

EXAMENSARBETE Graph-Augmented RAG for Multimodal Document Question Answering**STUDENTER** Yaliang Cai, Haobo Zu**HANDLEDARE** Pierre Nugues (LTH), Xuhao Zhang (Huawei Sweden R&D)**EXAMINATOR** Xuan-Son Vu (LTH)

Ett kunskapsnätverk för bättre svar från dokument

POPULÄRVETENSKAPLIG SAMMANFATTNING **Yaliang Cai, Haobo Zu**

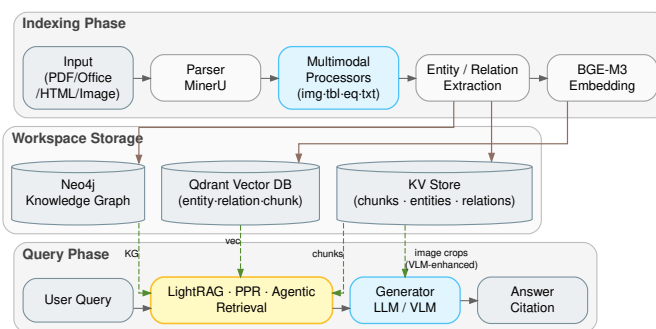
Tekniska dokument innehåller viktig information i text, tabeller och bilder. Vi har utvecklat ett lokalt AI-system som söker efter relevanta delar och kopplar samman information innan det svarar. Tester visar bättre resultat på dokumentfrågor och frågor som kräver flera informationssteg.

När man frågar om ett tekniskt dokument finns svaret inte alltid i ett enda stycke. En metod kan beskrivas i löpande text, medan det avgörande resultatet visas i en tabell, en figur eller en formel längre bak i dokumentet. Detta är vanligt i forskningsartiklar, tekniska rapporter och manualer, där olika delar av dokumentet kompletterar varandra. Ett system som bara söker efter liknande formuleringar kan därför hitta rätt ämne, men missa just det underlag som behövs för att svara korrekt.

I vårt examensarbete har vi utvecklat ett lokalt AI-system som organiserar dokumentens innehåll i en kunskapsgraf. I grafen kopplas begrepp och relationer till de textstycken, tabeller, bilder och formler där informationen finns, så att samband mellan olika delar av dokumentet blir synliga.

När en användare ställer en fråga kombineras vanlig textsökning med sökning genom grafens kopplingar. På så sätt kan systemet hitta stöd som ligger på olika platser eller nås genom en kedja av samband, snarare än bara genom ytlig ordlikhet.

Det skapar en automatiserad bro mellan textuell och visuell information, vilket ger AI-systemet en mer komplett bild av hela sammanhanget.



En extra kontrollkomponent granskar om svaret verkligen stöds av källorna, vilket minskar risken för att systemet hittar på information. All bearbetning sker lokalt, så känsliga eller affärskritiska dokument aldrig behöver skickas till externa AI-tjänster.

Vi testade systemet på akademiska dokument och på flerstegsfrågor där svaret kräver att flera fakta vägs samman. Resultaten visar att grafbaserad sökning kan ge bättre svar än traditionella metoder när informationen är spridd över ett dokument. Vid bred sökning över stora samlingar av forskningsartiklar var metoden däremot fortfarande begränsad, eftersom grafer är dyrare att bygga och söka i än enkla textindex. Arbetet visar därför både möjligheterna och gränserna för kunskapsgrafer i AI-stöd.