



UNIVERSITAS
GADJAH MADA

Arsitektur dan komponen mesin pencari

Lukman Heryawan, PhD

22 Agustus 2022

Perkuliahahan (setelah UTS)



UNIVERSITAS GADJAH MADA

- Link elok: <https://elok.ugm.ac.id/course/view.php?id=10789> (Sistem temu balik informasi)
- 7 kali pertemuan:
 - Definisi dan kategori Sistem Temu Kembali Informasi
 - Arsitektur dan komponen mesin pencari
 - Metode akuisisi data (crawling)
 - Metode transformasi data
 - Metode pembuatan indeks
 - Metode kompresi indeks
 - Rekap (demo+presentasi)
 - UTS
- Penilaian:
 - Tugas - 30%
 - UTS - 20%



UNIVERSITAS
GADJAH MADA

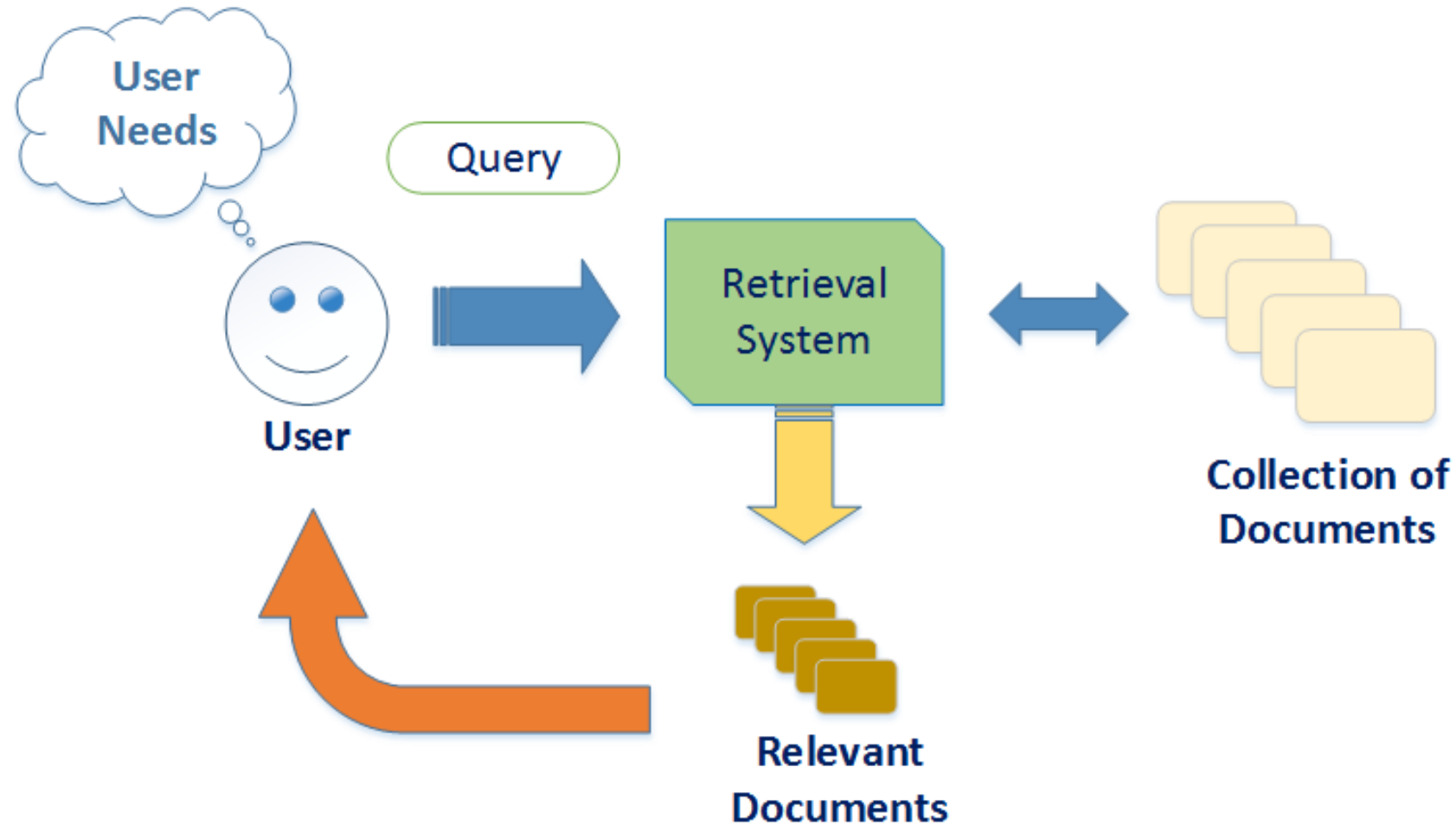
Materi

LOCALLY ROOTED, GLOBALLY RESPECTED

Information retrieval



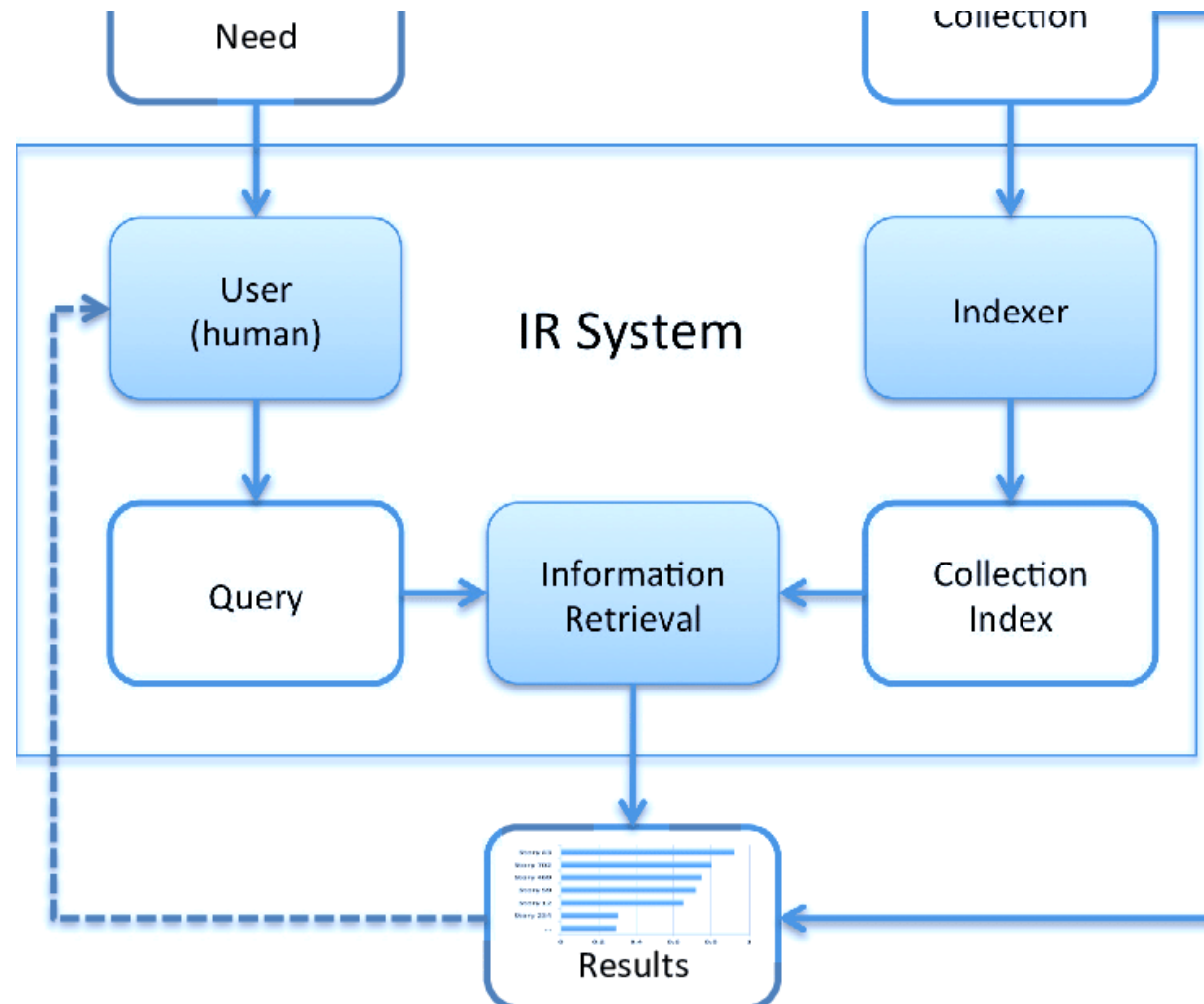
UNIVERSITAS GADJAH MADA



Architecture



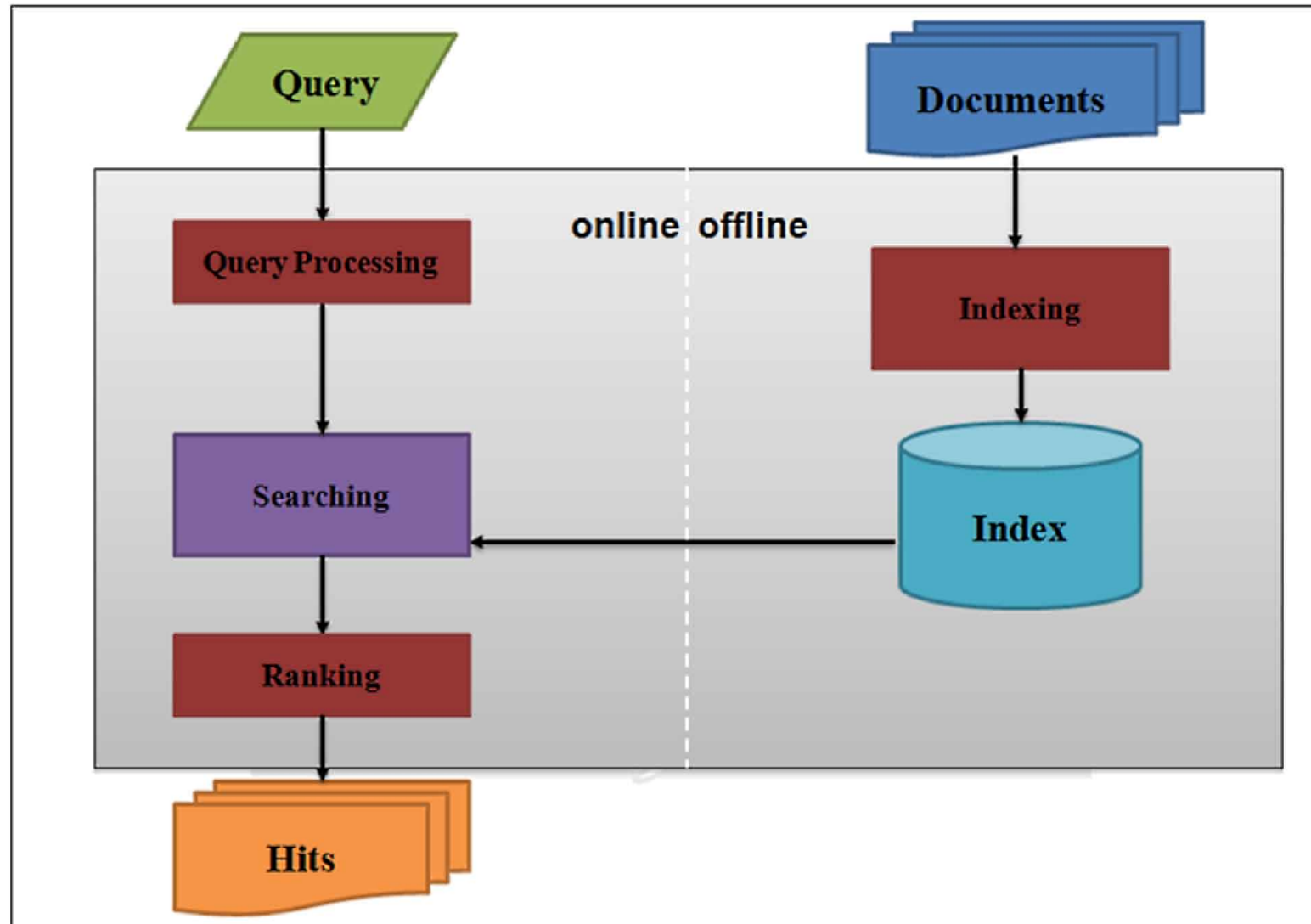
UNIVERSITAS GADJAH MADA



Architecture (cont.)



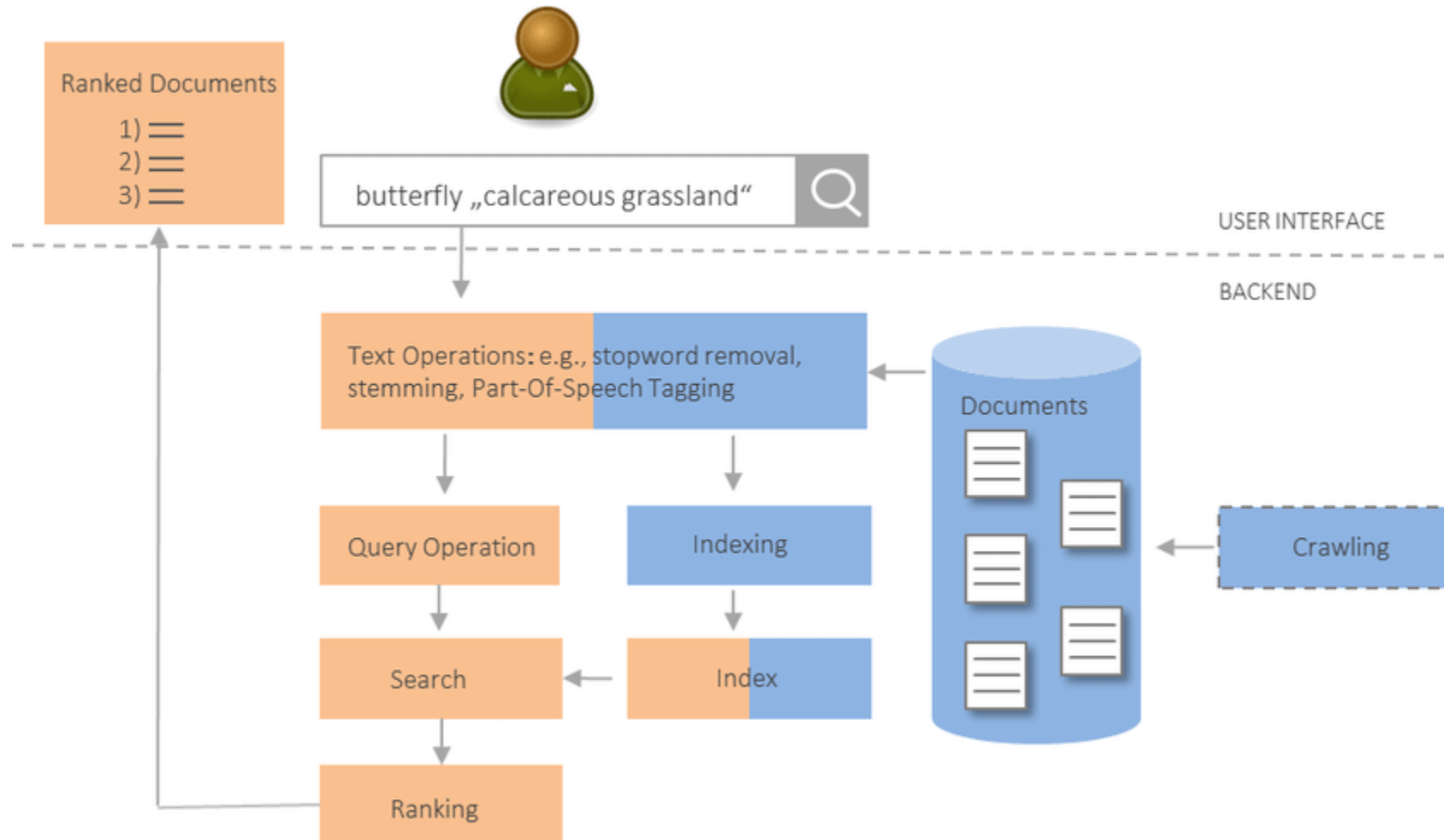
UNIVERSITAS GADJAH MADA



Architecture (cont.)



UNIVERSITAS GADJAH MADA



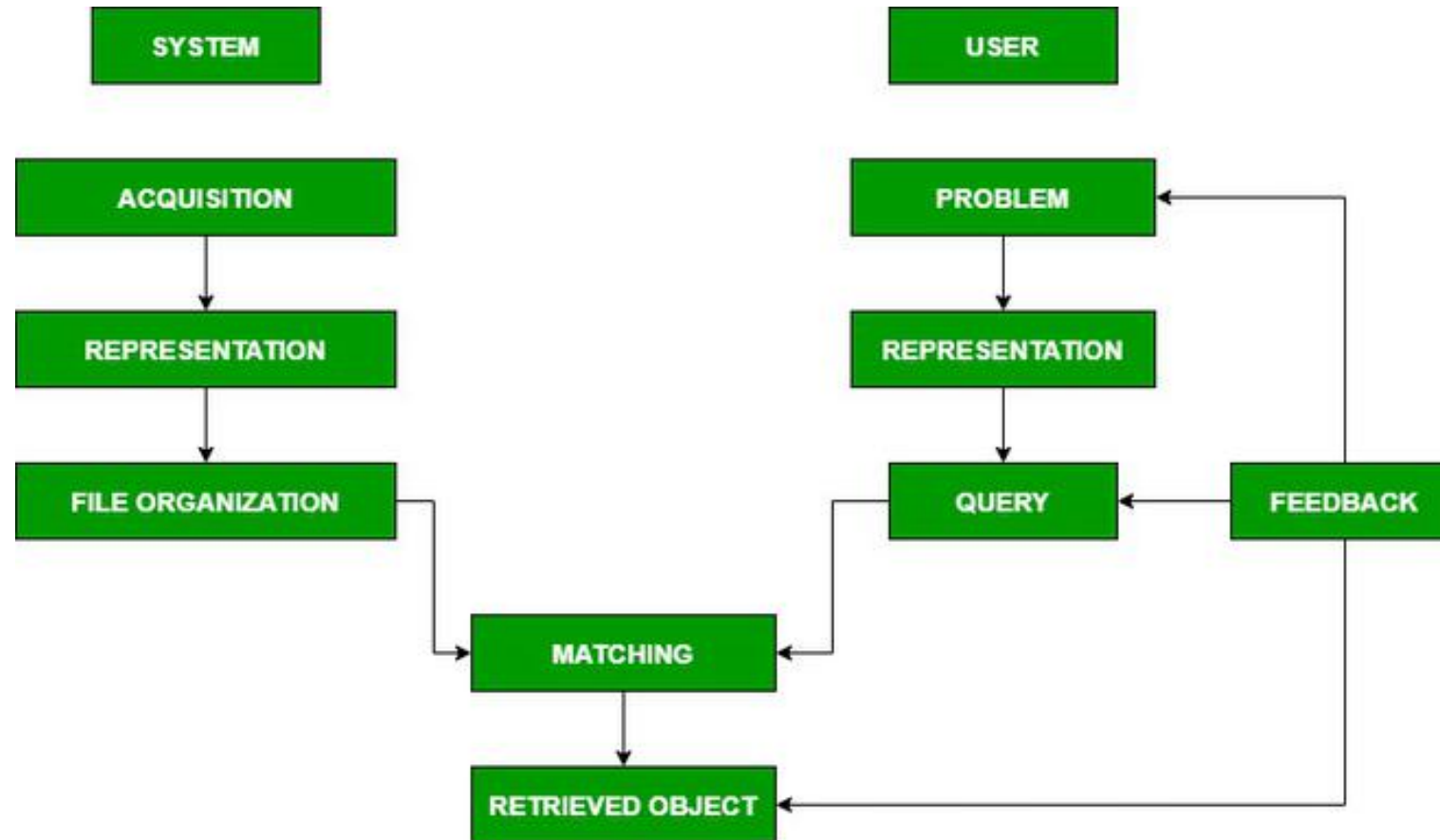
https://www.researchgate.net/figure/The-architecture-of-an-Information-Retrieval-system-based-on-Baeza-Yates-and_fig1_339551159

Components of Information Retrieval



UNIVERSITAS GADJAH MADA

- <https://www.geeksforgeeks.org/what-is-information-retrieval/>





UNIVERSITAS
GADJAH MADA

Case study

LOCALLY ROOTED, GLOBALLY RESPECTED

Tugas minggu ini (due date: 29 Agustus 2022, class presentation)



UNIVERSITAS GADJAH MADA

- Membuat **document spesifikasi STKI** yang akan dikembangkan (tugas 1)
- Membuat **document arsitektur dan komponen STKI** (tugas 2)
- Tugas berikutnya adalah:
 - **Document** metode dan hasil crawling - dataset (tugas 3)
 - **Document** metode dan hasil transformasi dataset (tugas 4)
 - **Document** metode dan hasil pembuatan indeks (tugas 5)
 - **Document** metode dan hasil kompresi indeks (tugas 6)
 - Paper yang menggabungkan tugas 1,2,3,4,5,6 + Presentasi (tugas 7)
- Tugas dapat dikerjakan secara perorangan atau kelompok (maksimal 3 orang)
 - Jika kelompok, maka semua anggota harus memiliki kontribusi dan melakukan presentasi+demo
 - Jika ada salah satu atau lebih anggota yang tidak berkontribusi dan tidak melakukan presentasi+demo, maka nilai kelompok tersebut dipotong 50%
- Document dibuat dalam file doc/docx atau editable file
- Paper ditulis menggunakan format ACM atau IEEE dan diharapkan menggunakan editor overleaf (<https://www.overleaf.com/>)
- Presentasi dibuat dalam file ppt/pptx/google slides/prezi
- Dokumentasi dikirim ke eLOK pada tugas yang bersesuaian
- Document 1,2,3,4,5,6, dataset, dan source code software diupload di github (<https://github.com/>)
- Dataset juga dapat diupload di <https://www.kaggle.com/labrpIdugm> (dapat menghubungi PIC Lab –Lukman)



Tanya jawab

- Email: lukmanh@ugm.ac.id
- Scholar profile: https://scholar.google.co.id/citations?user=V_iMAWYAAAAJ&hl=en



LOCALLY ROOTED, GLOBALLY RESPECTED