

نظام‌های بازیابی اطلاعات متنی

چارلز تی. میدو

برت آر. بویس

دونالد اچ. کرفت

کارول باری

ترجمه:

دکتر نجلا حریری

(دانشیار دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران)



عنوان و نام پدیدآور	: نظام‌های بازیابی اطلاعات متنی / مؤلفان [چارلز تی. میدو ... (و دیگران)] :
مشخصات نشر	: ترجمه نجلا حریری. تهران: چاپار، ۱۳۹۰.
مشخصات ظاهری	: ۶۰۸ ص.: مصور، جدول، نمودار.
شابک	: 978-600-5294-91-0
وضعیت فهرست‌نویسی	: فیبا
یادداشت	: [مؤلفان] چارلز تی. میدو، برت آر. بويس، دونالد اچ. گرفت، کارول باری.
یادداشت	: عنوان اصلی: Text information retrieval systems, 3rd ed, 2007.
یادداشت	: واژه‌نامه.
یادداشت	: کتابنامه: ص. ۵۵۱.
یادداشت	: نمایه.
موضوع	: ذخیره و بازیابی اطلاعات
موضوع	: متن‌پردازی
موضوع	: بازیابی اطلاعات
شنامه افزوده	: میدو، چارلز تی.
شنامه افزوده	: Meadow, Charles T.
شنامه افزوده	: حریری، نجلا، ۱۳۴۰ -، مترجم
رده‌بندی کنگره	: ۱۳۹۰ قن Z۶۹۹۹
رده‌بندی دیویی	: ۰۲۵/۰۴
شماره کتابشناسی ملی	: ۲۵۲۱۳۷۵



تهران: ۸۸۸۹۹۶۸۰ - ۹۱۲۱۹۶۰۲۱۴

نظام‌های بازیابی اطلاعات متنی

مؤلفان: چارلز تی. میدو، برت آر. بويس، دونالد اچ. گرفت، کارول باری
ترجمه: نجلا حریری

طراحی جلد و صفحه‌آرایی: رحیم کبیر صابر

چاپ اول، ۱۳۹۰

ناظر فنی چاپ: مریم تفضلی

تیراژ: ۲۰۰۰ نسخه

چاپ و صحافی: چاپخانه آربین

شابک: ۹۷۸-۶۰۰-۵۲۹۴-۹۱-۰ ISBN: 978-600-5294-91-0

۱۵۰۰۰ تومان

فهرست

۱۹	پیشگفتار مترجم
۲۱	مقدمه نویسندگان
	فصل اول
۲۵	مقدمه
۲۵	۱.۱ اطلاعات چیست؟
۲۷	۲.۱ بازیابی اطلاعات چیست؟
۳۲	۳.۱ بازیابی اطلاعات چگونه انجام می‌شود؟
۳۵	۱.۳.۱ گام‌های مربوط به کاربر
۴۰	۲.۳.۱ گام‌های مربوط به تولیدکننده پایگاه اطلاعاتی
۴۶	۳.۳.۱ طراحی و کارکرد نظام
۴۹	۴.۳.۱ چرا فرایند بازیابی اطلاعات بی‌نقص نیست؟
۵۱	۴.۱ چه کسانی از نظام‌های بازیابی اطلاعات استفاده می‌کنند؟
۵۲	۱.۴.۱ متخصصان اطلاع‌رسانی
۵۳	۲.۴.۱ کاربران نهایی متخصص موضوعی
۵۵	۳.۴.۱ کاربران نهایی غیرمتخصص موضوعی
۵۵	۵.۱ مشکلات طراحی و استفاده از نظام‌های بازیابی اطلاعات کدامند؟
۵۶	۱.۵.۱ طراحی
۵۷	۲.۵.۱ شناخت رفتار کاربر
۵۸	۶.۱ تاریخچه مختصر بازیابی اطلاعات
۵۹	۱.۶.۱ روش‌های سنتی بازیابی اطلاعات
۶۲	۲.۶.۱ نظام‌های بازیابی اطلاعات پیش از رایانه
۶۸	۳.۶.۱ نظام‌های رایانه‌ای خاص - منظوره
۷۱	۴.۶.۱ نظام‌های رایانه‌ای عام - منظوره
۷۲	۵.۶.۱ خدمات پایگاه‌های اطلاعاتی آنلاین
۷۷	۶.۶.۱ وب جهانی

۸۲..... منابع برای مطالعه بیشتر

فصل دوم

۸۳..... داده‌ها، اطلاعات و دانش

۸۳..... ۱.۲ مقدمه

۸۴..... ۲.۲ تعاریف

۸۵..... ۱.۲.۲ داده‌ها

۸۵..... ۲.۲.۲ اطلاعات

۸۸..... ۳.۲.۲ اخبار

۸۹..... ۴.۲.۲ دانش

۹۰..... ۵.۲.۲ هوش

۹۱..... ۶.۲.۲ معنا

۹۲..... ۷.۲.۲ خود

۹۳..... ۸.۲.۲ ربط و ارزش

۹۴..... ۳.۲ فراداده

۹۸..... ۴.۲ پایگاه دانش

۱۰۲..... ۵.۲ باور، باور موجه و دیدگاه

۱۰۶..... ۶.۲ خلاصه

فصل سوم

۱۰۷..... بازنمون اطلاعات

۱۰۷..... ۱.۳ اطلاعاتی که باید بازنمایی شود

۱۱۵..... ۲.۳ انواع بازنمون

۱۱۷..... ۱.۲.۳ زبان طبیعی

۱۱۹..... ۲.۲.۳ زبان طبیعی محدودشده

۱۲۱..... ۳.۲.۳ زبان مصنوعی

۱۲۲..... ۴.۲.۳ کدها، مقادیر و توصیفگرها

۱۲۴..... ۵.۲.۳ مدل‌های ریاضی متن

۱۲۸..... ۳.۳ ویژگی‌های بازنمون‌های اطلاعات

۱۲۸..... ۱.۳.۳ توان متمایزسازی

۱۲۹..... ۲.۳.۳ تشخیص مشابهت

۱۲۹..... ۳.۳.۳ توصیفی بودن

۱۳۰..... ۴.۳.۳ ابهام

۱۳۰..... ۵.۳.۳ اختصار

۱۳۱..... ۴.۳ رابطه بین موجودیت‌ها و ارزش‌های خصوصیت‌ها

۱۳۱..... ۱.۴.۳ کدهای سلسله‌مراتبی

۱۳۲.....	۲.۴.۳ اندازه‌گیری‌ها.....
۱۳۴.....	۳.۴.۳ توصیفگرهای اسمی.....
۱۳۵.....	۴.۴.۳ زبان تصریفی.....
۱۳۵.....	۵.۴.۳ متن کامل.....
۱۳۶.....	۶.۴.۳ نشانگرها و پیوندهای آشکار.....
۱۳۷.....	۵.۳ خلاصه.....

فصل چهارم

۱۳۹.....	ارزش‌ها و محتوای خصوصیت‌ها.....
۱۳۹.....	۱.۴ انواع نمادهای خصوصیت‌ها.....
۱۴۰.....	۱.۱.۴ اعداد.....
۱۴۱.....	۲.۱.۴ رشته کاراکترها: نام‌ها.....
۱۴۳.....	۳.۱.۴ سایر انواع رشته کاراکترها.....
۱۴۳.....	۲.۴ روابط رده‌ای.....
۱۴۴.....	۱.۲.۴ رده‌بندی سلسله‌مراتبی.....
۱۴۶.....	۲.۲.۴ روابط شبکه‌ای.....
۱۴۸.....	۳.۲.۴ عضویت در رده: دودویی، احتمالی، و فازی.....
۱۵۲.....	۳.۴ تغییر شکل ارزش‌ها.....
۱۵۳.....	۱.۳.۴ تغییر شکل کلمات از طریق ریشه‌گیری.....
۱۵۹.....	۲.۳.۴ تغییر شکل کلمات بر اساس اصوات.....
۱۶۱.....	۳.۳.۴ تغییر شکل کلمات از طریق معنا.....
۱۶۴.....	۴.۳.۴ تغییر شکل تصاویر گرافیکی.....
۱۷۰.....	۵.۳.۴ تبدیل و بازنمایی صوت.....
۱۷۳.....	۴.۴ منحصر به فرد بودن ارزش‌ها.....
۱۷۵.....	۵.۴ ابهام در ارزش‌های خصوصیت‌ها.....
۱۷۹.....	۶.۴ نمایه‌سازی متن.....
۱۸۲.....	۷.۴ کنترل واژگان.....
۱۸۳.....	۱.۷.۴ عناصر کنترل.....
۱۸۵.....	۲.۷.۴ اشاعه واژگان کنترل شده.....
۱۸۶.....	۸.۴ اهمیت دیدگاه.....
۱۸۸.....	۹.۴ خلاصه.....

فصل پنجم

۱۸۹.....	مدل‌های ساختار مجازی داده‌ها.....
۱۸۹.....	۱.۵ مفهوم مدل‌های داده‌ها.....
۱۹۴.....	۲.۵ عناصر و ساختارهای اصلی داده‌ها.....

۱۹۴.....	متغیرهای اسکالر و ثابت‌ها	۱.۲.۵
۱۹۵.....	متغیرهای بُرداری	۲.۲.۵
۱۹۵.....	ساختارها	۳.۲.۵
۱۹۵.....	آرایه‌ها	۴.۲.۵
۱۹۶.....	چندتایی‌ها	۵.۲.۵
۱۹۸.....	رابطه‌ها	۶.۲.۵
۱۹۹.....	متن	۷.۲.۵
۲۰۲.....	مدل‌های ساختاری متداول	۳.۵
۲۰۲.....	مدل ترتیبی خطی	۱.۳.۵
۲۰۴.....	مدل رابطه‌ای	۲.۳.۵
۲۰۷.....	مدل‌های سلسله‌مراتبی و شبکه‌ای	۳.۳.۵
۲۰۹.....	کاربرد مدل‌های اصلی	۴.۵
۲۱۰.....	فراמתن	۱.۴.۵
۲۱۳.....	فایل‌های صفحه‌گسترده	۲.۴.۵
۲۱۶.....	مدل موجودیت - رابطه	۵.۵
۲۲۰.....	خلاصه	۶.۵

فصل ششم

۲۲۱.....	ساختار فیزیکی داده‌ها	
۲۲۱.....	مقدمه‌ای بر ساختارهای فیزیکی	۱.۶
۲۲۳.....	ساختارهای رکوردها و اثرات آن‌ها	۲.۶
۲۲۳.....	ساختارهای اصلی	۱.۲.۶
۲۲۷.....	فضا - زمان و نرخ تراکنش	۲.۲.۶
۲۲۸.....	مفاهیم اساسی ساختار فایل	۳.۶
۲۲۸.....	ترتیب رکوردها	۱.۳.۶
۲۲۹.....	یافتن رکوردها	۲.۳.۶
۲۳۰.....	روش‌های سازماندهی	۴.۶
۲۳۱.....	فایل‌های ترتیبی	۱.۴.۶
۲۳۴.....	ساختارهای فایل نمایه‌دار	۲.۴.۶
۲۳۷.....	لیست‌ها	۳.۴.۶
۲۴۴.....	ساختارهای درختی	۴.۴.۶
۲۴۷.....	ساختارهای دستیابی مستقیم	۵.۴.۶
۲۴۹.....	تجزیه عناصر داده‌ای	۵.۶
۲۵۱.....	تجزیه عبارتی	۱.۵.۶
۲۵۲.....	تجزیه کلمه‌ای	۲.۵.۶
۲۵۳.....	تجزیه کلمه‌ای و عبارتی	۳.۵.۶

۶۶	ساخترهای ترکیبی	۲۵۴
۱۶۶	نمایه‌های تودرتو	۲۵۴
۲۶۶	ساختر مستقیم با زنجیره‌ها	۲۵۷
۳۶۶	روش دستیابی ترتیبی نمایه‌دار	۲۵۸
۷۶	خلاصه	۲۶۰

فصل هفتم

۲۶۳	پرسش از نظام بازیابی اطلاعات	۲۶۳
۱.۷	مقدمه	۲۶۳
۲.۷	انواع زبان‌ها	۲۶۶
۳.۷	منطق پرسش	۲۶۸
۱.۳.۷	مجموعه‌ها و زیر مجموعه‌ها	۲۷۰
۲.۳.۷	عبارت‌های رابطه‌ای	۲۷۱
۳.۳.۷	منطق بولی پرسش	۲۷۲
۴.۳.۷	مجموعه‌های رتبه‌بندی شده و فازی	۲۷۸
۵.۳.۷	اندازه‌گیری‌های مشابهت	۲۸۲
۴.۷	عملیاتی که انجام می‌شود	۲۸۴
۱.۴.۷	اتصال به یک نظام بازیابی اطلاعات	۲۸۴
۲.۴.۷	انتخاب پایگاه اطلاعاتی	۲۸۶
۳.۴.۷	کاوش فایل مقلوب یا اصطلاحنامه	۲۸۷
۴.۴.۷	ایجاد یک زیر مجموعه از پایگاه اطلاعاتی	۲۹۲
۵.۴.۷	کاوش برای رشته‌ها	۲۹۳
۶.۴.۷	تحلیل مجموعه	۲۹۷
۷.۴.۷	رج‌کردن، نمایش و فرمت رکوردها	۲۹۹
۸.۴.۷	نحوه کار در مورد رکوردهای ساختار نیافته	۳۰۰
۹.۴.۷	دانلود	۳۰۱
۱۰.۴.۷	سفارش اسناد	۳۰۳
۱۱.۴.۷	ذخیره، فراخوانی و ویرایش کاوش‌ها	۳۰۴
۱۲.۴.۷	کاوش آگاهی‌رسانی جاری	۳۰۴
۱۳.۴.۷	خلاصه هزینه‌ها	۳۰۷
۱۴.۴.۷	اتمام یک نشست کاوش	۳۰۷
۵.۷	مبنای محاسبه هزینه کاوش‌ها	۳۰۷

فصل هشتم

۳۱۱	تفسیر و اجرای پرسش	۳۱۱
۱.۸	مشکلات تفسیر زبان پرسش	۳۱۱

۳۱۳.....	تجزیه زبان دستوری.....	۱.۱.۸
۳۱۸.....	تجزیه زبان طبیعی.....	۲.۱.۸
۳۲۱.....	پردازش انتخاب از منو.....	۳.۱.۸
۳۲۲.....	اجرای دستورهای بازیابی.....	۲.۸
۳۲۳.....	انتخاب پایگاه اطلاعاتی.....	۱.۲.۸
۳۲۳.....	کاوش از فایل مقلوب.....	۲.۲.۸
۳۲۵.....	ایجاد مجموعه و زیرمجموعه.....	۳.۲.۸
۳۲۷.....	کوتاه‌سازی و کاراکترهای جهانی.....	۴.۲.۸
۳۲۹.....	کوتاه‌سازی از سمت چپ.....	۵.۲.۸
۳۳۴.....	تحلیل رکورد و دستورهای ارایه.....	۳.۸
۳۳۴.....	عملکردهای تحلیل مجموعه.....	۱.۳.۸
۳۳۶.....	نمایش فرمت و رج کردن.....	۲.۳.۸
۳۴۰.....	پرینت در حالت آفلاین.....	۳.۳.۸
۳۴۱.....	اجرای سایر دستورات.....	۴.۸
۳۴۱.....	سفارش.....	۱.۴.۸
۳۴۲.....	ذخیره، فراخوانی و ویرایش کاوش‌ها.....	۲.۴.۸
۳۴۲.....	آگاهی‌رسانی جاری.....	۳.۴.۸
۳۴۳.....	جمع‌زدن هزینه‌ها و صدور صورتحساب.....	۴.۴.۸
۳۴۵.....	اتمام یک نشست کاوش.....	۵.۴.۸
۳۴۶.....	ارایه بازخورد به کاربران و پیغام‌های خطا.....	۵.۸
۳۴۶.....	پاسخ به دستورهای خطا.....	۱.۵.۸
۳۴۸.....	نشان دادن اندازه مجموعه.....	۲.۵.۸
۳۴۸.....	نمایش رکورد.....	۳.۵.۸
۳۴۸.....	تحلیل مجموعه.....	۴.۵.۸
۳۴۹.....	هزینه.....	۵.۵.۸
۳۴۹.....	راهنمایی.....	۶.۵.۸

فصل نهم

۳۵۳.....	کاوش متن.....	
۳۵۳.....	مشکلات ویژه کاوش متن.....	۱.۹
۳۵۵.....	نکته‌ای در مورد واژگان و نمادها.....	۱.۱.۹
۳۵۶.....	وب معنایی.....	۲.۱.۹
۳۵۹.....	برخی ویژگی‌های متن و کاربردهای آن‌ها.....	۲.۹
۳۶۰.....	اجزای متن.....	۱.۲.۹
۳۶۱.....	کلمات مهم - نمایه‌سازی.....	۲.۲.۹
۳۶۵.....	جملات مهم - چکیده‌نویسی.....	۳.۲.۹

۴.۲.۹	اندازه‌گیری‌های کل متن	۳۶۸
۳.۹	زبان دستوری برای کاوش متن	۳۶۹
۱.۳.۹	بیان عضویت در یک مجموعه	۳۷۰
۲.۳.۹	بیان ظهور کلمه یا عبارت	۳۷۰
۳.۳.۹	بیان مجاورتی	۳۷۰
۴.۳.۹	کاوش متن وب-پایه	۳۷۵
۴.۹	وزن دادن به اصطلاحات	۳۷۶
۱.۴.۹	نمایش‌سازی با وزن‌ها	۳۷۸
۲.۴.۹	تخصیص خودکار وزن‌ها	۳۷۹
۳.۴.۹	اصلاح و بهبود وزن‌ها	۳۸۱
۵.۹	فنون تداعی کلمات	۳۸۱
۱.۵.۹	واژه‌نامه‌ها و اصطلاحنامه‌ها	۳۸۲
۲.۵.۹	اصطلاحنامه‌های کوچک	۳۸۳
۳.۵.۹	آمارهای هم‌ظهوری کلمات	۳۸۴
۴.۵.۹	ریشه‌گیری و تلفیق	۳۸۶
۶.۹	فنون تداعی متن یا رکورد	۳۸۷
۱.۶.۹	اندازه‌گیری‌های مشابهت	۳۸۸
۲.۶.۹	خوشه‌بندی	۳۹۱
۳.۶.۹	تطبیق امضاء	۳۹۶
۴.۶.۹	روش‌های متمایزکننده	۴۰۰
۷.۹	سایر پردازش‌های کلمات در متن	۴۰۳
۱.۷.۹	کلمات بازدارنده	۴۰۳
۲.۷.۹	جایگزینی کلمات با ریشه‌ها یا کلمات هم‌پیوند	۴۰۵
۳.۷.۹	فراوانی به عنوان ملاکی برای اهمیت کلمات	۴۰۵
۴.۷.۹	نکاتی در مورد محاسبه میزان ارتباط استاد	۴۰۶

فصل دهم

محاسبه ربط و رتبه‌بندی توسط نظام	۴۱۱
۱.۱۰ ارزش وضعیت بازیابی	۴۱۱
۲.۱۰ رتبه‌بندی	۴۱۲
۳.۱۰ روش‌های ارزیابی "ارزش وضعیت بازیابی"	۴۱۳
۱.۳.۱۰ مدل فضای برداری	۴۱۴
۲.۳.۱۰ مدل احتمالی	۴۱۶
۳.۳.۱۰ مدل بولی گسترش‌یافته	۴۱۷
۴.۱۰ ارزش وضعیت بازیابی (RSV) در بازیابی عملیاتی	۴۲۲

فصل یازدهم

۴۲۵.....	بازخورد و از سرگیری کاوش.....
۴۲۵.....	۱.۱۱ مفاهیم بنیادی بازخورد و از سرگیری کاوش.....
۴۲۸.....	۲.۱۱ توالی دستورها.....
۴۳۱.....	۳.۱۱ اطلاعات موجود به عنوان بازخورد.....
۴۳۱.....	۱.۲.۱۱ انتخاب فایل یا پایگاه اطلاعاتی.....
۴۳۲.....	۲.۳.۱۱ کاوش اصطلاح یا مرور.....
۴۳۴.....	۳.۳.۱۱ کاوش رکوردها و تشکیل مجموعه.....
۴۳۶.....	۴.۳.۱۱ نمایش رکوردها و مرور.....
۴۳۸.....	۵.۳.۱۱ دستیابی به رکوردها.....
۴۳۹.....	۶.۳.۱۱ درخواست اطلاعات در مورد نظام بازیابی.....
۴۴۰.....	۷.۳.۱۱ برقراری پارامترهای ارتباطی.....
۴۴۱.....	۸.۳.۱۱ موارد دیگری غیر از رشته اقدامات و چرخه‌ها.....
۴۴۲.....	۴.۱۱ تنظیمات در کاوش.....
۴۴۳.....	۱.۴.۱۱ اصلاح انتخاب واژه‌ها.....
۴۴۳.....	۲.۴.۱۱ اصلاح منطق تشکیل مجموعه.....
۴۴۴.....	۳.۴.۱۱ اصلاح اندازه مجموعه نهایی.....
۴۴۴.....	۴.۴.۱۱ اصلاح مانعیت (دقت)، جامعیت (بازیافت) یا کاربردپذیری کلی.....
۴۴۵.....	۵.۱۱ بازخورد از کاربر به نظام.....

فصل دوازدهم

۴۴۹.....	کاوش از پایگاه‌های اطلاعاتی چندگانه و نقشه‌برداری.....
۴۴۹.....	۱.۱۲ مفاهیم اساسی.....
۴۵۱.....	۲.۱۲ کاوش از پایگاه‌های اطلاعاتی چندگانه.....
۴۵۲.....	۱.۲.۱۲ ماهیت رکوردهای تکراری.....
۴۵۵.....	۲.۲.۱۲ یافتن رکوردهای تکراری.....
۴۶۰.....	۳.۲.۱۲ پوشش پایگاه‌های اطلاعاتی چندگانه.....
۴۶۲.....	۳.۱۲ نقشه‌برداری.....
۴۶۶.....	۴.۱۲ ارزش نقشه‌برداری.....

فصل سیزدهم

۴۶۹.....	راهبرد کاوش.....
۴۶۹.....	۱.۱۳ تأمل در ماهیت کاوش.....
۴۷۱.....	۱.۱.۱۳ کاوش اقلام مشخص.....
۴۷۱.....	۲.۱.۱۳ کاوش اطلاعات مشخص.....

۴۷۲.....	کاوش عمومی اطلاعات.....	۳.۱.۱۳
۴۷۲.....	اکتشاف در پایگاه اطلاعاتی.....	۴.۱.۱۳
۴۷۳.....	ماهیت راهبرد کاوش.....	۲.۱۳
۴۷۵.....	هدف کاوش.....	۱.۲.۱۳
۴۷۵.....	طرح کلی اقدامات.....	۲.۲.۱۳
۴۷۶.....	عناصر اطلاعاتی اصلی کاوش.....	۳.۲.۱۳
۴۷۷.....	طرح ویژه اقدامات.....	۴.۲.۱۳
۴۷۸.....	انواع راهبردها.....	۳.۱۳
۴۷۹.....	طبقه‌بندی بر اساس هدف.....	۱.۳.۱۳
۴۸۰.....	طبقه‌بندی بر اساس طرح اقدامات.....	۲.۳.۱۳
۴۸۳.....	تاکتیک‌ها.....	۴.۱۳
۴۸۴.....	تاکتیک‌های نظارت.....	۱.۴.۱۳
۴۸۴.....	تاکتیک‌های ساختار فایل.....	۲.۴.۱۳
۴۸۴.....	تاکتیک‌های فرمول‌بندی کاوش.....	۳.۴.۱۳
۴۸۴.....	تاکتیک‌های اصطلاح.....	۴.۴.۱۳
۴۸۵.....	خلاصه.....	۵.۱۳

فصل چهاردهم

۴۸۷.....	رابط کاربر نظام بازیابی اطلاعات.....	
۴۸۷.....	مدل عمومی جریان پیغام.....	۱.۱۴
۴۹۱.....	منابع ابهام.....	۲.۱۴
۴۹۳.....	نقش واسطه کاوش.....	۳.۱۴
۴۹۵.....	تعیین نیاز اطلاعاتی.....	۱.۳.۱۴
۴۹۶.....	تنظیم یک راهبرد کاوش.....	۲.۳.۱۴
۴۹۶.....	برگردان بیان نیاز اطلاعاتی به پرسش.....	۳.۳.۱۴
۴۹۶.....	تفسیر و ارزیابی برون‌داد.....	۴.۳.۱۴
۴۹۷.....	از سرگیری کاوش در درون یک طرح راهبردی.....	۵.۳.۱۴
۴۹۷.....	تفسیر راهبرد در صورت لزوم.....	۶.۳.۱۴
۴۹۸.....	راهنمایی در استفاده از نظام بازیابی اطلاعات.....	۷.۳.۱۴
۴۹۸.....	واسطه خودکار کاوش.....	۴.۱۴
۴۹۸.....	پیشرفت‌های اولیه.....	۱.۴.۱۴
۵۰۰.....	عملکردهای واسطه‌ای کاملاً خودکار.....	۲.۴.۱۴
۵۰۱.....	عملکردهای تعاملی واسطه.....	۳.۴.۱۴
۵۰۴.....	رابط کاربر به عنوانی بخشی از کلیه نظام‌ها.....	۵.۱۴
۵۰۶.....	رابط کاربر در موتورهای کاوش وب.....	۶.۱۴

فصل پانزدهم

سنجش و ارزیابی ۵۰۹

۱.۱۵ میانی اندازه‌گیری ۵۰۹

۱.۱.۱۵ مدیر داده‌ها ۵۱۱

۲.۱.۱۵ مدیر پرسش ۵۱۲

۳.۱.۱۵ فرایند فرمول‌بندی پرسش ۵۱۳

۴.۱.۱۵ استنتاج نیاز اطلاعاتی ۵۱۴

۵.۱.۱۵ پایگاه اطلاعاتی ۵۱۴

۶.۱.۱۵ کاربران ۵۱۵

۲.۱۵ ربط، ارزش و سودمندی ۵۱۶

۱.۲.۱۵ ربط به عنوان دأرا بودن ارتباط ۵۱۶

۲.۲.۱۵ جنبه‌های ارزش ۵۱۷

۳.۲.۱۵ ربط به عنوان سودمندی ۵۱۹

۴.۲.۱۵ حفظ هر دو اندازه‌گیری جداگانه ربط ۵۲۰

۵.۲.۱۵ مقیاس اندازه‌گیری ربط ۵۲۳

۶.۲.۱۵ به‌کارگیری مقیاس‌ها ۵۲۴

۷.۲.۱۵ تردیدهایی در مورد ربط به عنوان یک مقیاس ۵۲۶

۳.۱۵ مقیاس‌های مبتنی بر ربط ۵۲۷

۱.۳.۱۵ مانعیت (Pr) ۵۲۷

۲.۳.۱۵ جامعیت (Re) ۵۲۹

۳.۳.۱۵ رابطه جامعیت و مانعیت ۵۳۰

۴.۳.۱۵ اندازه‌گیری‌های اثربخشی کلی بر اساس جامعیت و مانعیت ۵۳۲

۴.۱۵ اندازه‌گیری‌های فرایند ۵۳۶

۱.۴.۱۵ برگردان پرسش ۵۳۷

۲.۴.۱۵ خطاهای بیان پرسش ۵۳۷

۳.۴.۱۵ متوسط زمان به ازای هر دستور یا تصمیم کاربر ۵۳۸

۴.۴.۱۵ مدت انجام کاوش ۵۳۸

۵.۴.۱۵ تعداد دستورات یا گام‌ها در یک کاوش ۵۳۸

۶.۴.۱۵ هزینه کاوش ۵۳۸

۷.۴.۱۵ اندازه مجموعه شکل گرفته نهایی ۵۳۹

۸.۴.۱۵ تعداد رکوردهای بازیابی شده توسط کاربر ۵۴۰

۹.۴.۱۵ الگوهای استفاده از زبان ۵۴۰

۱۰.۴.۱۵ اندازه‌گیری‌های ترتیب رتبه‌بندی ۵۴۴

۵.۱۵ اندازه‌گیری‌های نتیجه ۵۴۵

۱.۵.۱۵ مانعیت ۵۴۶

۵۴۶.....	جامعیت	۲.۵.۱۵
۵۴۷.....	کارآمدی	۳.۵.۱۵
۵۴۷.....	ارزیابی کلی کاربر	۴.۵.۱۵
۵۴۸.....	اندازه‌گیری‌های محیطی	۶.۱۵
۵۴۸.....	انتخاب رکوردها و پایگاه اطلاعاتی	۱۶.۱۵
۵۴۹.....	محتوای رکورد	۲۶.۱۵
۵۴۹.....	مطالعه و ارزیابی کاربران	۳۶.۱۵
۵۵۰.....	نتیجه‌گیری	۷.۱۵
۵۵۱.....	منابع	
۵۶۳.....	واژه‌نامه فارسی به انگلیسی	
۵۷۹.....	واژه‌نامه انگلیسی به فارسی	
۵۹۵.....	نمایه موضوعی	

www.ketab.ir

پیشگفتار مترجم

بازیابی اطلاعات به یک تعبیر، بنیادی‌ترین مفهوم در حوزه علم اطلاعات است، چرا که یافتن اسناد و اطلاعات مرتبط با نیازهای اطلاعاتی کاربران، در واقع جوهره اصلی کلیه فعالیت‌های کتابداری و اطلاع‌رسانی در تمامی دوران‌ها محسوب می‌شود. سهولت و کارآمدی فرایند ذخیره و بازیابی اطلاعات که امروزه شاهد آن هستیم، دستاوردی انقلابی است که از پیشرفت‌های فناوری اطلاعات و ارتباطات حاصل شده است. این فرایند، مفاهیم پیچیده‌ای را شامل می‌شود که توضیح و تشریح آن‌ها موضوع کتاب‌های خارجی متعددی بوده است. در ایران نیز در دهه‌های اخیر کتاب‌هایی در زمینه بازیابی اطلاعات منتشر شده است، اما انگیزه ترجمه کتاب حاضر، جای خالی یک منبع جامع در مورد نظام‌های بازیابی اطلاعات متنی بود که تمامی مباحث مربوط به نظام‌های بازیابی رایانه‌ای را پوشش بدهد. کتاب حاضر در پانزده فصل^۱ به شرح زیر تنظیم شده است:

فصل اول مقدمه‌ای شامل مباحث کلی مانند طراحی و کارکرد نظام، استفاده‌کنندگان نظام و انواع نظام‌های بازیابی اطلاعات است. در فصل دوم به مفاهیم زیربنایی در بازیابی اطلاعات مانند داده‌ها، اطلاعات و دانش پرداخته شده است. فصل سوم بازنمون اطلاعات و چگونگی آن را مطرح می‌کند. فصل چهارم، به

۱. متن اصلی کتاب دارای شانزده فصل است. فصل ۱۵ متن انگلیسی کتاب به معرفی نمونه‌هایی از نظام‌های بازیابی اختصاص دارد که به دلیل تحول سریع در رابط کاربر نظام‌های بازیابی و ناپایداری جزئیات ویژگی‌های ظاهری نظام‌ها، از ترجمه آن فصل صرف‌نظر گردید.

ارزش‌های خصوصیت‌ها و مباحث مرتبط مانند نمایه‌سازی و کنترل واژگان اختصاص دارد. در فصول پنجم و ششم به ترتیب به ساختار مجازی و ساختار فیزیکی داده‌ها پرداخته شده است. فصل‌های هفتم و هشتم به مباحث پرسش از نظام بازیابی اطلاعات و انواع زبان‌ها و منطق پرسش و تفسیر و اجرای پرسش اختصاص یافته‌اند. کاوش متن و محاسبه ربط و رتبه‌بندی نتایج کاوش در فصول نهم و دهم مطرح گردیده. بازخورد و ازسرگیری کاوش در فصل یازدهم و کاوش از پایگاه‌های اطلاعاتی چندگانه و نقشه‌برداری در فصل دوازدهم مطرح شده‌اند. فصل سیزدهم به راهبرد کاوش و ماهیت و انواع آن می‌پردازد و رابط کاربر نظام بازیابی اطلاعات موضوع بحث در فصل چهاردهم است. فصل پانزده که آخرین فصل کتاب است به مبحث مهم ارزیابی نظام‌های بازیابی اطلاعات و روش‌های و رویکردهای آن اختصاص یافته است.

کتاب نظام‌های بازیابی اطلاعات متنی چنان که نویسندگان کتاب نیز در مقدمه نویسندگان اشاره کرده‌اند می‌تواند منبع سودمندی برای مطالعه دانشجویان رشته‌های مرتبط با اطلاع‌رسانی و فناوری اطلاعات و برخی رشته‌های دیگر مانند ارتباطات و غیره باشد. در واقع کلیه افرادی که به نحوی با نظام‌های بازیابی اطلاعات سروکار دارند، اعم از طراحان نظام یا کاربرانی که علاقمند به درک کارکردهای نظام‌های بازیابی اطلاعات هستند، کتاب را سودمند خواهند یافت. کتاب حاضر می‌تواند در تدریس دروس ذخیره و بازیابی اطلاعات، نظام‌های بازیابی اطلاعات و دروس مشابه دیگر در رشته‌های فناوری اطلاعات، کتابداری و اطلاع‌رسانی و سایر رشته‌های مرتبط نیز مورد استفاده قرار بگیرد.

مقدمه نویسندگان

بازیابی اطلاعات یک فرایند ارتباطی است. کاربران نظام‌ها یا خدمات اطلاعاتی از طریق این فرایند می‌توانند اسناد، رکوردها، تصاویر گرافیکی یا موسیقی را که نیاز اطلاعاتی یا علائق آن‌ها را ارضاء می‌کند، بازیابی نمایند. موارد فوق را به طور کلی می‌توان اسناد نامید. ویژگی اصلی این نوع ارتباطات این است که این اسناد، قبل از این که استفاده شوند، به صورت پیام‌هایی تولید و ذخیره شده‌اند و برای کاوش و بازیابی در دسترس کاربران قرار گرفته‌اند. کاربران باید بدانند چگونه نیاز اطلاعاتی خود را ابراز کنند که البته نحوه بیان نیاز اطلاعاتی نیز همیشه مشخص نیست.

کاربران باید با چه کسی یا چه چیزی ارتباط برقرار کنند؟ فرایند ارتباطی، کاربران را با کتابدار، راهنمای موزه یا متخصص تشخیص اثر انگشت یا هر فرد دیگری که مسئول مجموعه‌ای از اسناد است، پیوند می‌دهد. این کار چگونه انجام می‌گیرد؟ برقراری ارتباط می‌تواند از طریق ارتباط مستقیم با این واسطه‌ها و یا از طریق ارتباط غیرمستقیم و با استفاده از ابزارهای بازیابی انجام بگیرد که توسط این واسطه‌ها و به همین منظور طراحی شده‌اند. فهرست کتابخانه یکی از انواع این ابزارهاست. یک نظام رایانه‌ای نیز در جامعه مدرن امروز نمونه دیگری از این ابزارهاست. این فرایند ارتباطی معمولاً شامل پردازش متن یا رشته‌هایی از کلمات و اعداد و یا کدهایی است که برای هر دو طرف فرایند آشنا هستند و می‌توانند برای توصیف محتوا و سایر خصوصیت‌های اسناد به کار بروند و اسناد را به نیازهای کاربران که با اصطلاحات مشابهی توصیف شده‌اند، پیوند بدهند.

این مسأله چندان هم ساده نیست و دانش تخصصی ویژه‌ای طلب می‌کند. مثلاً یافتن این که آیا یک اختراع قبلاً ثبت شده است یا نه، مستلزم دانش فنی در زمینه آن اختراع و نیز نحوه رده‌بندی پروانه‌های ثبت اختراع است. این کار ممکن است مستلزم استفاده از یک زبان ویژه یعنی زبان پرسش، برای بیان نیاز اطلاعاتی به رایانه باشد.

افراد معمولاً مشکلاتی در درک نقطه نظرات یکدیگر دارند و این مشکلات زمانی بیشتر می‌شود که مکالمه در مورد یک مسأله فنی است که تنها یکی از دو طرف دانش کافی در مورد آن دارد و در این جاست که ارتباط ممکن است به نحو جدی خدشه‌دار شود. حال اگر یکی از دو طرف ارتباط، رایانه باشد، مسأله پیچیده‌تر خواهد بود. به خصوص زمانی که ارتباط کاربران با رایانه به زبان طبیعی باشد و رایانه تنها در حد ابتدایی توانایی درک زبان کاربران را داشته باشد.

زمانی که اطلاعات از مکان ذخیره‌سازی آن بازیابی شد، وظیفه ارزیابی آن نیز به عهده کاربر خواهد بود. ما نمی‌توانیم هر اطلاعات بازیابی‌شده در مورد یک پروانه ثبت اختراع، یا هر مقاله در مورد سازمانی را که قصد سرمایه‌گذاری در آن را داریم، یا هر اثر انگشت بازیابی‌شده را بپذیریم، بلکه باید به نحو معقولی از بهترین بودن یا دست‌کم مناسب بودن اطلاعات اطمینان داشته باشیم. اگر نتوانیم آن چه را که به دنبال آن بودیم بازیابی کنیم، تکلیف چیست؟ چگونه فرایند را بهبود خواهیم داد؟ چگونه اسناد را از نظر کامل بودن، صحت و قابلیت اعتماد آن‌ها ارزیابی خواهیم کرد؟ نظام‌های بازیابی اطلاعات به این پرسش‌ها پاسخ نمی‌دهند، اما آگاهی از ساختار و عملکرد این نظام‌ها می‌تواند رویکردهای متفاوتی را برای کاوش در پیش روی کاربران قرار دهد. متخصصان حوزه بازیابی اطلاعات بر این باورند که کاربران در کاوش‌های خود چندان کمال‌گرا نیستند. یک دانش‌آموز دبیرستان معمولاً هر سند آسان‌فهم بازیابی‌شده در زمینه مورد کاوش خود را می‌پذیرد. اما اگر شیوه مدیران تجاری، وکلا یا پزشکان نیز چنین باشد، چه باید کرد؟ این متخصصان باید از نحوه کاوش، بازیابی و ارزیابی اطلاعات آگاه باشند.

هدف این کتاب، آموزش نحوه کاوش به کاربران نیست، بلکه کتاب بر آن است که نحوه کارکرد نظام‌های بازیابی اطلاعات را به کاوش‌کنندگان بیاموزد. شاید

بتوان گفت کتاب حاضر یک راهنما برای طراحان نظام‌های بازیابی اطلاعات است و به مشکلات مربوط به این زمینه و راه‌حل‌های موجود می‌پردازد. برای کاوش‌کننده نیز کتاب حاضر نحوه کارکرد نظام را تشریح می‌کند. کتاب عمدتاً در مورد نظام‌های بازیابی رایانه‌ای است اما اصول مطرح‌شده، در مورد بازیابی غیر مکانیکی نیز صدق می‌کند.

خواننده کتاب می‌تواند دانشجوی تحصیلات تکمیلی یا دانشجوی سال‌های بالاتر کارشناسی در رشته‌های اطلاعات، رایانه یا کتابداری و نیز ارتباطات یا روزنامه‌نگاری باشد. کارکنان حوزه خدمات اطلاع‌رسانی نیز کتاب را سودمند خواهند یافت. فرض بر این است که خواننده کتاب از یک پیش‌زمینه اطلاعاتی محدود در زمینه ریاضی و رایانه برخوردار باشد اما تلاش بر این بوده است که کلیه مفاهیم ریاضی پیچیده با اصطلاحات غیرفنی بیان شوند.

کتاب با مروری کلی بر بازیابی اطلاعات شروع می‌شود: چگونه نیاز اطلاعاتی به پرسش منجر می‌شود؟ استاد چگونه تولید و وارد نظام ذخیره و بازیابی می‌شوند؟ و اجزای اصلی یک نظام بازیابی اطلاعات کدام است. سپس با توضیح ماهیت اطلاعات و مفاهیم مرتبط دیگری مانند داده‌ها، خبر و دانش ادامه می‌یابد. این توضیحات صرفاً تعاریف ساده نیستند، زیرا آن چه اهمیت دارد تفاوت میان آنهاست: آن چه که برای یک نفر، خبر محسوب می‌شود برای دیگری ممکن است جنبه خبری نداشته باشد یا آن چه که برای فردی اطلاعات تلقی می‌شود برای دیگری ممکن است این ویژگی را نداشته باشد. عکس‌العمل کاربران در برابر نظام‌هایی که مدعی فراهم‌آوری اطلاعات هستند، بازتابی از برداشت آنها از اطلاعات است. اطلاعات چیز ملموسی نیست و با نمادها بازنمایی می‌شود. فصل سوم کتاب به توصیف برخی از ابزارهای بازنمایی اطلاعات اختصاص یافته است.

سپس به این بحث خواهیم پرداخت که نمادهایی که اطلاعات را بازنمایی می‌کنند، چگونه در رایانه سازماندهی می‌شوند و رایانه چگونه دستورات یا درخواست‌های پردازش (پرسش‌ها) را تفسیر و اجرا می‌کند. به دنبال آن، مفاهیم پیشرفته‌تری مانند رتبه‌بندی برون‌داد به ترتیب مورد علاقه کاربر، نقشه‌برداری نتایج یک کاوش به کاوش دیگر و استفاده از بازخورد درباره کاوش از هر دو سوی کاربر به رایانه و رایانه به کاربر، مطرح می‌شود.

هر چند کتاب حاضر عمدتاً به نحوه کاوش نمی‌پردازد، یک فصل به راهبرد کاوش اختصاص یافته است و به دنبال آن بحث رابط کاربر مطرح می‌شود که موضوع آن چگونگی وظایف هر یک از طرفین ارتباط رایانه-انسان در برقراری این ارتباط و کمک به طرف مقابل جهت تحقق هدف مشترک بازیابی نتایج راضی‌کننده از کاوش است. بخش نهایی به اندازه‌گیری و ارزیابی عملکرد نظام بازیابی اطلاعات می‌پردازد.

ویرایش سوم کتاب به این سبب انتشار می‌یابد که ویرایش دوم اندکی پس از ظهور وب جهانی و پیش از کاربرد گسترده و تبدیل آن به مهم‌ترین منبع اطلاعات، انتشار یافته بود. با ظهور وب، موتورهای کاوش نیز به ظهور رسیدند که هر چند با تعریف نظام‌های بازیابی اطلاعات مطابقت دارند، عموماً متفاوت از نظام‌های سنتی بازیابی اطلاعات هستند. تلاش کتاب بر این بوده است که به این تفاوت‌ها و برخی از امتیازات موتورهای کاوش نیز اشاره کند.

علاوه بر ما چهار نفر، افراد دیگری نیز در تولید کتاب و رفع خطاهای آن در نسخه‌های اولیه نقش داشته‌اند که البته هنوز هم احتمال خطا وجود دارد. نویسندگان از لوان فروند از دانشگاه تورنتو برای بازخورد مفید وی، همچنین از جو کوکس و کارکنان اینفوروم در گروه مطالعات اطلاع‌رسانی دانشگاه تورنتو، کمپانی دایالوگ، سازمان اطلاعات علمی و ملیا مالون دستیار پژوهشی در دانشگاه ایالتی لونیویانا تشکر می‌کنند.

چارلز تی میدو
برت آر بویس
دونالد اچ کرفت
کارول باری