



مؤسسه فرهنگی هنری
دیارگران تهران

به نام خدا



مؤسسه فرهنگی هنری
دیارگران تهران

ذخیره و بازیابی اطلاعات مفاهیم و رویکردها

مؤلف

دکتر علی احسانی

(دکترای مدیریت فناوری اطلاعات)



هرگونه چاپ و تکثیر از محتویات این کتاب بدون اجازه کتبی ناشر ممنوع است. متخلفان به موجب قانون حمایت حقوق مؤلفان، مصنفان و هنرمندان تحت پیگرد قانونی قرار می گیرند.

◀ عنوان کتاب: ذخیره و بازیابی اطلاعات مفاهیم و رویکردها

◀ مولف: دکتر علی احسانی

◀ ناشر: موسسه فرهنگی هنری دیباگران تهران

◀ صفحه آرای: فریوش عبدالهی

◀ طراح جلد: داریوش فرسای

◀ نوبت چاپ: اول

◀ تاریخ نشر: ۱۳۹۸

◀ چاپ و صحافی: درج عقیق

◀ تیراژ: ۱۰۰ جلد

◀ قیمت: ۷۵۰۰۰۰ ریال

◀ شابک: ۹۷۸-۶۲۲-۲۱۸-۲۴۳-۴

نشانی واحد فروش: تهران، میدان انقلاب،

خ کارگر جنوبی، روبروی پاساژ مهستان،

پلاک ۱۲۵۱

تلفن: ۲۲۰۸۵۱۱۱-۶۶۴۱۰۰۴۶

فروشگاههای اینترنتی دیباگران تهران:

WWW.MFTBOOK.IR

www.dibbook.ir

www.dibagarantehran.com

نشانی تلگرام: @mftbook

نشانی اینستاگرام دیبا [dibagaran_publishing](https://www.instagram.com/dibagaran_publishing)

هر کتاب دیباگران، یک فرصت جدید شغلی.

هر گوشی همراه، یک فروشگاه کتاب دیباگران تهران.

از طریق سایتها و اپ دیباگران، در هر جای ایران به کتابهای ما دسترسی دارید.

سرشناسه: احسانی، علی، ۱۳۶۶-
عنوان و نام پدیدآور: ذخیره و بازیابی اطلاعات، مفاهیم
و رویکردها
مشخصات نشر: تهران: دیباگران تهران: ۱۳۹۸
مشخصات ظاهری: ۲۸۴ص: مصور،
شابک: ۹۷۸-۶۲۲-۲۱۸-۲۴۳-۴
وضعیت فهرست نویسی: فیبا
موضوع: ذخیره و بازیابی اطلاعات
موضوع: Information storage and retrieval systems
موضوع: حفاظت داده ها
موضوع: Data protection
موضوع: کامپیوترها-حافظه-مدیریت
موضوع: Memory management (computer science)
رده بندی کنگره: QA ۷۶/۹
رده بندی دیویی: ۰۰۵/۷۴
شماره کتابشناسی ملی: ۶۰۲۸۵۶۹

فهرست مطالب

بخش اول : مفاهیم و اصطلاحات ذخیره و بازیابی ۱۰

فصل اول : رسانه های ذخیره سازی ۱۱

۱-۱- سیستم کامپیوتری ۱۱

۲-۱- حافظه ۱۲

۳-۱- نوار مغناطیسی ۱۳

۴-۱- گپ (GAP) ۱۳

۵-۱- دیسک مغناطیسی ۱۴

۶-۱- تقسیمات دیسک ۱۵

۷-۱- پارامترهای دیسک ۱۷

۸-۱- طبله (DRUM) ۱۹

فصل دوم: سیستم فایل و بلاک بندی ۲۰

۱-۲- سیستم فایل ۲۰

۲-۲- قالب های رکورد منطقی ۲۱

۳-۲- اشاره گر (نشانه رو) ۲۲

۴-۲- فایل ۲۳

۵-۲- ساختار فایل ۲۳

۶-۲- بلاک بندی ۲۴

۷-۲- فایل در محیط فیزیکی ۳۰

۸-۲- تکنیکهای تهیه پشتیبان ۳۲

۹-۲- سطوح نشانی دهی ۳۴

۱۰-۲- بافر ۳۷

فصل سوم: پارامترهای ارزیابی رسانه ها ۴۱

۱-۳- ظرفیت واقعی نوار ۴۱

۲-۳- محاسبه نرخ انتقال واقعی نوار ۴۱

۴۴ ۳-۳- محاسبه زمان کل پردازش فایل
۴۶ ۳-۴- تکنیک های بهبود کارایی سیستم فایل
۴۷ ۳-۵- توزیع فایل روی چند دیسک

بخش دوم: ساختار فایل ۴۹

فصل چهارم: ساختارهای فایل (پایل-ترتیبی) ۵۰

۵۰ ۴-۱- انواع ساختارهای فایل
۵۱ ۴-۲- انواع فایل
۵۲ ۴-۳- شرح اصول عملیات ششگانه
۵۵ ۴-۴- فایل با ساختار پایل (بی نظم)
۵۷ ۴-۵- معرفی ساختار ترتیبی

فصل پنجم: ساختارهای شاخص دار ۶۴

۶۴ ۵-۱- معرفی ساختارهای شاخص دار
۶۴ ۵-۲- انواع شاخص
۶۶ ۵-۳- ساختار ترتیبی شاخص دار (INDEXED SEQUENTIAL)
۷۳ ۵-۴- ساختار چند شاخصی (فایل شاخص دار)

فصل ششم: ساختار فایل مستقیم ۷۷

۷۷ ۶-۱- معرفی ساختار مستقیم
۷۷ ۶-۲- روش های حل مشکل تصادف و درج سرریزی ها
۷۹ ۶-۳- موارد استفاده فایل مستقیم
۸۱ ۶-۴- باکت بندی

بخش سوم: مفاهیم تکمیلی در ذخیره سازی اطلاعات ۸۳

فصل هفتم: آشنایی با ذخیره اطلاعات ۸۴

۸۵ ۷-۱- ذخیره اطلاعات
۹۲ ۷-۲- تکامل معماری ذخیره سازی
۹۳ ۷-۳- زیرساخت مرکز داده
۹۷ ۷-۴- مجازی سازی و رایانش ابری

فصل هشتم: مرکز داده ۹۹

۹۹	۱-۸- برنامه کاربردی
۱۰۰	۲-۸- سیستم مدیریت پایگاه داده
۱۰۰	۳-۸- میزبان (محاسب)
۱۰۸	۴-۸- اتصال
۱۱۱	۵-۸- ذخیره سازی
۱۱۲	۶-۸- دسترسی میزبان به داده ها
۱۱۴	۷-۸- ذخیره سازی اتصال مستقیم (DAS)
۱۱۶	۸-۸- صف فرمان
۱۱۷	۹-۸- آشنایی با درایوهای فلش

فصل نهم: حفاظت از داده ها: RAID ۱۲۰

۱۲۱	۱-۹- روش های اجرای RAID
۱۲۲	۲-۹- مولفه های ارابه RAID
۱۲۲	۳-۹- روشهای RAID
۱۲۶	۴-۹- سطوح RAID
۱۳۳	۵-۹- تأثیر RAID بر عملکرد دیسک
۱۳۵	۶-۹- مقایسه RAID

فصل دهم: سیستم های ذخیره سازی هوشمند ۱۳۷

۱۳۷	۱-۱۰- مؤلفه های یک سیستم ذخیره سازی هوشمند
۱۴۴	۲-۱۰- تأمین ذخیره ساز
۱۵۱	۳-۱۰- انواع سیستم های ذخیره سازی هوشمند

فصل یازدهم: فن آوری های ذخیره سازی تحت شبکه ۱۵۵

۱۵۶	۱-۱۱- کانال فیبر: بررسی اجمالی
۱۵۷	۲-۱۱- SAN و تکامل آن
۱۵۸	۳-۱۱- معماری کانال فیبر
۱۶۱	۴-۱۱- توپولوژی های FC SAN
۱۶۴	۵-۱۱- مجازی سازی در SAN
۱۶۸	۶-۱۱- IP SAN

۱۶۹	۷-۱۱- ذخیره سازی متصل به شبکه (NAS).....
۱۷۰	۸-۱۱- سرورها با اهداف عمومی در مقابل دستگاه های NAS.....
۱۷۱	۹-۱۱- مزایای NAS.....
۱۷۱	۱۰-۱۱- مؤلفه های NAS.....
۱۷۲	۱۱-۱۱- پروتکل به اشتراک گذاری فایل NAS.....
۱۷۴	۱۲-۱۱- عوامل مؤثر بر عملکرد NAS.....
۱۷۵	۱۳-۱۱- مجازی سازی سطح فایل.....

فصل دوازدهم: رایانش ابری ۱۷۸

۱۷۸	۱-۱۲- فن آوری های ایجاد CLOUD.....
۱۷۹	۲-۱۲- ویژگی های رایانش ابری.....
۱۸۰	۳-۱۲- مزایای محاسبات ابری.....
۱۸۱	۴-۱۲- مدل های سرویس ابر.....
۱۸۳	۵-۱۲- مدل استقرار ابر.....
۱۸۶	۶-۱۲- زیرساخت رایانش ابری.....
۱۹۰	۷-۱۲- چالش های ابری.....
۱۹۱	۸-۱۲- ملاحظات پذیرش ابر.....

فصل سیزدهم: ایمن سازی زیرساخت های ذخیره سازی ۱۹۴

۱۹۴	۱-۱۳- چارچوب امنیت اطلاعات.....
۱۹۵	۲-۱۳- ریسک سه گانه.....
۱۹۹	۳-۱۳- دامنه های امنیتی ذخیره سازی.....
۲۰۸	۴-۱۳- ایمن سازی زیرساخت های ذخیره سازی در محیط های مجازی و ابری.....

بخش چهارم: بازیابی اطلاعات ۲۱۳

فصل چهاردهم: الگوی معماری بانک های اطلاعاتی و NOSQL ۲۱۴

۲۱۴	۱-۱۴- وضعیت بانک های اطلاعاتی از گذشته تا کنون.....
۲۱۶	۲-۱۴- الگوهای معماری داده NOSQL.....

فصل پانزدهم: آشنایی با سیستم های بازیابی اطلاعات ۲۴۳

۲۴۴	۱-۱۵- تعریف سیستم بازیابی اطلاعات.....
۲۴۶	۲-۱۵- اهداف سیستم های بازیابی اطلاعات.....

۲۴۶	۳-۱۵- روش ها و معیارهای ارزیابی الگوریتم های بازیابی اطلاعات
۲۴۹	۴-۱۵- جستجوی اطلاعات
۲۵۳	۵-۱۵- پردازش زبان طبیعی (NLP)
۲۵۷	۶-۱۵- متن کاوی (TEXT MINING)
۲۶۳	۷-۱۵- انواع ایندکس یا اندیس در پایگاه داده (INDEX)

فصل شانزدهم: کلان داده ها ۲۷۰

۲۷۰	۱-۱۶- خصوصیات کلان داده
۲۷۲	۲-۱۶- چالش های حوزه کلان داده
۲۷۳	۳-۱۶- فناوری های مرتبط با کلان داده
۲۸۰	۴-۱۶- یادگیری عمیق

خط مشی کیفیت انتشارات مؤسسه فرهنگی هنری دیباگران تهران در عرصه کتاب‌هایی است که بتواند خواسته‌هایی به روز جامعه فرهنگی و علمی کشور را تا حد امکان پوشش دهد. هر کتاب دیباگران تهران، یک فرصت جدید شغلی

حمد و سپاس ایزد منان را که با الطاف بی‌کران خود این توفیق را به ما ارزانی داشت تا بتوانیم در راه ارتقای دانش عمومی و فرهنگی این مرز و بوم در زمینه چاپ و نشر کتب علمی دانشگاهی، علوم پایه و به ویژه علوم کامپیوتر و انفورماتیک گام‌هایی هرچند کوچک برداشته و در انجام رسالتی که بر عهده داریم، مؤثر واقع شویم.

گسترده‌گی علوم و توسعه روزافزون آن، شرایطی را به وجود آورده که هر روز شاهد تحولات اساسی چشمگیری در سطح جهان هستیم. این گسترش و توسعه نیاز به منابع مختلف از جمله کتاب را به عنوان قدیمی‌ترین و راحت‌ترین راه دستیابی به اطلاعات و اطلاع‌رسانی، بیش از پیش روشن می‌نماید.

در این راستا، واحد انتشارات مؤسسه فرهنگی هنری دیباگران تهران با همکاری جمعی از اساتید، مؤلفان، مترجمان، متخصصان، پژوهشگران، محققان و نیز پرسنل ورزیده و ماهر در زمینه امور نشر درصدد هستند تا با تلاش‌های مستمر خود برای رفع کمبودها و نیازهای موجود، منابعی پُر بار، معتبر و با کیفیت مناسب در اختیار علاقمندان قرار دهند.

کتابی که در دست دارید با همت "جناب آقای علی احسانی" و تلاش جمعی از همکاران انتشارات میسر گشته که شایسته است از یکایک این گرامیان تشکر و قدردانی کنیم.

کارشناسی و نظارت بر محتوا: زهره قزلباش

در خاتمه ضمن سپاسگزاری از شما دانش‌پژوه گرامی درخواست می‌نماید با مراجعه به آدرس dibagaran.mft.info (ارتباط با مشتری) فرم نظرسنجی را برای کتابی که در دست دارید تکمیل و ارسال نموده، انتشارات دیباگران تهران را که جلب رضایت و وفاداری مشتریان را هدف خود می‌داند، یاری فرمایید.

امیدواریم همواره بهتر از گذشته خدمات و محصولات خود را تقدیم حضورتان نماییم.

مدیر انتشارات

مؤسسه فرهنگی هنری دیباگران تهران
bookmarket@mft.info

مقدمه مولف:

بار الها سپاس بیکران تو را، که به ما توفیق خدمتگزاری به نسل جوان و آینده ساز ایران اسلامی را عطا فرمودی و همواره چتر لطف و عنایت بی نهایت تو بر ما، سایه مهر خود را گسترده است.

پیشرفت فناوری، گسترش اینترنت و استفاده از خدمات الکترونیکی و فناوری اطلاعات منجر به تولید حجم انبوهی از داده ها و اطلاعات در دامنه ها و زمینه های گوناگون مانند پزشکی، علمی، تجارت، نجوم و غیره شده است. با توجه به این حجم گسترده و رو به رشد، نمی توان آن ها را به وسیله پایگاه داده ها و ابزارهای سنتی پردازش و تحلیل کرد و نیازمند معماری جدید برای گردآوری، انتقال، ذخیره سازی و پردازش داده ها خواهیم بود. رویکرد کتاب حاضر معطوف به بحث ذخیره و بازیابی اطلاعات خواهد بود. فعالیت‌هایی که برای تحلیل و سازماندهی مدارک و منابع صورت می‌گیرد، ذخیره اطلاعات، و تلاش‌هایی که برای یافتن یک یا چند مدرک از میان انبوه مدارک ذخیره شده انجام می‌شود، بازیابی اطلاعات نام دارد. نظام‌هایی که این جریان‌ها در آنها روی می‌دهد نظام‌های ذخیره و بازیابی اطلاعات خوانده می‌شوند. کتاب حاضر به گونه ای تالیف گردیده است که بتواند پاسخگوی نیاز تمامی کاربران چه در سطح دانشگاهی و چه در سطح کاربردی باشد. فصل بندی کتاب به شکلی صورت پذیرفته است که بیان کننده سیر رشد فناوری های ذخیره و بازیابی اطلاعات باشد. فصول ابتدایی با بیان سیستم های ذخیره اطلاعات و فایل سیستم ها آغاز می گردد و این رشد تا فصول انتهایی و بخش کلان داده ها ادامه خواهد یافت. کتاب حاضر برای دانشجویان رشته فناوری اطلاعات، مهندسی کامپیوتر و علوم کامپیوتر و تمامی علاقه مندان این حوزه چه در سطح دانشگاهی و چه سطح عمومی مفید می باشد اما مطالعه این کتاب برای خوانندگانی که با مفاهیم و رویکردهای ذخیره و بازیابی اطلاعات آشنایی داشته و قصد مطالعه این مباحث را در سطح پیشرفته دارند توصیه نمی گردد.

کتاب فوق الذکر پس از چندین سال بررسی، پژوهش و تجربیات اینجانب در زمینه ذخیره و بازیابی اطلاعات و مطالعه و تدریس درس ذخیره و بازیابی اطلاعات در دانشگاه و تلمذ از محضر اساتید بزرگی که شاید در این مختصر جای بیان تک تک اسامی این عزیزان مقدور نباشد، صورت پذیرفته است. در انتها جای دارد از مدیریت محترم انتشارات مجتمع فنی تهران و تمامی همکارانی که در گردآوری این کتاب بنده را یاری نمودند قدردانی نمایم.

امید است این اثر بتواند نقشی هر چند کوچک در خدمت به جامعه تعلیم و تربیت کشور داشته باشد.