده آل
۲۱۵		ریزش کاذب
۲۱۶		شاخص F
۲۱۶		شاخص E
۲۱۷		شاخصهای کاربر-مدار
۲۱۸		پژوهشهای ارزیابی بازیابی اطلاعات
۲۱۹		کرانفیلد ۱
۲۱۹		کرانفیلد ۲
۲۲۰		مدلاوز
۲۲۰		اسمارت
۲۲۰		استایرز
۲۲۱		ترک
۲۲۲		جدول اختصارات
۲۲۵		لیست موتورهای کاوش براساس نوع
۲۲۹		صد موتور کاوش آکادمیک
۲۳۹		فهرست منابع و مآخذ

توسعه همه‌جانبه‌ی کشور عزیزمان ایران، مستلزم توسعه‌ی تحقیق، پژوهش و بهره‌گیری از فناوری‌های نوین است که بخش عمده‌ای از این مسئولیت بزرگ، بر عهده‌ی استادان، محققان، دانشجویان و مؤسسات علمی و تحقیقاتی است. جهاددانشگاهی نیز به عنوان نهادی برخاسته از انقلاب اسلامی با اهداف علمی، پژوهشی و فرهنگی، چاپ و نشر کتب علمی و دانشگاهی که نتیجه‌ی تحقیق، پژوهش و تلاش علمی و اندیشمندانه‌ی همکاران محترم عضو هیئت علمی دانشگاه‌ها و مراکز تحقیقاتی می‌باشد را در کنار دیگر مسئولیت‌های خطیر خود، رسالتی مهم تلقی می‌نماید. امید که خداوند متعال ما را در انجام این مهم، بیش از پیش یاری فرماید.

دکتر رضا کامیاب مقدس
رئیس جهاددانشگاهی استان کرمان

پیشگفتار

رشد سریع و توسعه اینترنت و منابع اطلاعاتی باعث تغییرات بسیار زیادی در دنیای بازیابی اطلاعات شده است. حجم روز افزون منابع اطلاعاتی الکترونیکی و وبی همچنین نیاز شدید کاربران به منابع اطلاعاتی مرتبط، مقوله بازیابی اطلاعات را بیش از گذشته با اهمیت کرده است، تا جائیکه پیدایش انواع مدل های بازیابی اطلاعات، موتورهای جستجو پیشرفته و ابر موتورهای جستجو را باعث شده است. با وجود منابع نسبتاً خوبی که در حوزه بازیابی اطلاعات وجود دارد، نیاز به منبع جامع و روزآمد احساس می شود، تا بازیابی از منابع جدید را هم در برگیرد. از این رو کتاب حاضر با هدف رویکرد جدید بازیابی اطلاعات از منابع اطلاعات وبی تدوین شده است. در این اثر علاوه بر مفاهیم، استانداردها، مدل ها، روش ها و فنون بازیابی اطلاعات در منابع اطلاعاتی، به روش های پیشرفته بازیابی اطلاعات با استفاده از موتورهای جستجو و ابر موتورهای جستجو در منابع وبی بخصوص وب نامرئی و وب معنی پرداخته شده است. این کتاب، بر اساس سرفصل های درس ذخیره و بازیابی اطلاعات مقطع کارشناسی ارشد و درس نظام های بازیابی اطلاعات پیشرفته مقطع دکتری علم اطلاعات تهیه شده است. همچنین منبع مناسبی برای دانشجویان رشته های مهندس کامپیوتر و مهندسی فن آوری اطلاعات و سایر متخصصان و کاربران منابع اطلاعاتی الکترونیکی و وبی می باشد. بدون شک این اثر خالی از ایراد و اشکال نمی باشد. پیشاپیش از راهنمایی های ارزنده و نظرات اصلاحی اساتید و دانشجویان محترمی که این منبع را مورد استفاده قرار می دهند و ما را در رفع اشکالات احتمالی یاری می دهند تشکر و قدر دانی می کنیم. بر خود لازم می دانیم از تمام کسانی که در چاپ این کتاب متحمل زحمت شده اند، بخصوص خانواده هایمان و مدیریت محترم انتشارات جهاد دانشگاهی کرمان صمیمانه قدر دانی نمائیم. اثر حاضر هر چند لایق ولی امرمان حضرت حجت نیست، لذا عشق و ارادت به حضرتش جرئت تقدیم این کار به پیشگاهش را به ما داد. امیدواریم مقبول قرار گیرد.

دکتر عادل سلیمانی نژاد

آمنه سلیمانی