

Introduction to Information Retrieval

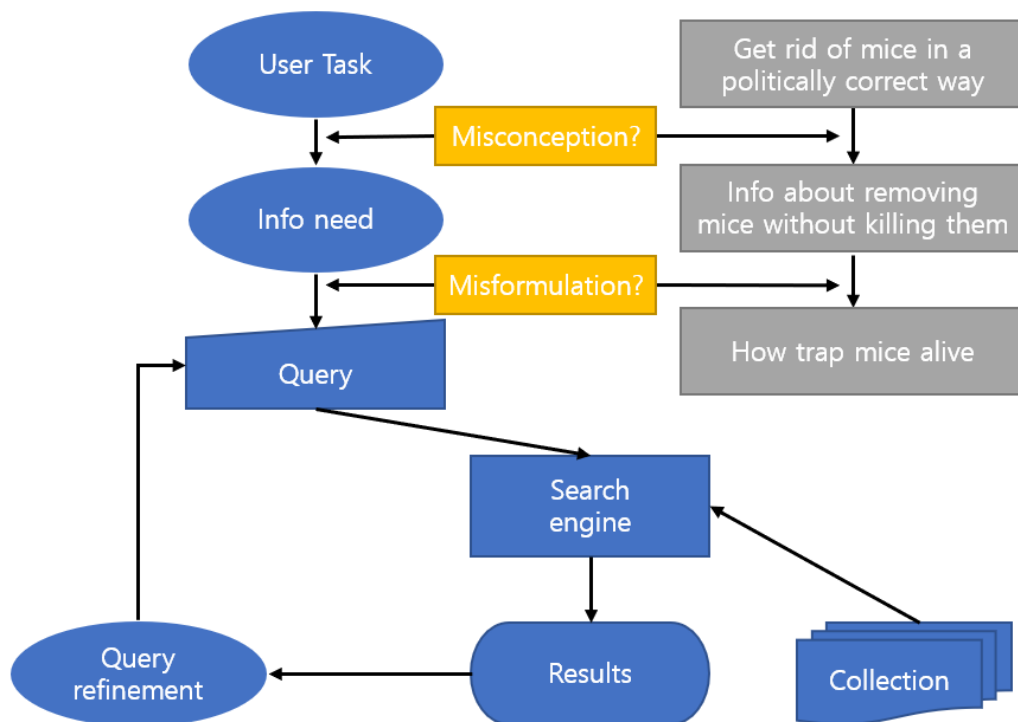
Information Retrieval

Information Retrieval (IR) : finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers)

Basic assumptions of Information Retrieval

- Collection : A set of documents
- Goal of IR : Retrieve documents with information that is **relevant** to user's **information need** and helps the user complete a **task**

The classic search model



classic search model 설명 : info need를 위한 적절한 query를 search engine에 날리면 Collection에서 query와 연관이 있는 정보 즉, 문서를 output으로 제공하여 results를 만들어 낸다. Results가 마음에 들지 않는다면 query를 다시 정제(query refinement)하여 새로운 query를 search engine에 날린다.

misconception은 사용자가 행하고 싶은 task(user task)와 사용자가 행하고 싶은 task를 하기 위해 필요한 정보(info need)가 일치하는지를 의미한다.

misformulation은 필요한 정보를 찾기위해 query를 날리는데 이때 query의 단어를 잘 못 선택한 것을 의미한다.

ex)

user task : get rid of mice in a politically correct way

info need : info about removing mice without killing them

query : how trap mice alive

misconception: not **whether or not i kill them** but **whether i do it humanely**

misformulation : how trap mice alive → how trap mice without killing

출처 : https://www.youtube.com/watch?v=kNkCfaH2rxc&list=PLaZQkZp6WhWwoDuD6pQCmgVy.DbUWI_ZUj