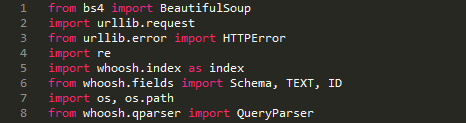
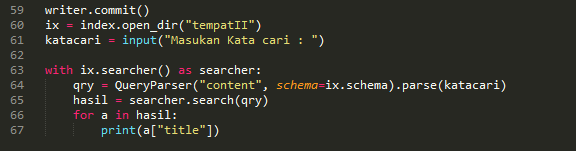
Timothius Tirtawan / 215116485

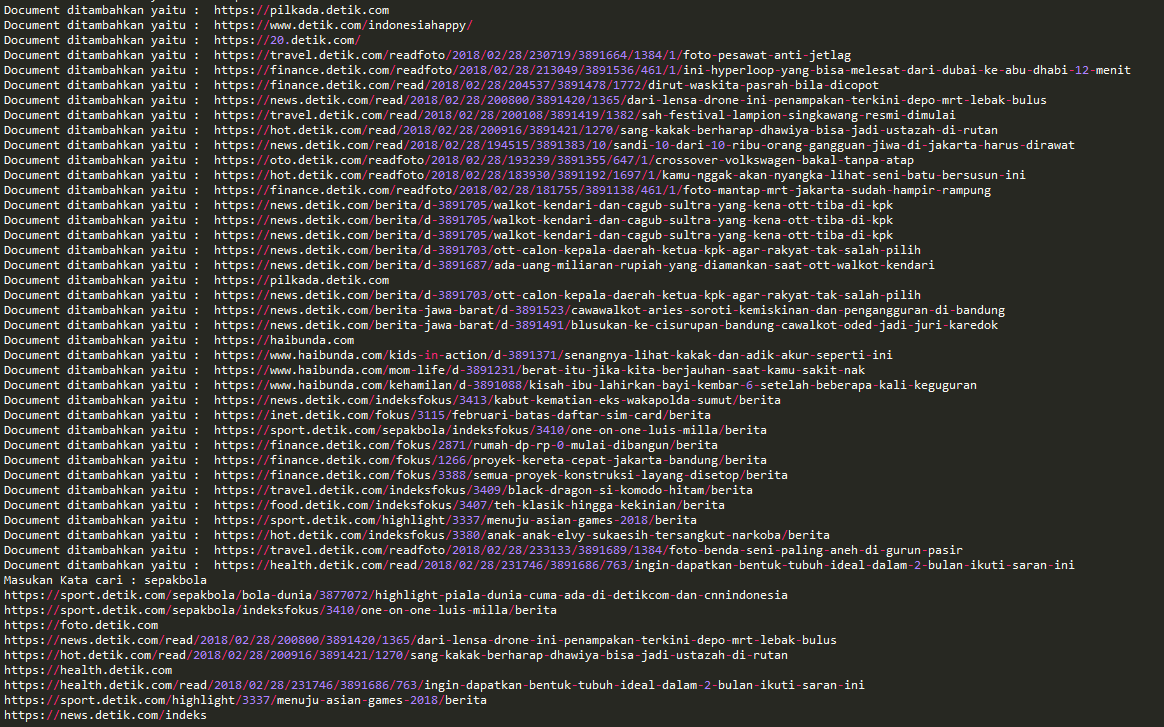
Untuk melakukan inverted index, Python menyediakan library yang cukup besar yaitu Whoosh. Whoosh sendiri mencakup untuk indexing, inverted indexing, searching. Library-library yang membantu untuk inverted index ialah : 

Beautiful soup untuk mengambil isi content pada web yang akan dituju, urllib untuk membuka web tersebut, error nya untuk try catch agar program tidak berhenti pada saat proses fetching. Re untuk regex, whoosh.index untuk indexing inverted. Whoosh field untuk membuat schema lalu os untuk membuat folder yang ada pada file python tersebut. Qparser berguna untuk metode searching. Pada awal program harus dibuat schema terlebih dahulu untuk menentukan struktur penyimpanan inverted index. schema dibuat dengan struktur title sebagai id dan content sebagai isi web text. Untuk fetching web dan crawl judul web menggunakan simple crawler. Web root yang akan di fetching adalah <http://detik.com>. Lalu membuat folder untuk menempatkan hasil inverted index. Pada folder tersebut memiliki writer. Lalu untuk setiap web yang ada pada Eku akan disimpan dengan penulisan writer pada folder “tempatII”.



Setelah membuat inverted indexnya, untuk mencari kata yang ada pada Whoosh menggunakan yang namanya QueryParser dengan memasukan struktur apa yang mau di cari (Pada kasus ini bisa ID / title , Content / isi) setelah proses fetching maka user akan diminta inputan berupa string untuk pencarian kata. Jika kata berhasil ditemukan maka akan ditampilkan link yang menuju pada web yang memiliki kata tersebut.

Eksekusi program



Hasil yang didapatkan dengan mencari kata “sepakbola” adalah seperti yang terlihat pada gambar tersebut.