

Rekayasa Data dan Multimedia (INF-503)

Indexing dan Proses Indexing

Pengajar:

Alim Misbullah, S.Si., M.S.

Jurusan Informatika, FMIPA

Universitas Syiah Kuala

<https://elearning.unsyiah.ac.id/>

Definisi Indexing



- *Indexing* adalah suatu proses pengindeksan dokumen yang akan disediakan informasi kepada *user* untuk mempercepat proses pencarian/pemanggilan kembali dengan *query* tertentu.
- Tahapan-tahapan indexing:
 - *Parsing* document
 - Pembuangan *stopwords*
 - *Stemming*
 - *Term weighting*
 - *Inverted index*

Tugas 1

- Gunakanlah salah satu dari toolkit berikut untuk melakukan proses indexing
 - Text Indexing
 1. Swish-e (Install Package on Ubuntu)
 2. <https://github.com/apurva91/Verbose-The-Indexer>
 3. <https://github.com/nilayjain/text-search-engine>
 4. <https://github.com/RediSearch/redisearch-py>
 - Image Indexing
 1. <https://github.com/ChiyenLee/py-image>
 2. <https://github.com/meng1994412/CBIR>
 3. <https://github.com/kudeh/image-search-engine>

Tugas 1

- Langkah-langkah untuk mengerjakan tugas 1
 1. Buka account Github
 2. Clone salah satu tugas yang ingin anda kerjakan untuk text indexing dan image indexing
 3. Buat repository baru di account github anda, kemudian upload tugas anda beserta screenshot yang sudah berhasil anda jalankan
 4. Sertakan link referensi/sumber asli di dalam repository anda
 5. Waktu pengerjaan tugas selama 2 minggu
 6. Mohon didiskusikan jika ada pertanyaan melalui group WA