



Information Retrieval



Musthofa Galih Pradana, M.Kom., Kharisma Wiati Gusti, S.T., M.T.,
Adi Prasetyo, S.Kom., M.Kom., Ir. Pujo Hari Saputro, S.Kom., M.T.,
Dhina Puspasari Wijaya, M.Kom., CEC., Dadang Heksaputra, S.Kom., M.Kom.,
Rendy Syahputra, S.Kom., M.Kom., Kaharuddin, S.Kom., M.Kom.,
Nurhafifah Matondang, S.Kom., MM., M.T.I., Andri Pramuntadi, M.Kom.,
Paradise S.Kom., M.Kom.

Information Retrieval

Musthofa Galih Pradana, M.Kom.

Kharisma Wiati Gusti, S.T.,M.T.

Adi Prasetyo, S.Kom.,M.Kom.

Ir. Pujo Hari Saputro, S.Kom.,M.T.

Dhina Puspasari Wijaya, M.Kom.,CEC

Dadang Heksaputra, S.Kom., M.Kom.

Rendy Syahputra, S.Kom, M.Kom.

Kaharuddin, S.Kom., M.Kom.

Nurhafifah Matondang, S.Kom.,MM.,M.T.I

Andri Pramuntadi, M.Kom.

Paradise S.Kom.,M.Kom.



Information Retrieval

Copyright © PT Penamuda Media, 2024

Penulis:

Musthofa Galih Pradana, M.Kom.
Kharisma Wiati Gusti, S.T.,M.T.
Adi Prasetyo, S.Kom.,M.Kom.
Ir. Pujo Hari Saputro, S.Kom.,M.T.
Dhina Puspasari Wijaya, M.Kom.,CEC
Dadang Heksaputra, S.Kom., M.Kom.
Rendy Syahputra, S.Kom, M.Kom.
Kaharuddin, S.Kom., M.Kom.
Nurhafifah Matondang, S.Kom.,MM.,M.T.I
Andri Pramuntadi, M.Kom.
Paradise S.Kom.,M.Kom.

ISBN:

9786238686261

Penyunting dan Penata Letak:

Tim PT Penamuda Media

Desain Sampul:

Tim PT Penamuda Media

Penerbit:

PT Penamuda Media

Redaksi:

Casa Sidoarum RT03 Ngentak, Sidoarum Godean Sleman Yogyakarta

Web: www.penamudamedia.com

E-mail: penamudamedia@gmail.com

Instagram: [@penamudamedia](https://www.instagram.com/penamudamedia)

WhatsApp: [+6285700592256](https://wa.me/6285700592256)

Cetakan Pertama, Juli 2024

viii + 176 halaman; 15 x 23 cm

Hak cipta dilindungi undang-undang

Dilarang memperbanyak maupun mengedarkan buku dalam bentuk dan
dengan cara apapun tanpa izin

Kata Pengantar

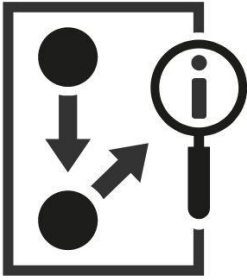
Dengan penuh antusiasme, saya mempersembahkan buku "Information Retrieval" ini kepada para pembaca yang ingin menjelajahi dunia informasi dan teknologi dengan lebih dalam. Buku ini bukan hanya sekadar panduan, melainkan juga jendela yang mengungkap kompleksitas dan keajaiban di balik sistem yang membantu kita menemukan informasi dalam hiruk-pikuk data digital. Semoga buku ini memberikan wawasan yang berharga dan memperkaya pemahaman Anda tentang informasi retrieval, sebuah bidang yang tak terhindarkan dalam dunia modern yang dipenuhi oleh data. Selamat membaca dan menikmati perjalanan pengetahuan yang menarik ini!

Daftar Isi

Daftar Isi	v
Pengantar Information retrieval	1
A. Peranan Data dan Dokumen.....	2
B. Fungsi Information retrieval	4
C. Fokus Information retrieval	6
D. Proses Information retrieval.....	7
E. Aplikasi Information retrieval.....	11
Pra Proses Teks Korpus	15
A. Pendahuluan	15
B. Tahapan.....	17
C. Manfaat Pra Proses Teks Korpus.....	26
Pengindeksan (<i>Inverted Index</i>).....	27
A. Apa itu Pengindeksan?.....	27
B. Pengertian <i>Inverted index</i>	28
C. Manfaat <i>Inverted index</i>	31
D. Struktur <i>Inverted index</i>	31
E. Membuat <i>Inverted index</i>	33
F. Pencarian Menggunakan <i>Inverted index</i>	35
G. Tantangan Penggunaan Metode <i>Inverted index</i>	39
H. Kesimpulan	43
Pembobotan Kata TF-IDF.....	44
A. Pembobotan Kata	44
B. <i>Term</i> Frekuensi (TF)	46

C. Kombinasi Term Frequency (TF) dengan Inverse Document Frequency (IDF)	52
Boolean Model	56
A. Pengantar ke Model <i>Boolean</i>	56
B. Konsep Dasar	59
C. Prinsip Kerja	65
Evaluasi Information retrieval (IR).....	71
A. Pentingnya Evaluasi dalam Information Retrieval ...	71
B. Metrik Evaluasi dalam Information Retrieval	77
Sentiment analysis.....	87
A. Pengertian Sentiment analysis	87
B. Karakteristik <i>Sentiment Analysis</i>	89
C. Tipe-tipe <i>Sentiment Analysis</i>	90
D. Mengapa Sentiment Analysis Penting?.....	91
E. Tahapan Proses <i>sentiment analysis</i>	92
Cross Language Information Retrieval (CLIR)	98
A. Definisi dan Konsep Dasar CLIR	98
B. Perkembangan CLIR.....	101
C. Pendekatan CLIR.....	102
D. Tantangan Utama CLIR.....	107
Bab 11 Question Answering System (QAS).....	111
A. Pemahaman QA <i>System</i>	111
B. Kelebihan dan Kekurangan Penerapan QAS	112
C. Contoh Penerapan QAS	114
D. Cara Kerja Sistem Pertanyaan Menjawab (QAS) Dalam Domain Tertutup.....	117

E. Peran CNN dan mekanisme self-attention dalam QAS tertutup.....	119
<i>Essay auto scoring</i>	122
A. Definisi.....	122
B. Teknologi Essay auto scoring	123
C. Metode Auto Scoring Essay.....	128
D. Proses Essay auto scoring	131
E. Evaluasi Essay auto scoring.....	132
F. Keterbatasan dan Tantangan dalam EAS.....	133
Summarization	134
A. Definisi Summarization	135
B. Pentingnya Summarization	135
C. Teknik dan Algoritma Extractive Summarization ..	138
D. Teknik dan Algoritma Abstractive Summarization	141
E. Proses Pembuatan <i>Summarization</i>	143
Daftar Pustaka.....	146
Tentang Penulis.....	159



Bab 1 Pengantar Information Retrieval

Musthofa Galih Pradana, M.Kom

Temu kembali informasi merupakan langkah penting dalam menjelajahi kompleksitas informasi di era digital. Ini tak hanya sebatas menggali konsep dasar, melainkan memahami inti dari keberadaan Temu Kembali Informasi dalam ekosistem informasi saat ini. Temu Kembali Informasi adalah lebih dari sekedar mencari data, melainkan merangkum proses esensial dalam menemukan kembali informasi yang relevan dari gudang informasi yang melimpah, sesuai kebutuhan pengguna. Konsep-konsep seperti "*query*", "dokumen", "relevansi", dan "evaluasi" menjadi fondasi pemahaman tentang bagaimana sistem Temu Kembali Informasi beroperasi (Jiafeng Guo, Yixing Fan, Qingyao Ai, 2020). Urgensi Temu Kembali Informasi dalam kehidupan sehari-hari, mulai dari dunia akademis hingga korporat. Kemampuannya untuk mempercepat akses informasi, mendukung pengambilan keputusan, serta mendorong inovasi.

Sebagai contoh, jika mencari informasi tentang mobil di dealer, pasti akan menemukan tipe mesin, fitur tambahan, atau warna mobil, namun mencari informasi tentang hari pembuatan mobil mungkin tidak akan ada, walaupun kadang-kadang hal tersebut disebutkan sebagai penilaian keandalan mobil. Disinilah proses penemuan kembali informasi berperan dalam menggali dan menemukan informasi secara lebih mendalam.

A. Peranan Data dan Dokumen

Peranan data dan dokumen saat ini sudah sangat penting dan vital keberadaannya. Hal ini dapat berimplikasi kepada beberapa hal, salah satunya adalah proses temu kembali informasi, dimana informasi yang terkandung dalam dokumen maupun data menjadi sangat penting untuk dapat ditemukan kembali makna dan hakikat di dalamnya. Berikut ini adalah alasan penting data dan dokumen dalam ranah temu kembali informasi:

1. Basis Pengetahuan dan Referensi

Data dan dokumen menyediakan basis pengetahuan yang diperlukan untuk berbagai kebutuhan penelitian, analisis, dan pengambilan keputusan. Dokumen seperti jurnal ilmiah, laporan, dan buku menyediakan informasi yang mendalam dan valid yang dapat diandalkan sebagai referensi.

Information Retrieval



Buku "Information Retrieval" membawa pembaca ke dalam dunia yang kompleks namun penting dari pengambilan informasi. Mulai dari sejarah perkembangannya hingga teknologi terkini, buku ini mengupas beragam metode, algoritma, dan konsep yang digunakan untuk mengambil informasi secara efisien dari berbagai sumber. Dengan penjelasan yang mendalam namun mudah dipahami, pembaca akan dibimbing melalui landasan teoritis hingga aplikasi praktis dalam domain yang semakin relevan dalam era digital ini.



PT Penerbit Penamuda Media
Godean, Yogyakarta
085700592256
@penamuda_media
penamuda.com