

캡스톤 디자인(1) 중간 발표

Web Crawling, NLP, WordCloud를 이용한
기사 통합 검색 시스템

목차

1. 프로젝트 소개

- ❖ 개요, 기능, 아키텍처

2. 진행 상황

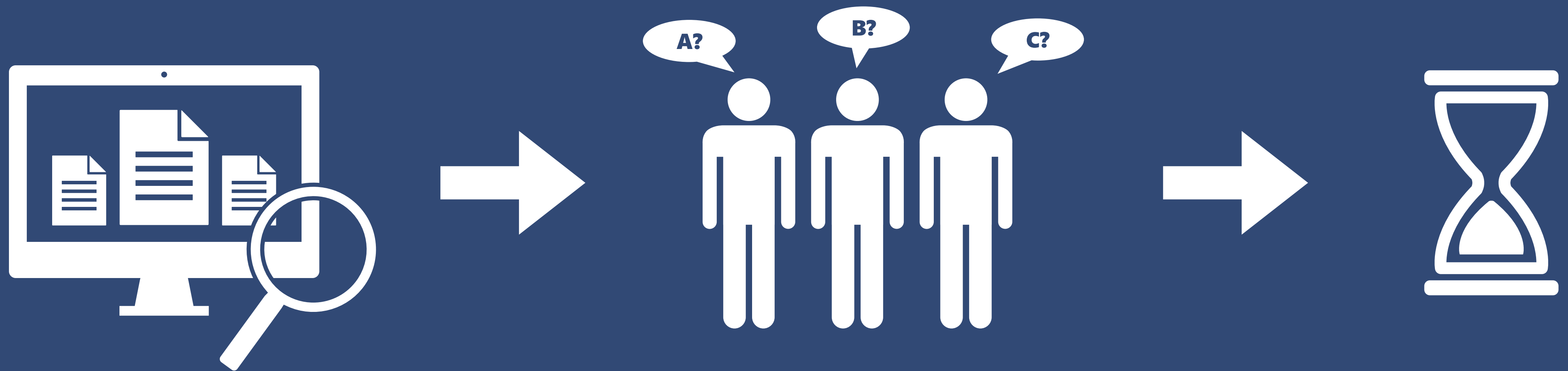
- ❖ 진행 과정, 학습 내용

3. 향후 계획

- ❖ 최종 발표까지 예상 결과, 향후 진행 계획

1. 프로젝트 소개

프로젝트 개요



프로젝트 기능

뉴스 매체 선택

단어의 유의어, 상위어, 하위어 포함 여부 선택

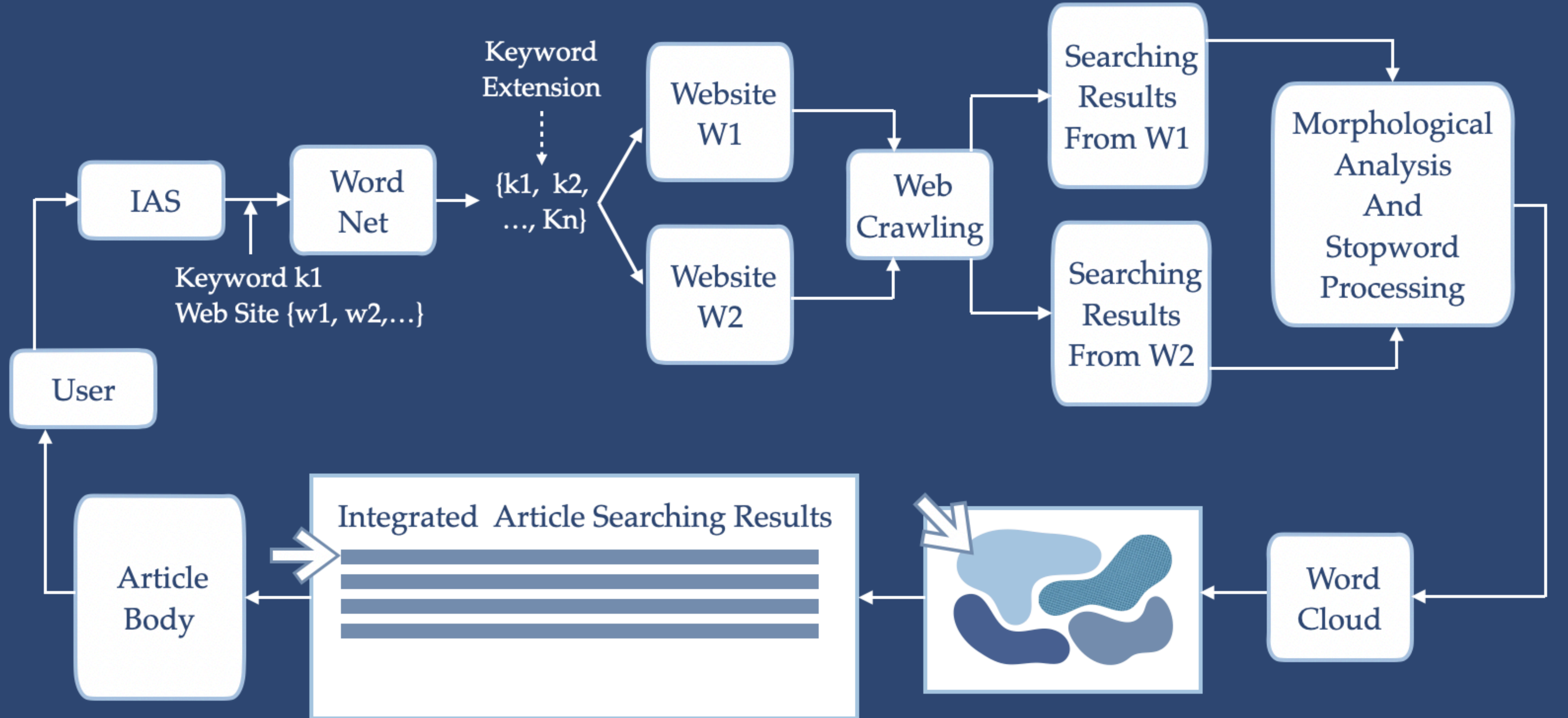
단어 검색

단어 검색

해당 단어가 포함된 뉴스들의 정보로 wordcloud 생성

wordcloud의 단어 클릭 시 해당 단어를 포함한 뉴스들의 링크 제공

프로젝트 아키텍처



2. 진행 상황

진행 과정

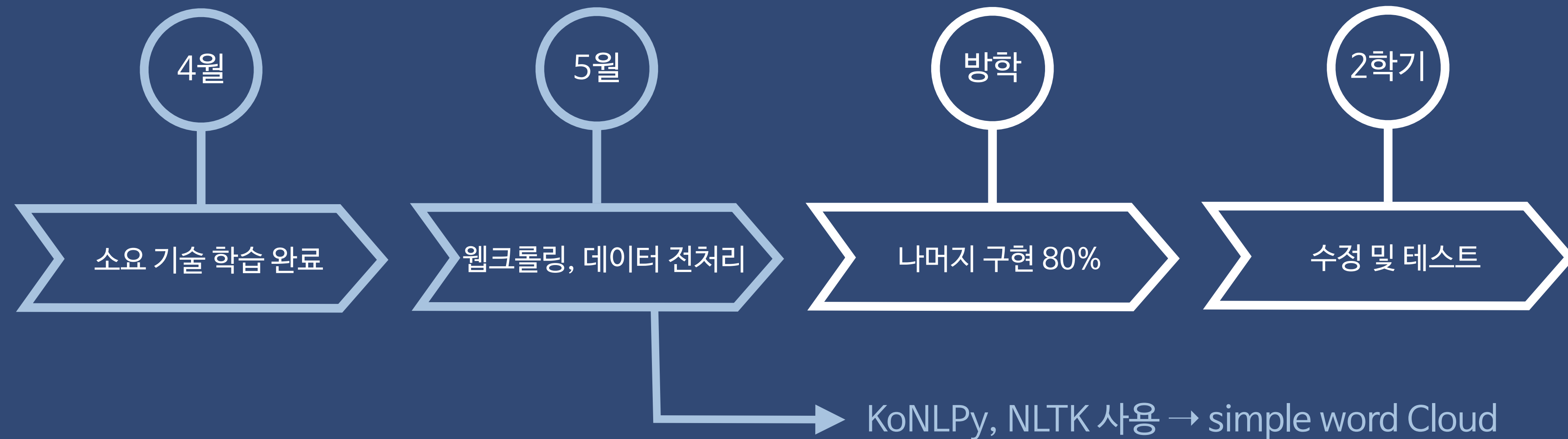


소요 기술, 소요 자원 검토 및 학습

- ❖ Python, selenium, beautifulsoup4, KoNLPy, NLTK, word cloud
 - ❖ Selenium, beautifulsoup4 학습
 - ❖ KoNLPy, NLTK 활용
 - ❖ Word cloud 분석 → interactive word cloud

3. 향후 계획

최종 발표까지 예상 결과



향후 진행 계획

