

JERY's WORD

(유사도 추정 인공신경망을 이용한
기사 검색, 댓글 시각화 서비스)

설계서

2020년 4월 30일

담당교수 권기학 교수님, 문현준 교수님
소 속 세종대학교 컴퓨터공학과
조 이 름 JERY(6조)
조 원 16011002 박보람 (조장)
16010941 최유진
16011022 손모은
16011080 이주연

- 목차 -

1. 분석단계.....	3
1.1. 사용자요구사항정의서.....	3
1.2. 유스케이스명세서	6
2. 설계단계.....	17
2.1 클래스 설계서.....	17
2.2 유사도 추정 설계서	28
2.3 사용자 인터페이스 설계서	29
2.4 컴포넌트 설계서.....	35
2.5 인터페이스 설계서.....	41
2.6 아키텍처 설계서.....	47
2.7 총괄시험 계획서.....	49

1. 분석단계

1.1. 사용자요구사항 정의서

■ 제·개정 이력

날짜	버전	작성자	승인자	내용

R1	사용자 요구사항 정의서				
시스템명	JERY's WORD	서브시스템명	뉴스 검색 시스템, 댓글 시각화 시스템, 기간 설정 시스템, 모델 학습 시스템		
단계 명	분석	작성일자	2020. 05. 02	버전	1.0

서브시스템 명			뉴스 기사 유사도 처리						
요구 사항 ID	요구사항 명	구분	요구사항 설명	요구사항 출처	제약 사항	중요도	해결방안	검수기준	비고
JW_01	연관도 높은 뉴스 검색	기능 성능	사용자는 입력한 키워드와 연관이 높은 뉴스기사 리스트를 언론사 구분 없이 유사도 순으로 받아볼 수 있어야 한다.	과제요 구사항 명세서. pdf 1-가	유사도가 측정 가능한 키워드로 검색 / 검색 결과가 존재 해야 함	상	뉴스용어 분석에 최적화 되도록 학습시킨 인공신경망 모델을 통해 키워드 유사도 향상	키워드와 기사 간의 실제 연관/유사도 확인 / 유사도 순으로 정렬된 기사 리스트 확인	

JW_02	뉴스 유사도 확인	기능 성능	사용자는 검색을 통해 받아본 기사들과 실제로 입력한 키워드 간의 유사도(%)를 확인할 수 있어야 한다.	과제요 구사항 명세서. pdf 1-가	유사도가 측정 되어야 함	상	학습모델을 통해 도출된 유사도 (%)를 웹페이지에 디스플레이	유사도 측정 결과로 도출된 값의 정확도 확인	
JW_03	검색 속도	성능	사용자는 검색결과를 5초 이내에 받아볼 수 있어야 한다.	과제요 구사항 명세서. pdf 3-나-1	-	중 상	최적의 알고리즘 구축 / 성능을 내기예	검색결과 디스플레이 속도 ≤5초	
JW_04	기사 원문 리다이 렉트	기능	사용자가 선별된 리스트의 목록 중 특정 기사를 클릭했을 시 해당 기사의 원문을 보여줄 수 있어야 한다.	과제요 구사항 명세서. pdf 1-가	-	하	새로운 브라우저 창을 열어 기사 사이트로 이동	-	

서브시스템 명		댓글 시각화 처리							
요구 사항 ID	요구사항 항목	구분	요구사항 설명	요구사항 출처	제약 사항	중요도	해결방안	검수기준	비고
JW_11	기사 댓글 시각화	기능 성능	검색결과로 나온 기사 리스트의 댓글을 추합하여 빈도수에 따라 시각화(워드 클라우딩)한 이미지를 3 개 디스플레이 할 수 있어야 한다.	과제요구 사항명세서.pdf 1-가	검색된 기사의 댓글이 존재 해야 함	상	파이썬 오픈소스를 이용하여 댓글 클라우딩 자동화	검색된 기사내의 댓글로 형성됐는지 / 의미있는단어 기준으로 생성되었는 지 여부	
JW_12	댓글 클라우드 관련 기사 탐색	기능	사용자가 특정 기간의 댓글 클라우드 이미지를 클릭했을 시 해당 기간 동안의 기사 목록을 확인할 수 있어야 한다.	과제요구 사항명세서.pdf 1-가	-	상	해당 기간 동안의 기사 리스트 디스플레이	워드 클라우딩으로 보여지는 댓글 반응과 실제 기사 댓글 반응의 부합 여부 확인	
JW_13	댓글 클라우드 디스플레이 속도	성능	사용자는 시각화된 댓글을 10 초 이내에 받아볼 수 있어야 한다.			중상	성능을 내기에 충분한 HW 확보	댓글 클라우드 디스 플레이 속도 ≤10 초	

서브시스템 명		기사 기간 설정							
요구 사항 ID	요구사항 명	구 분	요구사항 설명	요구사항 출처	제약 사항	중 요 도	해결방안	검수기준	비 고
JW_21	검색결과기간 설정	기 능	사용자는 기간 설정을 통해 해당 기간 동안 작성된 기사 목록과 댓글 클라우드를 받아볼 수 있어야 한다.	과제요 구사항 명세서. pdf 1-가	초기값은 3 일로 이는 사용자가 임의로 지정할 수 없음 / 개발자 측에서 콤보 박스 형태로 지정해놓은 날짜 선택	상	기간에 해당 되는 기사 목록과 워드 클라우드 디스 플레이	사용자가설정 한 기간과 기사 작성날짜의 부합 여부 확인	

1.2. 유스케이스 명세서

■ 제·개정 이력

날짜	버전	작성자	승인자	내용

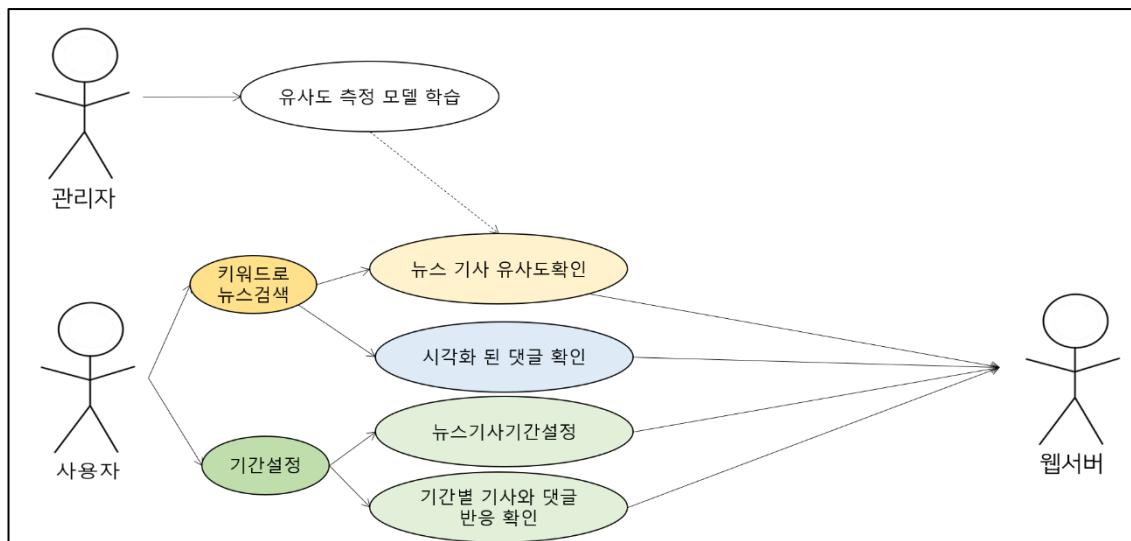
R2	유스케이스 명세서			
시스템명	JERY's WORD	서브시스템명	뉴스 검색 시스템, 댓글 시각화 시스템, 기간 설정 시스템, 모델 학습 시스템	
단계 명	분석	작성일자	2020. 05. 02	버전 1.0

1) 서브시스템 목록

서브시스템 ID	서브시스템명	서브시스템 설명
JW_ASM_010	뉴스 검색 시스템	유사도를 기준으로 뉴스 기사 검색
JW_CVM_020	댓글 시각화 시스템	기사 내 댓글 시각화
JW_DSM_030	기간 설정 시스템	검색 기간 설정
JW_MLM_030	모델 학습 시스템	유사도 측정

2) 유스케이스 다이어그램(UCD)

- 전체 다이어그램



UCD ID	JW_UCD_U01	UCD 명	뉴스 검색
관련 서브시스템 ID	JW_ASM_010, JW_CVM_020	관련 서브시스템명	뉴스 검색 시스템, 댓글 시각화 시스템


```

graph TD
    User((사용자)) --> JW_ASM_010_01([JW_ASM_010-01 키워드로 뉴스 검색])
    JW_ASM_010_01 --> JW_ASM_010_02([JW_ASM_010-02 뉴스 기사 유사도 확인])
    JW_ASM_010_02 --> WebServer((웹서버))
    User --> JW_CVM_020_01([JW_CVM_020-01 시각화 된 댓글 확인])
    JW_CVM_020_01 --> WebServer

```

This detailed diagram for '뉴스 검색' (News Search) shows the specific use cases: 'JW_ASM_010-01 키워드로 뉴스 검색' (Search news by keyword) and 'JW_CVM_020-01 시각화 된 댓글 확인' (Confirm visualized comments). Both of these lead to the 'JW_ASM_010-02 뉴스 기사 유사도 확인' (Confirm news article user similarity) step, which then interacts with the '웹서버' (Web Server).

UCD ID	JW_UCD_U02	UCD 명	뉴스 기간 설정
관련 서브시스템 ID	JW_DSM_020	관련 서브시스템명	기간 설정 시스템
<pre> graph LR User((User)) --> DSM01((JW_DSM_030-01 뉴스 기사 기간 설정)) User --> DSM02((JW_DSM_030-02 기간 별 기사와 댓글반응 확인)) DSM01 --> WebServer((Web Server)) DSM02 --> WebServer </pre>			

UCD ID	JW_UCD_U03	UCD 명	유사도 측정 모델 학습
관련 서브시스템 ID	JW_MLM_030	관련 서브시스템명	모델 학습 시스템
<pre> graph LR Manager((Manager)) --> MLM01((JW_MLM_040-01 유사도 측정 모델 학습)) MLM01 --> WebServer((Web Server)) </pre>			

3) 유스케이스 목록

유스케이스 ID	유스케이스 명	유스케이스 설명	관련액터 ID	관련 UCD ID	관련 요구사항 ID
JW_ASM_010-01	뉴스 검색	키워드를 이용하여 뉴스를 검색한다.	JW_AC_010 JW_AC_020	JW_UCD_U01	JW_01
JW_ASM_010-02	뉴스 기사 유사도 확인	검색을 통해 나온 뉴스리스트의 각 기사별로 키워드와의 유사도를 출력한다.	JW_AC_010 JW_AC_020	JW_UCD_U01	JW_02

JW_CVM_020-01	시각화된 댓글 확인	키워드 검색으로 도출된 기사의 댓글을 분석하여 워드 클라우드 이미지를 출력한다.	JW_AC_010 JW_AC_020	JW_UCD_U01	JW_11
JW_DSM_030-01	뉴스 기사 기간 설정	기사가 작성된 기간을 조정한다. 드롭 다운 박스 내 9~7 일전 / 6~4 일전 / 3~0 일전 목록 중 하나의 값을 선택한다.	JW_AC_010 JW_AC_020	JW_UCD_U02	JW_21
JW_DSM_030-02	기간 별 기사와 댓글 반응 확인	설정된 기간에 따른 기사 리스트와 댓글 반응을 출력한다.	JW_AC_010 JW_AC_020	JW_UCD_U02	
JW_MLM_040-01	유사도 측정 모델 학습	기존의 인공 신경망 모델에 뉴스 기사를 학습시켜 뉴스 용어 분석에 최적화 시킨 모델을 구현한다.	JW_AC_030 JW_AC_040	JW_UCD_U03	

4) 액터 목록

액터 ID	액터명	액터유형	액터설명
JW_AC_010	사용자	주요	웹을 통해 실 서비스를 사용
JW_AC_020	웹 서버	보조	사용자 인터페이스와 사용자의 키워드를 통해 데이터 수집 및 댓글 시각화 담당
JW_AC_030	관리자	주요	유사도 측정 모델 학습 간접 담당
JW_AC_040	유사도 측정 모델	보조	사용자의 검색 키워드를 받아 뉴스 기사와의 유사도 측정 및 키워드와 유사한 단어 도출

5) 유스케이스 기술서

유스케이스 ID	JW_ASM_010-01	유스케이스명	뉴스 검색
<hr/>			
1. 주요 액터			
<ul style="list-style-type: none">- 사용자			
<hr/>			
2. 이해관계자와 관심사항			
<ul style="list-style-type: none">- 사용자는 키워드 입력으로 높은 연관성의 기사를 찾길 원한다.- 기사는 언론사에 관계없이 유사/연관도 순으로 검색되어야 한다.			
<hr/>			
3. 전제조건			
<ul style="list-style-type: none">- 사용자가 키워드를 입력한 상태여야 한다.			
<hr/>			
4. 종료조건			
<ul style="list-style-type: none">- 키워드에 맞는 추천 기사리스트를 반환한다.			
<hr/>			
5. 기본 시나리오			
<ol style="list-style-type: none">1) 사용자가 web browser를 사용해 JERY's WORD 기본 홈페이지에 접속한다.2) 사용자가 검색창에 알고 싶은 뉴스 기사의 키워드를 입력한다.3) 사용자가 엔터키를 누르거나 검색아이콘을 클릭한다.4) 웹서버는 메인 검색결과 화면을 디스플레이 한다.5) 유스케이스 뉴스 기사 유사도 확인(JW_ASM_010-02) 과 시각화된 댓글 확인(JW_CVM_020-01)이 실행된다.6) 본 유스케이스가 종료된다.			
<hr/>			
6. 대안 시나리오			
<ol style="list-style-type: none">1) 키워드가 단어가 아닌 문장부호가 들어간다면 오류 메시지를 디스플레이 한다. (단 띄어쓰기는 제외한다.) 또는 빈칸에서 엔터키를 누르거나 검색을 누르면 오류 메시지를 디스플레이 한다.			

2) 사용자가 오류 메시지를 인식하고 유스케이스를 마친다.

7. 구현 시 고려사항

- 검색 키워드는 한글 단어로 제한하며 단어 개수에는 제한이 없다.
- 두 개 이상의 단어를 키워드로 검색할 경우 띄어쓰기로 구분한다.

8. 발생 빈도

- 일 1회 이상 지속적으로 발생한다.

유스케이스 ID	JW_ASM_010-02	유스케이스명	뉴스 기사 유사도 확인
----------	---------------	--------	--------------

1. 주요 액터

- 사용자

2. 이해관계자와 관심사항

- 사용자는 검색된 기사의 유사도를 확인하길 원한다.

3. 전제조건

- 키워드와 연관된 기사 목록이 10개 이상 존재해야 한다.
- 유스케이스 JW_ASM_010-01이 실행된 이후여야 한다.

4. 종료조건

- 각 기사마다 측정된 유사도 퍼센티지(num)를 반환한다.

5. 기본 시나리오

- 1) 사용자가 입력한 키워드와 자사의 모델을 사용하여 유사한 키워드 10개를 추출한다.
- 2) 기존 키워드와 추출한 키워드를 종합하여 뉴스를 검색하여 총 100개의 기사를 간추린다.

- 3) 그 중 키워드와의 기사제목 유사도를 측정하여 상위 10 개의 기사를 추려낸다.
- 4) 최상단 검색창은 기존 키워드값 출력을 유지한다.
- 5) 페이지 상부에 10 개의 기사를 위에서 아래로 디스플레이 한다.
- 6) 기사 오른쪽에 유사도도 함께 디스플레이 한다.
- 7) 기사 아래에 작게 출판사와 발행일을 출력한다.
- 8) 기사 제목을 누르면 해당 기사 원문 페이지로 이동할 수 있다.
- 9) 본 유스케이스가 종료된다.

6. 대안 시나리오

- 1) 유사도가 측정되지 않는 단어는 오류 메시지를 디스플레이 한다.
- 2) 기사가 없는 경우 오류 메세지를 디스플레이 한다.
- 3) 사용자가 오류 메시지를 인식하고 유스케이스를 마친다.

7. 구현 시 고려사항

- 키워드가 입력된 날짜를 기준으로 현재부터 9 일 전까지의 기사 목록을 크롤링 해온 후 이를 유사도 측정 데이터로 활용한다.
- 기사는 유사도가 높은 순으로 정렬되어 보여야 한다.

8. 발생 빈도

- 뉴스 검색이 이루어질 때마다 발생한다. (일 1 회 이상 지속적)

유스케이스 ID	JW_CVM_020-01	유스케이스명	시각화 된 댓글 확인
----------	---------------	--------	-------------

1. 주요 액터

- 사용자

2. 이해관계자와 관심사항

- 사용자는 검색된 기사의 댓글 반응을 한 눈에 알아보길 원한다.
- 시간에 따라 댓글 반응이 어떻게 변화하는 지도 살펴보길 원한다.

3. 전제조건

- 키워드로 검색을 해야 한다.
- 즉, 유스케이스 JW_ASM_010-01 이 실행된 이후여야 한다.
- 검색된 기사에 댓글이 충분히 존재하는 경우를 고려한다.

4. 종료조건

- 시각화 된 댓글 이미지 파일(.png)을 반환한다.

5. 기본 시나리오

- 1) 현재 시점에서 0~3 일전, 4~6 일전, 7~9 일전 기사끼리 그룹화되어 이뤄진다.
- 2) 워드 클라우드 이미지는 각 기간별로 1 개씩 형성된다.
- 3) 기사의 댓글의 단어 빈도수를 이용하여 시각화 한 이미지가 형성된다.
- 4) 형성된 이미지를 페이지 하단 댓글 시각화 레이아웃 구간에 디스플레이 한다.
- 5) 각 이미지를 누르면 해당 기간 기사 조회 페이지로 이동가능하다(유스케이스 JW_DSM_030-01로 이어진다.)
- 6) 본 유스케이스가 종료된다.

6. 대안 시나리오

- 1) 댓글이 없는 경우 기본(오류표시) 이미지를 디스플레이 한다.

7. 구현 시 고려사항

- 댓글 클라우드는 검색 결과로 나온 기사의 댓글들로 구성되어야 한다.
- 의미 있는 단어(명사, 형용사 등)를 제외한 불필요한 형태소는 댓글 클라우딩 과정에서 제외한다.

8. 발생 빈도

- 뉴스 검색이 이루어질 때마다 발생한다. (일 1회 이상 지속적)

유스케이스 ID	JW_DSM_030-01	유스케이스명	기간 설정
----------	---------------	--------	-------

1. 주요 액터

- 사용자

2. 이해관계자와 관심사항

- 사용자는 특정 기간의 기사와 댓글 반응을 얻어 보길 원한다.

3. 전제조건

- 사용자가 댓글 반응 클라우드 이미지 중 한 가지를 클릭했을 때 유스케이스가 실행된다.

4. 종료조건

- 설정된 기간에 적합한 새로운 기사리스트과 댓글 클라우드 이미지를 반환한다.

5. 기본 시나리오

- 1) 기사와 댓글 클라우드가 집계되는 기간의 초기값은 3 일이다.
- 2) 사용자가 좌측 상단의 날짜 선택 버튼을 누른다.
- 3) 날짜 선택 버튼 밑으로 다음과 같은 세 가지 항목이 보여진다.
0~3 일간/4~6 일간/7~9 일간
- 4) 뉴스 검색(JW_ASM_010-01)을 통해 수집한 데이터 내에서 사용자가 선택한 기간에 적합한 기사 리스트를 얻어 댓글 클라우드 이미지와 함께 디스플레이 한다.
- 5) 본 유스케이스가 종료된다.

6. 대안 시나리오

- 1) 유사도가 측정되지 않는 단어는 오류 메시지를 디스플레이 한다.
- 2) 기사가 없는 경우 오류 메세지를 디스플레이 한다.
- 3) 댓글이 없는 경우 기본(오류표시) 이미지를 디스플레이 한다.

7. 구현 시 고려사항

- 초기값은 3 일로 이는 사용자가 임의로 지정할 수 없다.
- 기간 선택은 드롭다운 박스 형태로 9~7 일전/6~4 일전/3~0 일전 3 가지 목록으로 구현한다.

8. 발생 빈도

- 일 1 회 이상 지속적으로 발생한다.

유스케이스 ID	JW_MLM_040-01	유스케이스명	유사도 측정 모델 학습
1. 주요 액터			
	- 관리자		
2. 이해관계자와 관심사항			
	- 관리자는 키워드의 유사도 측정에 사용되는 모델을 학습시키고자 한다.		
	- 관리자는 새로운 이슈에도 적용될 수 있도록 정기적으로 기존 모델을 업데이트 시키는 학습을 시키고자 한다.		
3. 전제조건			
	- 크롤링이 가능한 언론사 사이트의 분야별 url 이 start_urls 에 설정되어 있어야 한다.		
	- 모델의 상용화가 가능한 성능이 나오도록 충분한 양의 기사를 크롤링하여 학습시킨다.		
4. 종료조건			

- 최초 학습된 / 업데이트 학습된 모델이 저장 된다.

5. 기본 시나리오

- 1) 관리자가 관련 파이썬 파일을 실행시킨다.
- 2) 크롤링이 수행되고 학습할 데이터가 저장된다.
- 3) 데이터를 사용하여 모델을 학습시키고 저장한다.
- 4) 본 유스케이스가 종료된다.

6. 대안 시나리오

- 1) 추후에 업데이트를 시킬 때 연합뉴스 홈페이지가 크롤링을 금지한다면 대안 사이트를 찾아 사용한다.
- 2) 데이터 크기나 하드웨어 에러 등으로 모델 학습/업데이트가 실패한다면 기존 모델을 유지한다.

7. 구현 시 고려사항

- 분야별로 지원하지 않는 주제에 관해서는 제외하고 학습시킬 수 있다.(예 : 연예)
- 기사의 제목과 본문 텍스트 데이터 중 명사만 학습시킨다.
- 학습시킬 기사 양은 3000 개 이상을 기본으로 한다.

8. 발생 빈도

- 서비스 시작 전 1 회와 주 1 회정도 업데이트 한다.

2. 설계단계

2.1 클래스 설계서

■ 제·개정 이력

날짜	버전	작성자	승인자	내용

D1	클래스 설계서				
시스템명	JERY's WORD	서브시스템명	뉴스 검색 시스템, 댓글 시각화 시스템, 기간 설정 시스템, 모델 학습 시스템		
단계 명	분석	작성일자	2020. 05. 02	버전	1.0

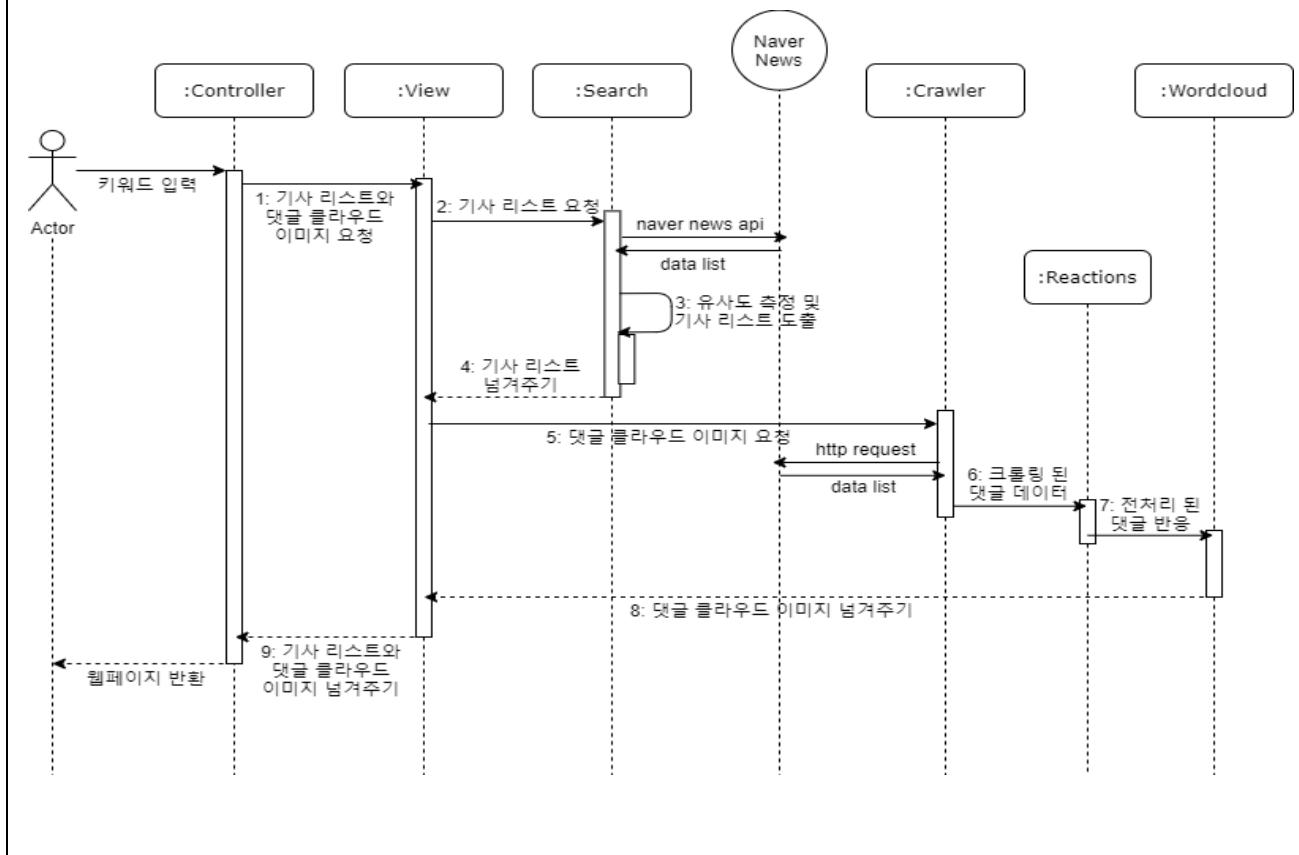
1) 설계 클래스 목록

설계 클래스 ID	설계 클래스명	관련 유스케이스 ID
JW_DC_010-01	SEA001Controller	JW_ASM_010-01, JW_ASM_010-02, JW_CVM_020-01, JW_DSM_030-01, JW_DSM_030-02
JW_DC_010-02	SEA002View	JW_ASM_010-01, JW_ASM_010-02, JW_CVM_020-01, JW_DSM_030-01, JW_DSM_030-02
JW_DC_010-03	SEA003Search	JW_ASM_010-01, JW_ASM_010-02
JW_DC_010-04	SEA004Crawler	JW_CVM_020-01, JW_DSM_030-01, JW_DSM_030-02
JW_DC_010-05	SEA005Reactions	JW_CVM_020-01, JW_DSM_030-02

JW_DC_010-06	SEA006Wordcloud	JW_ASM_010-02, JW_CVM_020-01, JW_DSM_030-01, JW_DSM_030-02
JW_DC_010-07	SEA007Article	JW_ASM_010-02, JW_DSM_030-01, JW_DSM_030-02
JW_DC_020-01	MOD001Articles	JW_MLM_040-01
JW_DC_020-02	MOD002learnModel	JW_MLM_040-01
JW_DC_020-03	MOD003article_url_spider	JW_MLM_040-01

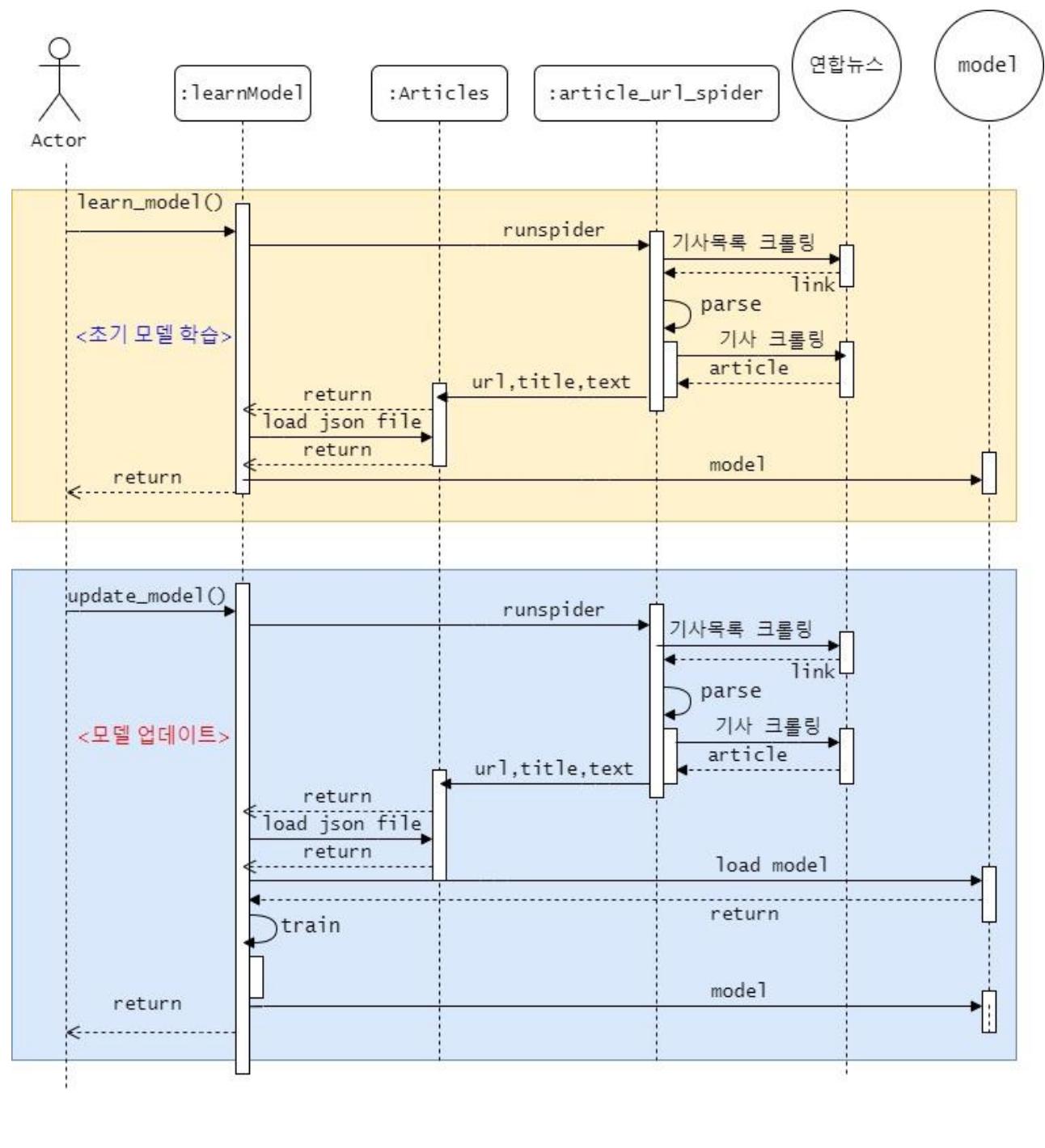
2) 시퀀스도

시퀀스도 ID	JW_SD_010-01	시퀀스도명	키워드 검색 및 댓글 시작화
관련 유스케이스 ID	JW_ASM_010-01, JW_ASM_010-02, JW_CVM_020-01		
주요 액터	사용자	주요 객체(또는 클래스)	View, Search, Crawler, Wordcloud



시퀀스도 ID	JW_SD_010-02	시퀀스도명	기간 설정
관련 유스케이스 ID	JW_DSM_030-01, JW_DSM_030-02		
주요 액터	사용자	주요 객체(또는 클래스)	View, Search, Wordcloud
<pre> sequenceDiagram participant Actor participant Controller participant View participant Search participant Wordcloud Actor->>Controller: 기간 설정 activate Controller Controller->>View: 1: 기사 리스트 요청 activate View View->>Search: 2: 설정 기간의 기사 리스트 요청 activate Search Search->>Wordcloud: 3: 새로운 기사 리스트 넘겨주기 activate Wordcloud Wordcloud->>View: 4: 댓글 클라우드 이미지 요청 activate View View-->>Actor: 5: 댓글 클라우드 이미지 반환 activate Actor Actor->>Controller: 6: 기사 리스트와 댓글 클라우드 이미지 넘겨주기 deactivate All </pre>			

시퀀스도 ID	JW_SD_020-01	시퀀스도명	모델 학습 및 업데이트
관련 유스케이스 ID	JW_MLM_040-01		
주요 액터	관리자	주요 객체(또는 클래스)	learnModel, article_url_spider, Articles



3) 설계 클래스도

설계 클래스도 ID	JW_DCD_010-01	설계 클래스도명	키워드 검색 및 댓글 시각화
관련 유스케이스 ID	JW_ASM_010-01, JW_ASM_010-02, JW_CVM_020-01, JW_DSM_030-01, JW_DSM_030-02		
<pre> classDiagram class Controller { index result_page detail_page } class View { topic top10 article_list() photo_list() detail() select_period() } class Search { topic set_request() get_articles(res) get_result() search() } class Article { source url originalurl title time } class Wordcloud { period articles wordcloud(Reactions) } class Reactions { comment num url } class Crawler { number_of_comment urls period get_reactions() } Controller "1" --> "use" View View "1" --> "use" Search View --> "use" Wordcloud View --> "use" Crawler Search --> "have" Article Article --> "1" --> "use" Wordcloud Wordcloud --> "use" Reactions Crawler --> "have" Reactions Reactions --> "use" Wordcloud </pre>			

설계 클래스도 ID	JW_DCD_020-01	설계 클래스도명	모델 학습
관련 유스케이스 ID	JW_MLM_040-01		
<pre> classDiagram class Articles { url title text } class learnModel { tokenizing(dir) learn_model() update_model() } class article_url_spider { start_urls parse(res) parse_article(res) } learnModel --> "use" Articles learnModel --> "use" article_url_spider article_url_spider --> "store" Articles </pre>			

4) 설계 클래스 정의

설계 클래스 ID	JW_DC_010-	설계 클래스명	SEA001Controller	
속성				
속성명	가시성	타입	기본값	설명
index	public	html	없음	검색어를 입력하는 검색창 화면
result_page	public	html	없음	검색어 입력 후 기사리스트 와 댓글 워드클라우드가 나열되는 화면
detail_page	public	html	없음	댓글 워드클라우드를 하나 클릭했을 경우 해당 기간 동안의 기사리스트가 나열되는 화면
오퍼레이션				
오퍼레이션명	가시성	파라미터	반환타입	설명
-	-	-	-	-

설계 클래스 ID	JW_DC_010-02	설계 클래스명	SEA002View	
속성				
속성명	가시성	타입	기본값	설명
topic	public	String	없음	사용자가 입력한 키워드
top10	public	String	없음	키워드로 검색한 기사 중 유사도가 높은 상위 10 개 기사 리스트
오퍼레이션				
오퍼레이션명	가시성	파라미터	반환타입	설명

article_list()	public	topic	기사 리스트와 유사도	Search 객체 이용하여 키워드로 검색 결과인 기사 리스트 출력, 워드 클라우드 함수 호출
photo_list()	public	urls	워드 클라우드 이미지 3 개	Crawler 와 Wordcloud 객체 이용하여 댓글 크롤링 후 워드 클라우드 이미지 출력
detail()	public	Wordcloud	해당 워드 클라우드 기간의 기사 리스트	Wordcloud 객체의 articles 를 사용하여 해당 기간의 기사 리스트를 출력
select_period()	public	int	해당 워드 클라우드 기간의 기사 리스트	detail()을 이용하여 사용자가 선택한 기간의 기사 리스트와 해당 워드 클라우드를 출력

설계 클래스 ID	설계 클래스명			
속성				
속성명	가시성	타입	기본값	설명
topic	public	String	없음	사용자가 입력한 키워드
오퍼레이션				
오퍼레이션명	가시성	파라미터	반환타입	설명
set_request()	public	topic, period	request	api 호출을 위한 request 설정
get_articles()	public	response, topic	기사 list	네이버 api 이용 기사 list 받아옴.
get_result()	public	topic, period	기사 list, 유사도	기사 list 를 유사도에 따라 정제.

search()	public	-	기사 list, 유사도	키워드의 기간별 검색 결과인 기사 list 와 유사도를 받아옴.
----------	--------	---	--------------	-------------------------------------

설계 클래스 ID	JW_DC_010-04	설계 클래스명	SEA004Crawler	
속성				
속성명	가시성	타입	기본값	설명
number_of_comment	public	Int	0	한 기간의 기사들의 댓글 총 개수
urls	public	List of String	없음	기사들의 링크 모음
period	public	Int	없음	기사 검색 기간 1: 0~3 일전 2: 4~6 일전 3: 7~9 일전
오퍼레이션				
오퍼레이션명	가시성	파라미터	반환타입	설명
get_reactions()	public	-	List of Reactions	urls 를 이용하여 각 기사의 댓글을 최대 50 개 크롤링 후 Reactions 모음을 반환

설계 클래스 ID	JW_DC_010-05	설계 클래스명	SEA005Reactions	
속성				
속성명	가시성	타입	기본값	설명
comment	public	String	없음	한 기사의 댓글 모음. 최대 50 개

num	public	Int	0	한 기사의 댓글 개수
url	public	String	없음	기사 링크
오퍼레이션				
오퍼레이션명	가시성	파라미터	반환타입	설명
-	-	-	-	-

설계 클래스 ID	JW_DC_010-06	설계 클래스명	SEA006Wordcloud	
속성				
속성명	가시성	타입	기본값	설명
period	public	Int	없음	기사 검색 기간 1: 0~3 일전 2: 4~6 일전 3: 7~9 일전
articles	public	Article	없음	해당 기간의 기사 리스트
오퍼레이션				
오퍼레이션명	가시성	파라미터	반환타입	설명
wordcloud()	public	Reactions	wordcloud 이미지	한 기간의 댓글 모음을 이용하여 빈도수로 워드 클라우드를 생성하여 반환

설계 클래스 ID	JW_DC_010-07	설계 클래스명	SEA007Article	
속성				
속성명	가시성	타입	기본값	설명

source	public	String	없음	기사를 발행한 언론사
url	public	String	없음	기사의 링크
originalurl	public	String	없음	기사 원문의 링크
title	public	String	없음	기사의 제목
time	public	String	없음	기사의 발행시각
오퍼레이션				
오퍼레이션명	가시성	파라미터	반환타입	설명
-	-	-	-	-

설계 클래스 ID	JW_DC_020-01	설계 클래스명	MOD001Articles	
속성				
속성명	가시성	타입	기본값	설명
url	public	String	없음	기사의 링크
title	public	String	없음	기사의 제목
text	public	String	없음	기사 본문 내용
오퍼레이션				
오퍼레이션명	가시성	파라미터	반환타입	설명
-	-	-	-	-

설계 클래스 ID	JW_DC_020-02	설계 클래스명	MOD002learnModel	
속성				
속성명	가시성	타입	기본값	설명
-	-	-	-	-

오퍼레이션				
오퍼레이션명	가시성	파라미터	반환타입	설명
tokenizing()	public	학습시킬 데이터 파일 위치	List of String	tokenizing 후 명사만 골라 반환.
learn_model()	public	-	없음	학습 후 초기 모델 저장. 연합뉴스 크롤링한 json 파일을 tokenizing 하여 저장.
update_model()	public	미리 학습된 모델	없음	새로운 데이터로 기존 모델 업데이트

설계 클래스 ID	JW_DC_020-03	설계 클래스명	MOD003article_url_spider	
속성				
속성명	가시성	타입	기본값	설명
start_urls	public	List of String	없음	크롤링할 사이트의 링크모음(연합뉴스)
오퍼레이션				
오퍼레이션명	가시성	파라미터	반환타입	설명
parse()	public	response	json(List of Articles)	각 기사 목록 페이지에서 기사 링크 크롤링
parse_article()	public	response	Articles	기사 내부 제목, 본문 크롤링

2.2 유사도 추정 설계서

■ 제·개정 이력

날짜	버전	작성자	승인자	내용

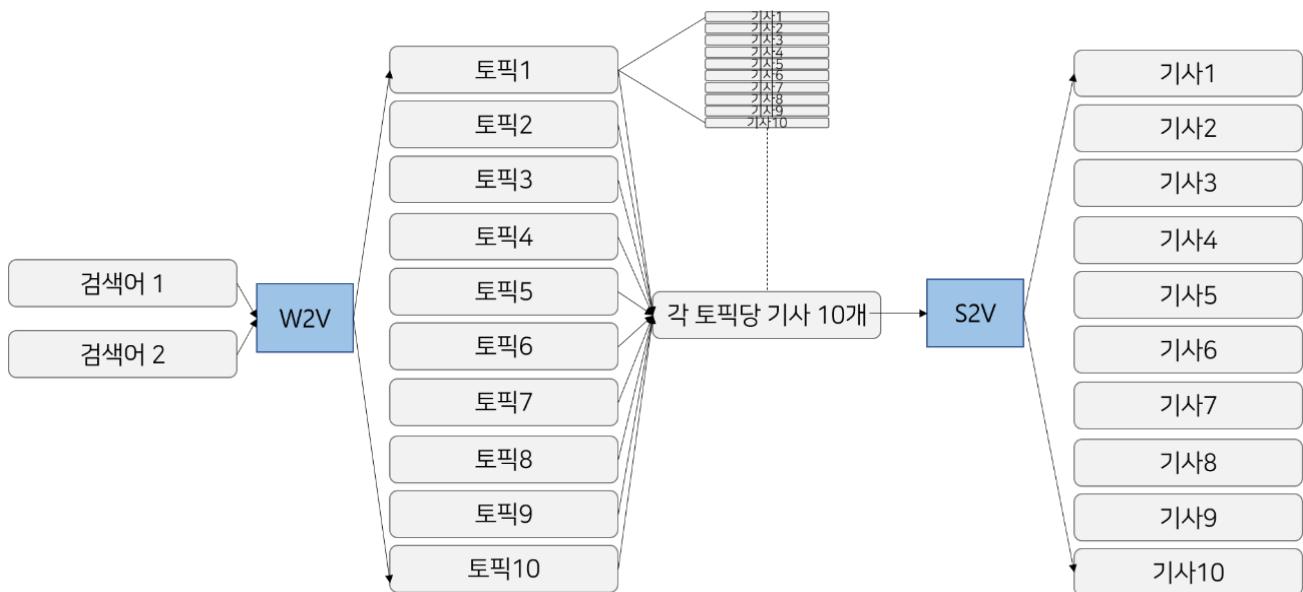
D2	유사도 추정 설계서				
시스템명	JERRY's WORD	서브시스템명	뉴스 검색 시스템, 댓글 시각화 시스템, 기간 설정 시스템, 모델 학습 시스템		
단계 명	설계	작성일자	2020. 05. 02	버전	1.0

1) 유사도 추정 관련 데이터 정의

모델 ID	모델 명	개요	관련 클래스
JW_SM_CR	데이터크롤링	연합뉴스 데이터크롤링	JW_DC_020-03
JW_SM_ST1	모델 학습	W2V	JW_DC_020-02
JW_SM_ST2	모델 학습	S2V	JW_DC_020-02
JW_SM_UD	모델 업데이트	1 주일마다 업데이트되는 뉴스 데이터	JW_DC_020-02

2) 유사도 프로세스

- i. W2V (WORD2VECTOR) : 두 검색어의 유사한 10 개의 토픽을 JW_SM_ST1 로부터 뽑아내는 모듈
- ii. S2V(SENTENCE2VECTOR) : 두 검색어와 W2V 을 통해 뽑아낸 토픽에 대한 기사들과의 유사성을 보여주는 모듈



2.3 사용자 인터페이스 설계서

■ 제·개정 이력

날짜	버전	작성자	승인자	내용

D3		사용자 인터페이스 설계서			
시스템명	JERY's WORD	서브시스템명	뉴스 검색 시스템, 댓글 시각화 시스템, 기간 설정 시스템, 모델 학습 시스템		
단계 명	설계	작성일자	2020. 05. 02	버전	1.0

1) 사용자 인터페이스 구조도

업무 영역 (Level 1)	Level 2	Level 3
시작페이지(JW_STP)		
	메인 view(JW_MV)	
		기사 페이지 연결
		워드클라우드(JW_WC)

2) 화면

화면 ID	화면명	관련 유스케이스 ID
JW_STP(시작페이지)	시작 페이지	JW_ASM_010-01
JW_MV(메인 view)	메인 페이지	JW_ASM_010-02, JW_CVM_020-01
JW_WC(워드클라우드)	워드클라우드 페이지	JW_DSM_030-01, JW_DSM_030-02

3) 출력물

출력물 ID	출력물명	관련 유스케이스 ID
OUT_ND	뉴스 타이틀	JW_ASM_010-01
OUT_SM	유사도 정도	JW_ASM_010-02
OUT_WC	워드클라우드 이미지	JW_CVM_020-01
OUT_CD	댓글 뉴스 타이틀, 신문사, 기사 등록시간, 날짜	JW_DSM_030-02
OUT_DT	기간 설정된 날짜	JW_DSM_030-01

4) 화면 상세설계

화면 ID	JW_STP(시작페이지)	화면명	시작페이지
관련 유스케이스	JW_ASM_010-01		

관련 시퀀스도 ID	없음		
화면유형	조회	메뉴경로	Level1
화면개요	<p>시작페이지를 보여준다.</p> <p>시작페이지로 들어오면 사이트 로고 이미지와 키워드를 통해 유사도를 검색할 수 있는 텍스트박스가 있다.</p> <p>돋보기 아이콘을 선택하면 메인 뷰로 들어간다.</p>		

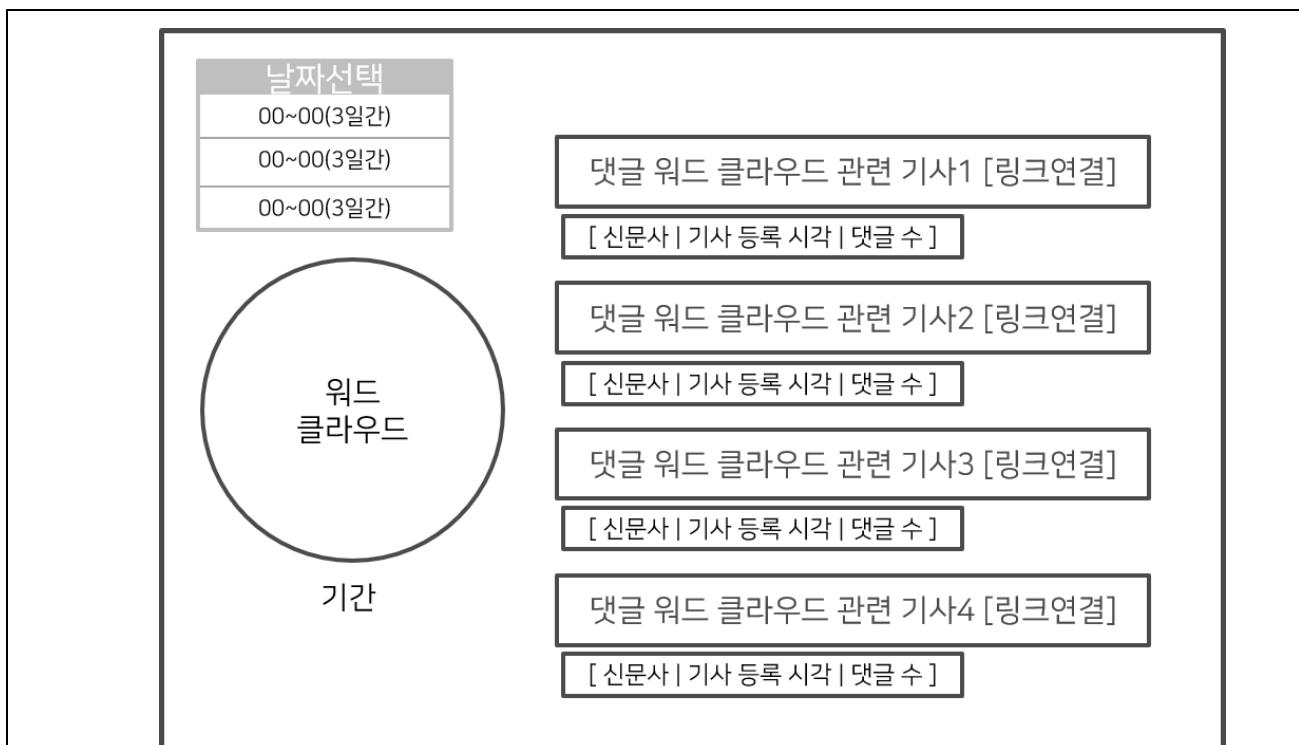


입출력 항목				
항목명	컨트롤명	타입 및 길이	속성	Validation Check
Keyword	Kw_keyboard_input	Char(16)	I(Input)	0
Logolimage	Lg_image_output	image	O(Output)	0
SearchImage	Sg_image)output	Image	O(Output)	0
처리 내용				
메인 뷰로 넘어간다.				
기술적 고려사항				
엔터키 입력이나 마우스 클릭 시 이동				

화면 ID	JW_MV(메인 view)	화면명	메인 페이지	
관련 유스케이스	JW_ASM_010-01, JW_ASM_010-02, JW_CVM_020-01			
관련 시퀀스도 ID	JW_SD_010-01			
화면유형	조회, 간단	메뉴경로	깊이 2	
화면개요	<p>상단에는 사용자가 입력한 검색어(키워드)1, 검색어 2 가 텍스트박스안에 들어간다.</p> <p>검색어와 유사도가 토픽 10 개로 크롤링한 후 높은 유사도순으로 기사타이틀 10 개가 보여진다.</p> <p>기사 제목 옆에는 검색어 1,2 와 기사의 유사도 검색을 통해 나온 기사의 유사도를 표시해준다.</p> <p>기사 기간별 댓글 반응변화를 3 개의 워드 클라우드 이미지를 통해 보여준다</p> <p>각 워드 클라우드 이미지는 기본 설정된 기간(3 일)의 이미지를 보여준다.</p>			
<p>The screenshot shows a search interface with two input fields for '검색어1, 검색어2'. Below the search button are three article cards: '검색어와 유사도가 높은 기사1 [링크연결]' with a similarity of 90%, '검색어와 유사도가 높은 기사2 [링크연결]' with 80%, and '검색어와 유사도가 높은 기사3 [링크연결]' with 70%. Below the articles is a section titled '기사 기간별 댓글 반응변화' featuring three circular word clouds labeled '워드 클라우드' with the text '00~00(3일간)' underneath each.</p>				
입출력 항목				
항목명	컨트롤명	타입 및 길이	속성	Validation Check

Keyword	Kw_keyboard_output	Char(16)	O(Output)	O
List_news_title (10 개)	Nt_output	Char(32)	O(Output)	O
Wordcloud1	W_output1	image	O(Output)	O
Wordcloud2	W_output2	image	O(Output)	O
Wordcloud3	W_output3	image	O(Output)	O
처리 내용				
기사 제목과 유사도를 보여주며 기사 제목 클릭 시 해당 기사를 볼 수 있는 페이지로 연결되고, 워드클라우드 이미지 클릭 시 워드클라우드 뷰로 넘어간다.				
기술적 고려사항				
기사 및 워드클라우드 생성 속도 처리가 필요하다.				

화면 ID	JW_WC(워드클라우드)	화면명	워드 클라우드페이지
관련 유스케이스	JW_DSM_030-01, JW_DSM_030-02		
관련 시퀀스도 ID	JW_SD_010-01, JW_SD_010-02		
화면유형	조회, 간신	메뉴경로	깊이 3
화면개요	<p>왼쪽 상단에 날짜선택 콤보박스를 보여준다. 콤보박스를 눌렀을 때 사용자가 원하는 날짜의 워드 클라우드 페이지로 들어간다.</p> <p>기간에 맞는 워드클라우드 이미지를 기간과 함께 보여준다.</p> <p>오른쪽에 워드 클라우드 이미지 관련 기사 타이틀과 신문사, 기사 등록 시간, 댓글 수를 10 개 보여준다.</p>		



입출력 항목

항목명	컨트롤명	타입 및 길이	속성	Validation Check
Wordcloud_wv	W_output	image	O(Output)	O
Comment_tile1	CT_output1	Char(32)	O(Output)	O
Comment_name1	CN_output1	Char(16)	O(Output)	O
Comment_date	CD_output1	Char(16)	O(Output)	O
Comment_number	CNB_output1	Char(16)	O(Output)	O

처리 내용

콤보박스를 통해 다른 날짜의 기간 워드클라우드 페이지로 들어간다.

기술적 고려사항

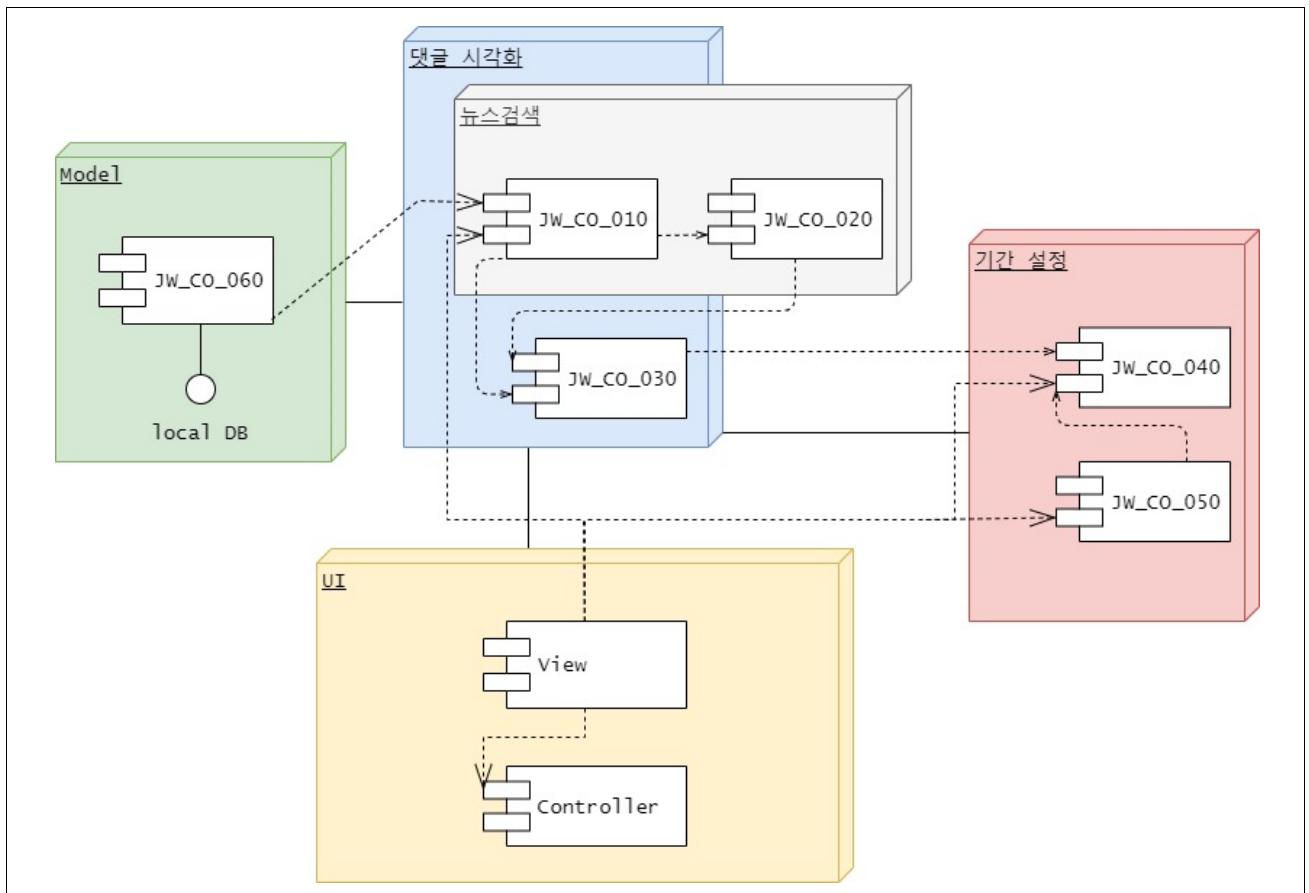
2.4 컴포넌트 설계서

■ 제·개정 이력

날짜	버전	작성자	승인자	내용

D4	컴포넌트 설계서			
시스템명	JERY's WORD	서브시스템명	뉴스 검색 시스템, 댓글 시각화 시스템, 기간 설정 시스템, 모델 학습 시스템	
단계 명	설계	작성일자	2020. 05. 02	버전 1.0

1) 컴포넌트 구조도



2) 컴포넌트 목록

컴포넌트 ID	컴포넌트명	개요	관련 유스케이스 ID
JW_CO_010	검색	키워드로 뉴스검색	JW_ASM_010-01
JW_CO_020	기사확인	검색 결과로 뉴스기사와 유사도 확인	JW_ASM_010-02
JW_CO_030	시각화 된 댓글확인	검색된 뉴스의 댓글을 시각화 시킨 워드 클라우드 확인	JW_CVM_020-01
JW_CO_040	기간 설정	뉴스 상세 페이지에서 콤보박스에서 다른 두개의 기간을 선택할 수 있다. 선택하면 해당 기간의 뉴스기사를 재조회한다.	JW_DSM_030-01
JW_CO_050	기간별 뉴스 기사	워드클라우드 하나를 누르면 해당 기간의 뉴스 기사만 모아준다. (뉴스 상세 페이지)	JW_DSM_030-02
JW_CO_060	모델 학습	크롤링한 뉴스 텍스트 데이터로 모델을 초기에 학습시키거나 정기적으로 업데이트한다.	JW_MLM_040-01

3) 컴포넌트 명세

관련 서브시스템 ID	JW_ASM_010, JW_CVM_020		
컴포넌트 ID	JW_CO_010	컴포넌트명	검색
컴포넌트 개요	메인화면 검색창에 사용자가 키워드를 입력한다. 그 키워드로 뉴스검색을 할 수 있다. 이후에는 뉴스 기사 & 유사도 확인 댓글 시각화 워드클라우드로 연결된다.		
내부 클래스			
ID	클래스명	비고	

JW_DC_010-02	SEA002View	키워드를 입력 후 엔터/클릭 이벤트가 들어오면 JW_CO_020, JW_CO_030 을 호출한다.
--------------	------------	--

관련 서브시스템 ID	JW_ASM_010		
컴포넌트 ID	JW_CO_020	컴포넌트명	기사확인
컴포넌트 개요	키워드를 유사한 키워드 10개로 만든 뒤, 각각의 키워드로 10개씩 뽑은 총 100개의 기사를 유사도 상위 10개로 간추린다. 검색 결과인 뉴스 기사 10개 와 그것의 유사도를 받아 출력한다. topic과 3가지 기간으로 검색한 결과인 기사링크 list 3개도 받아서 저장해 놓았다가 Crawler, Wordcloud에 이용한다.		
내부 클래스			
ID	클래스명	비고	
JW_DC_010-03	SEA003Search	topic 으로 Search 객체를 생성한 후, 기사 list 와 유사도를 얻기 위해 search 오퍼레이션을 호출하고 View 에 반환한다. search 오퍼레이션은 내부의 다른 오퍼레이션과 Article 객체 리스트를 이용하여 전체 기간의 유사도 상위 10 개 기사 list 와 유사도를 반환한다. topic 으로 검색하는 과정에서 NAVER API 가 사용된다.	
JW_DC_010-06	SEA006Wordcloud	View 는 search 에 접근하기 전에 Wordcloud 객체를 3 개 먼저 생성한다. search 오퍼레이션이 작동되면서 기간별로 기사들이 Wordcloud 객체들에 저장된다.	
JW_DC_010-07	SEA007Article	Search 가 기사 데이터를 저장하여 사용하기 위한 클래스로 객체리스트로 제작되어 View 에 전달된다. 또한 전체 기사 목록이 기간에 따라 나누어 Wordcloud 객체에 저장된다.	

관련 서브시스템 ID	JW_CVM_020		
컴포넌트 ID	JW_CO_030	컴포넌트명	시각화 된 댓글확인

컴포넌트 개요	View는 Crawler 객체를 이용하여 기사들의 댓글 list인 Reactions를 얻고 이를 View가 Wordcloud에 넘겨 워드 클라우드 이미지를 얻는다. 추후 워드 클라우드 이미지 재활용을 위해 연결이 유지될 동안만 Wordcloud의 wc_image에 저장해 놓는다.		
내부 클래스			
ID	클래스명	비고	
JW_DC_010-04	SEA004Crawler	한 기간의 기사 링크 list를 매개변수로 하여 Crawler 객체를 생성한 후 get_reactions 오퍼레이션을 호출하고 View에 반환한다. get_reactions 오퍼레이션은 Reactions 객체 리스트를 이용하여 기사들의 댓글 list를 반환한다.	
JW_DC_010-05	SEA005Reactions	Crawler가 댓글 데이터를 저장하여 사용하기 위해 사용되는 클래스로 객체리스트로 제작되어 View에 전달된다.	
JW_DC_010-06	SEA006Wordcloud	Wordcloud 객체를 최초 1회 생성한 후 wordcloud 오퍼레이션을 호출하고 View에 반환한다. wordcloud 오퍼레이션은 댓글 list를 이용하여 워드클라우드 이미지를 반환한다.	

관련 서브시스템 ID	JW_DSM_030		
컴포넌트 ID	JW_CO_040	컴포넌트명	기간 설정
컴포넌트 개요	뉴스 상세 페이지에서 콤보박스에서 다른 두개의 기간을 선택할 수 있다. 선택하면 해당 기간의 뉴스기사를 재조회한다.		
내부 클래스			
ID	클래스명	비고	
JW_DC_010-02	SEA002View	select_period 오퍼레이션이 호출된다. 해당 기간에 맞는 period를 가진 WordCloud 객체를	
JW_DC_010-06	SEA006Wordcloud	WordCloud 객체가 가지고 있는 Article 리스트인 articles를 View가 접근하여 사용할 수	

JW_DC_010-07	SEA007Article	기사 데이터를 저장하여 사용하기 위해 사용되는 클래스로 객체리스트로 제작되어 View에
--------------	---------------	--

관련 서브시스템 ID	JW_DSM_030		
컴포넌트 ID	JW_CO_050	컴포넌트명	기간별 뉴스 기사
컴포넌트 개요	뉴스 상세 페이지에서 워드클라우드 하나를 누르면 해당 기간의 뉴스 기사만 모아준다. View에서 Wordcloud 객체에 저장된 articles와 이미 생성되어 저장되었던 워드클라우드 이미지를 사용한다.		
내부 클래스			
ID	클래스명	비고	
JW_DC_010-02	SEA002View	detail 오퍼레이션이 호출된다. 이미지 위치에 따라 기간을 계산하여 해당 기간에 맞는 period를 가진 WordCloud 객체를 사용한다.	
JW_DC_010-06	SEA006Wordcloud	WordCloud 객체가 가지고 있는 Article 리스트인 articles와 워드클라우드 이미지를 View 가 접근하여 사용할 수 있다.	
JW_DC_010-07	SEA007Article	기사 데이터를 저장하여 사용하기 위한 클래스로 객체리스트로 제작되어 View에 전달된다.	

관련 서브시스템 ID	JW_MLM_030		
컴포넌트 ID	JW_CO_060	컴포넌트명	모델 학습
컴포넌트 개요	크롤링한 뉴스 텍스트 데이터로 모델을 초기에 학습시키거나 정기적으로 업데이트한다.		
내부 클래스			
ID	클래스명	비고	
JW_DC_020-01	MOD001Articles	기사의 텍스트 데이터를 저장하여 모델 학습에 사용되는 클래스로 객체리스트로 제작되어 View에 전달된다.	

JW_DC_020-02	MOD002learnModel	모델을 처음에 학습시키거나 추후에 업데이트 시킬 때 learn_model이나 update_model 오퍼레이션이 호출된다. 오퍼레이션이 완료되면 모델이 서버에 저장된다.
JW_DC_020-03	MOD003article_url_spider	학습 데이터를 수집하기 위한 클래스로, parse 오퍼레이션을 호출하여 연합뉴스의 기사들의 제목과 본문을 다량 수집한다.

2.5 인터페이스 설계서

■ 제·개정 이력

날짜	버전	작성자	승인자	내용

D4	인터페이스 정의서				
시스템명	JERY's WORD	서브시스템명	뉴스 검색 시스템, 댓글 시각화 시스템, 기간 설정 시스템, 모델 학습 시스템		
단계명	설계	작성일자	2020. 05. 02	버전	1.0

1) 인터페이스 목록

송신 인터 페이 스 번호	일련 번호	송신시스템명	프로그램ID	전달			수신					관 련 요 구 사 항 	비 고
				처리형 태	인 터 페 이 스 방	발생 빈도	상 대 담 당 자 확	프로그램ID	수신시스템명	일련 번호	수신번호		

					식		인					D	
I_0 1	T_0 1	Keyword transmission	Keyword	Onlin e	미 정	검색 시		Keyword	Keyword Reception	R_0 1	R_01_0 1		
I_0 2	T_0 2	Similarity transmission	Similarity	Onlin e	미 정	I_0 1 실행 후		Similarity	Similarity Reception	R_0 2	R_02_0 1		
I_0 3	T_0 3	News_crawling transmission	News_crawling	Onlin e	미 정	I_0 2 실행 후		News_crawling	Register Reception	R_0 3	R_03_0 1		
I_0 4	T_0 4	Similarity_news transmission	Similarity_news	Onlin e	미 정	I_0 3 실행 후 10 번		Similarity_news	Similarity_news Reception	R_0 4	R_04_0 1		
I_0 5	T_0 5	Comment_crawlin g transmission	Comment_crawlin g	Onlin e	미 정	I_0 4 실행 후 10		Comment_crawlin g	Comment_crawlin g Reception	R_0 5	R_05_0 1		

					번							
I_0 6	T_0 6	Wordcloud transmission	Wordcloud	Onlin e	미 정	I_0 5실 행 후	Wordcloud	Wordcloud Reception	R_0 6	R_06_0 1		
I_0 7	T_0 7	Select_wordcloud transmission	Select_wordcloud	Onlin e	미 정	이미 지선 택후	Select_wordcloud	Select_wordcloud Reception	R_0 7	R_07_0 1		
I_0 8	T_0 8	Select_date transmission	Select_date	Onlin e	미 정	기간 콤보 박스 선택 후	Select_date	Select_date Reception	R_0 8	R_08_0 1		

2) 인터페이스 명세

인 터 페 이 스 번	데이터 송신시스템					송신 프로 그램 ID	데이터 수신시스템					수 신 플 로 그 램
	시스템명	데이터 상소명	속성명	데이터 타 입	길 이		데이터저장소명	속성명	데이터 타 입	길 이	시스템명	

ID												호
I_0 1	Keyword transmission	Keyword list	Keyword1	Char	16	K_1	Keyword list	Keyword1	Char	16	Keword Reception	
			Keyword2	Char	16	K_2		Keyword2	Char	16		
I_0 2	Similarity transmission	Similarity list	Similarity1	Char	16	S_1	Similarity list	Similarity1	Char	16	Similarity Reception	
			Similarity2	Char	16	S_2		Similarity2	Char	16		
			Similarity3	Char	16	S_3		Similarity3	Char	16		
			Similarity4	Char	16	S_4		Similarity4	Char	16		
			Similarity5	Char	16	S_5		Similarity5	Char	16		
			Similarity6	Char	16	S_6		Similarity6	Char	16		
			Similarity7	Char	16	S_7		Similarity7	Char	16		
			Similarity8	Char	16	S_8		Similarity8	Char	16		
			Similarity9	Char	16	S_9		Similarity9	Char	16		
			Similarity10	Char	16	S_10		Similarity10	Char	16		
I_0 3	News_crawli ng transmission	News_cra wling list	News_crawling	bin	dat a에 따 라	N_C	News_crawling list	News_crawling	bin	dat a에 따 라	News_cra wling Reception	
I_0 4	Similarity_ne ws transmission	Similarity _news	Similarity_news1	미정	미 정	S_N	Similarity_new s	Similarity_news1	미정	미 정	Similarity_ news Reception	
			Similarity_news2					Similarity_news2				
			Similarity_news3					Similarity_news3				
			Similarity_news4					Similarity_news4				
			Similarity_news5					Similarity_news5				
			Similarity_news6					Similarity_news6				
			Similarity_news7					Similarity_news7				

			Similarity_news8				Similarity_news8				
			Similarity_news9				Similarity_news9				
			Similarity_news10				Similarity_news10				
I_0 5	Comment_crawling transmission	Comment_crawling list	Comment_crawling1 (약 50개)	미정	미 정	C_C	Comment_crawling list	Comment_crawling1 (약 50개)	미정	미 정	Comment_crawling Reception
			Comment_crawling2 (약 50개)					Comment_crawling2 (약 50개)			
			Comment_crawling3 (약 50개)					Comment_crawling3 (약 50개)			
			Comment_crawling4 (약 50개)					Comment_crawling4 (약 50개)			
			Comment_crawling5 (약 50개)					Comment_crawling5 (약 50개)			
			Comment_crawling6 (약 50개)					Comment_crawling6 (약 50개)			
			Comment_crawling7 (약 50개)					Comment_crawling7 (약 50개)			
			Comment_crawling8 (약 50개)					Comment_crawling8 (약 50개)			

			Comment_crawling9 (약 50개)					Comment_crawlin g9 (약 50개)				
			Comment_crawling1 0 (약 50개)					Comment_crawlin g10 (약 50개)				
I_0 6	Wordcloud transmission	Wordclou d	Image	imag e	미 정	W_I	None	Image	미정	미 정	Wordcloud Reception	
I_0 7	Select_word cloud transmission	Select_w ordcloud	Select_image	even t	미 정	S_I	Select_wordclo ud	Select_image	미정	미 정	Select_wo rdcloud Reception	
I_0 8	Select_date transmission	Select_da te	Select_date	int	미 정	S_D	Select_date	Select_date	미정	미 정	Select_dat e Reception	

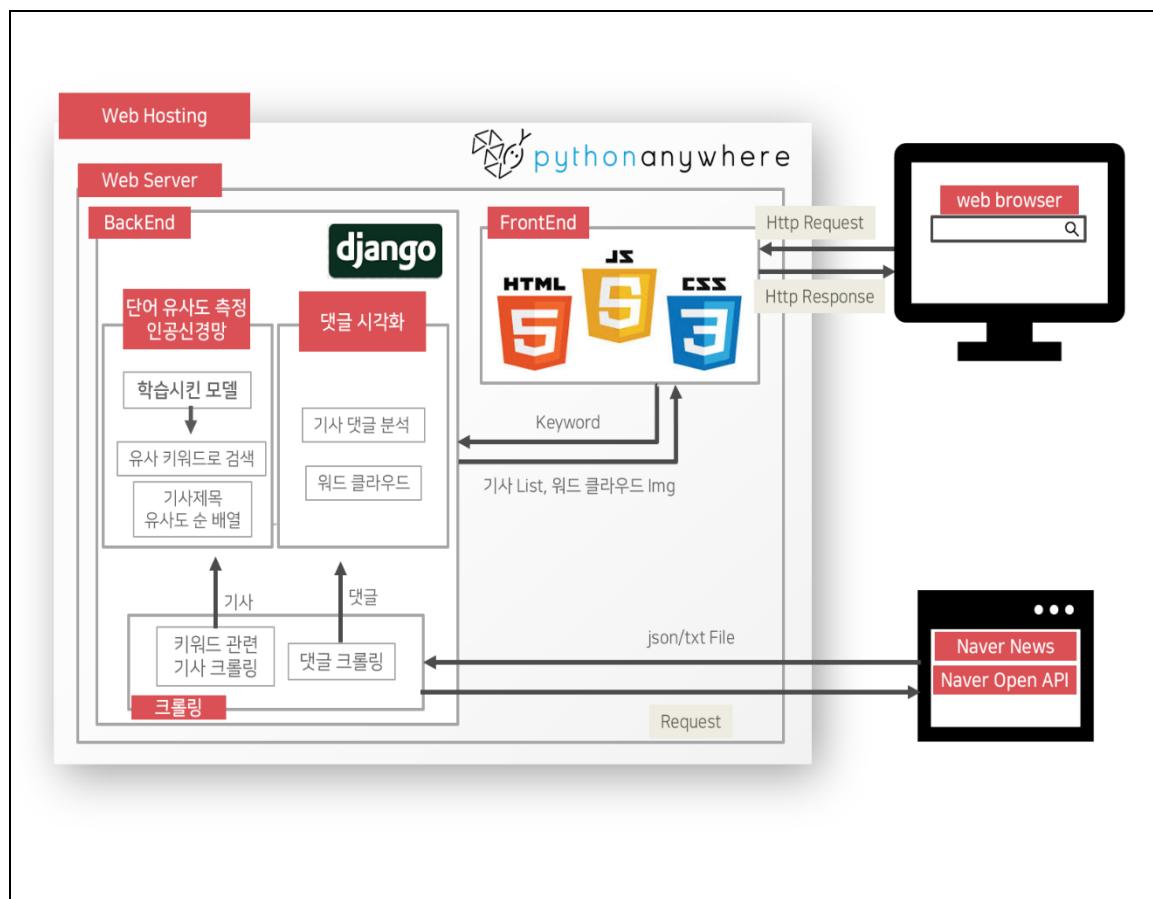
2.6 아키텍처 설계서

■ 제·개정 이력

날짜	버전	작성자	승인자	내용

D5	아키텍처 설계서				
시스템명	JERY's WORD	서브시스템명	뉴스 검색 시스템, 댓글 시각화 시스템, 기간 설정 시스템, 모델 학습 시스템		
단계명	설계	작성일자	2020. 05. 02	버전	1.0

1) 시스템 아키텍처



2) 아키텍처 요구사항 및 구현방안

요구사항 ID	VS-01
요구사항 내용	시스템의 품질 수준 확보를 위하여 결함예방, 품질만족 및 사용자 신뢰도 확보를 위하여 품질목표 수준을 설정하여 관리하여야 함
구현방안	
구분	수준
가용성	<ul style="list-style-type: none"> - 본 시스템은 1 일 24 시간, 365 일 상시 운영 가능성을 보장 - 발견된 결함은 시스템 오픈 이전에 모두 조치한 후 서비스 오픈 - 제공되기로 한 요구사항을 제공하였는지 여부는 각 기능 요구사항의 검증(테스트) 활동을 통해 예상된 결과가 도출되었을 경우 요구사항을 제공한 것으로 평가
학습성	<ul style="list-style-type: none"> - 모든 기능 및 정보에 대한 사용용이 및 편리성을 확보한 시스템 구축
이해성	<ul style="list-style-type: none"> - 기능이해도 : 정보접근의 편의성을 개선하고 시스템에서 제공하는 모든 기능에 대해 사용자 매뉴얼에서 정보를 제공해야 함 - 인터페이스 이해도 : 시스템에서 제공하는 인터페이스 기능과 방법을 초급자라도 쉽게 운영할 수 있는 직관적인 인터페이스로 제작해야 함. - 사용자가 원하는 기능을 쉽게 찾아서 사용할 수 있는 체계를 제공하여야 함 - 시스템은 콘텐츠의 모양이나 배치를 이해하기 쉽게 구성하여야 함
유지보수	<ul style="list-style-type: none"> - 시스템은 새로운 기능 구현을 위해 필요한 경우 업그레이드가 용이해야 하며, 보안사고 예방, 성능 개선 등이 발생하는 경우 패치를 제공하여야 함 - 유지보수에 용이하게 개발되어야 하며 시스템 확장 시 상호운영성, 이식성 등을 보장하여야 함

상호운용성	<ul style="list-style-type: none"> - 본 시스템은 관련된 정보시스템 및 기술표준과의 상호 운용성을 확보해야 함 - 시스템 인터페이스 요구사항은 기능 구현의 정확성뿐만 아니라 정보의 무결성, 데이터 정합성을 준수해야 함 - 시스템은 통상적인 운영시간 동안 가용성을 보장하여야 하며, 시스템 조건이 무엇이든지 간에 모든 채널에 동일한 자료 및 결과를 생성하고 인도해야 함 - 다양한 사용자 운영체제에 영향을 받지 않도록 시스템을 개발해야 함
-------	--

2.7 총괄시험 계획서

■ 제·개정 이력

날짜	버전	작성자	승인자	내용

D6	총괄시험 계획서				
시스템명	JERY's WORD	서브시스템명	뉴스 검색 시스템, 댓글 시각화 시스템, 기간 설정 시스템, 모델 학습 시스템		
단계명	설계	작성일자	2020. 05. 02	버전	1.0

1. 시험 대상 시스템

1.1 시스템 개요 (Overview of Application)

JERY's WORD의 응용시스템 개발부문은 2부분으로 구성되어있다.

검색어와의 유사성을 통한 기사 검색 시스템 구축

검색어와의 유사성을 통한 기사 검색 시스템은 검색어들과 유사성이 높은 키워드를 몇 가지 뽑고 그 키워드를 이용해 유사성이 높은 기사를 추려 사용자에게 보여준다.

기간별 댓글보기 서비스 구축

기간별 댓글보기 서비스는 검색일자로부터 3일간격으로 3개의 구간을 나눠 댓글을 분석해 반응변화를 한눈에 볼 수 있도록 하고 그 댓글이 달린 기사를 바로 볼 수 있도록 구성되어있다.

2. 가정(Assumptions) 또는 제약사항(Constraints)

- 시스템 개발 범위(scope)에 큰 변동이 없는 것을 가정한다.
- 성능 시험시 운영환경과 동일한 환경에서 시험이 가중한 것을 가정한다.
- 본 문서는 프로젝트의 응용시스템(소프트웨어)에 대한 시험 계획을 기술하고 있으며 하드웨어, 네트워크, 기타 기반 환경에 대한 시험은 본 계획과는 별도로 수행된다.

3. 시험 전략(Test Strategy)

3.1 시험 범위 (scope)

구분	시험항목
검색어와의 유사성을 통한 기사 검색 시스템	<ul style="list-style-type: none">- 검색어와 유사성 높은 키워드 뽑기- 검색어와 키워드를 조합해 나온 기사를 유사도 높은 순으로 추려냄
기간별 댓글보기 서비스	<ul style="list-style-type: none">- 단어 빈도수에 따른 댓글 분석- 데이터 시각화

3.2 개발 단계별 수행할 시험 종류

3.2.1 단위 시험

단위 시험은 개발되는 시스템 특성에 맞게 적절한 크기의 시스템 단위(unit)를 선정하고 각각의 단위가 하나의 독립적인 단위로서 그 기능을 적절히 수행하는가에 중점을 둔다. 이 프로젝트의 단위시험은 크게 2단계로 이루어진다. 1차 단위시험은 데이터 양을 줄인 상태로 모듈단위로 확인 할 수 있는 프로토타입을 개발하여 작성한 프로그램에 대해 모듈 단위로 시험을 수행하고, 1차 단위 시험이 완료되면 데이터 양을 늘리고 2차 단위시험을 수행해 남아있는 결함을 추가적으로 찾는다.

3.2.2 통합 시험

프로토타입을 통해 프로그램 내의 인터페이스를 시험하여 프로그램 통합그룹 또는 인터페이스 된 프로그램들이 기능대로 수행되는지 검증한다. 통합시험은 단위 시험을 통해 시스템의 단순 결함들이 걸러진 후 이루어지며, 구현이 완료된 전체

시스템의 통합된 기능을 시험한다.

3.2.3 시스템 시험

프로그램과 H/W, S/W를 포함하는 전체시스템을 대상으로 배포, 성능 등의 기능 외적 요인을 시험한다.

3.2.4 인수시험

사용자 승인시험은 개발된 시스템에 대해 운영 환경으로 전환하기 위한 준비를 완료하고, 운영환경과 동일한 환경하에서 시스템이 전체적으로 요구항을 만족시키는지 확인한다. 사용자입장에서 평가하여 그 시험결과에 대한 승인 여부를 결정하도록 한다.

	단위시험	통합시험	시스템시험	인수시험
시험대상	모듈, 화면	시스템 구성요소	전체 시스템	전체시스템
목적	단위모듈이 프로그램명세서의 기능에 부합되는지를 검증	인터페이스 검증	고객의 성능요구사항을 만족하는지 검증	시스템 인수여부 결정
시험환경	개발환경	개발환경	운영환경	운영환경

4. 시험 실행 계획

4.1 시험 수행 절차

업무 프로젝트 시험은 다음과 같은 절차에 따라 수행된다.

구분	프로젝트시험&개발팀	테스트환경구축팀	산출물
테스트계획	총괄시험계획	.	총괄시험계획서
	총괄시험계획 검토		
	시험교육		
단위시험	단위시험 케이스도출	테스트환경구축	단위시험케이스, 단위시험 결과서
	단위시험 수행		
	단위시험 결함검토		
	결함수행		
	결함수행여부확인		

통합시험	통합시험 시나리오/케이스도출	테스트환경구축	통합시험시나리오, 통합시험결과서	
	통합시험 수행			
	통합시험 결함검토			
	결함수행			
	결함수행여부 확인			
시스템시험	시스템시험 시나리오도출	테스트환경구축	시스템시험시나리오, 시스템시험결과서	
	시스템시험 수행			
	시스템시험 결함검토			
	결함수행			
	결함수행여부확인			
인수시험	인수시험 시나리오도출	테스트환경구축	인수시험 시나리오, 인수시험 결과서	
	인수시험 수행			
	인수시험 결함검토			
	결함수행			
	결함수행여부 확인			

4.2 발견된 문제점(결함)에 대한 수정절차

- 1) 시험 결과 오류 발생 또는 부적합이라고 판정 시 테스터는 이를 기록하여 시험 결함 내역을 기록한다.
- 2) 결함 검토회를 열어 시험 결과 도출된 오류 및 문제점들을 설명하고 결함 원인을 토의한다.
- 3) 개발자는 오류의 원인이 파악되면 해결방안을 찾아내어 결함이 수정 보완될 수 있도록 한다.
- 4) 오류 분석 결과 설계 변경이 필요한 경우 프로젝트관리 절차에 따라 변경을 실시한다.
- 5) 수정이 완료된 후 시험을 재실시하여 결함이 적절하게 수정되었는지 여부와 수정과정 중에 추가로 발견된 결함은 없는지 확인한다.

4.3 시험 수행 절차에 따른 역할

- 모든 절차는 팀원 모두가 함께한다.

구분	상세업무
총괄 시험 계획	시험 전략 및 수행절차 수립
	시험 조직 구성
	시험 전체 일정 수립
	총괄시험 계획서 작성
단위시험계획	응용개발팀별 단위시험 케이스 도출
통합시험 계획	통합시험시나리오작성
단위 시험	시험 환경 준비 - SW준비 - 시험 데이터 준비
	단위 시험 실행
	결합 수정
	결합 수정 여부 재검사
통합 시험환경 구축	시스템이관 - 응용프로그램 통합 및 환경 구축
통합 시험	통합시험 실행
	결합수정
	재검사
시스템 시험	시스템시험 계획
	시스템시험 설계
	시스템시험 실행
	시스템시험 결과보고
	시스템 시험 보완요청사항 보완
인수시험	인수시험 준비
	인수시험 실행
	인수시험 결과보고

4.4 시험 수행 일정

순번	주요활동	시험일자	완료일자	비고
1	총괄시험 계획	2020.04.30.	2020.04.30	
2	시스템시험 시나리오 작성	2020.05.04	2020.05.06	

3	통합시험 시나리오작성	2020.05.07	2020.05.08	
4	단위시험케이스 작성	2020.05.08	2020.05.09	
5	단위시험 환경 구축	2020.05.10	2020.05.13	
6	단위 시험 수행	2020.05.14	2020.05.15	
7	단위 시험 결함 수정 및 결과확인	2020.05.16	2020.05.17	
8	통합시험 환경구축	2020.05.18	2020.05.20	
9	통합시험 수행	2020.05.21.	2020.05.22	
10	통합시험 결함수정 및 결과확인	2020.05.23.	2020.05.24	
11	시스템시험 환경구축	2020.05.25.	2020.05.27	
12	시스템시험 수행	2020.05.28.	2020.05.29	
13	시스템시험 결함수정 및 결과확인	2020.05.30.	2020.06.01	
14	인수시험 시나리오 작성	2020.06.02	2020.06.02	
15	인수시험 환경구축	2020.06.02	2020.06.03	
16	인수시험 수행	2020.06.04.	2020.06.05	
17	인수시험 결함수정 및 결과확인	2020.06.08.	2020.06.12	