

C.E.M.S

Creative Elixir Members

자체 빅데이터 알고리즘 기반 AI 교육 플랫폼

외국인을 위한 TOPIK 교육 플랫폼 개발



Elixir

No Elixir, No Education



Disclaimer

본 자료는 회사의 영업활동에 대한 이해증진을 위해 (주)일릭서 (이하 "회사")에 의해 작성되었으며 이의 반출, 복사 또는 타인에 대한 재배포는 금지됨을 알려드립니다.

본 자료에 포함된 "예측정보"는 개별 확인 절차를 거치지 않은 정보들입니다. 이는 과거가 아닌 미래의 사건과 관계된 사항으로 회사의 향후 예상되는 경영현황 및 재무실적을 의미하고, 표현상으로는 '예상', '전망', '계획', '기대', '(E)' 등과 같은 단어를 포함합니다.

위 "예측정보"는 향후 경영환경의 변화 등에 따라 영향을 받으며, 본질적으로 불확실성을 내포하고 있는 바, 이러한 불확실성으로 인하여 실제 미래실적은 "예측정보"에 기재되거나 암시된 내용과 중대한 차이가 발생할 수 있습니다. 향후 전망은 현재 시장상황과 회사의 경영방향 등을 고려한 것으로 향후 시장환경의 변화와 전략수정 등에 따라 변경될 수 있으며, 별도의 고지 없이 변경될 수 있음을 양지하시기 바랍니다.

또한, 본 자료의 활용으로 인해 발생하는 손실에 대하여 회사 및 회사의 임직원들은 그 어떠한 책임도 부담하지 않음을 알려드립니다. (과실 및 기타의 경우 포함)

본 자료는 기업 비밀이 포함된 자료로서 무단 복제 및 배포를 삼가하여 주시기 바랍니다.

Contents

01 팀원 소개

02 기획 및 개발 준비

03 개발 과정

04 결과



01. 팀원 소개 (C.E.M.S of Elixir)

팀장
/AI, DB



장 정 권

“Elixir 없이는 교육도 없다.”

- AI
개체명&구문 분석기 모델 개발
문제 자동생성 알고리즘 (지문생성)
Data 멀티프로세싱 처리

- Database
문제DB & 사용결과 구조화

 More info

AI,
Back-end



김 재 희

“나는 오늘만 산다.”

- AI
Word2vec&Fasttext 모델 개발
문제 자동생성 알고리즘 (오답생성)
이미지 data 크롤링 및 분류

- Front-end / Back-end
회원탈퇴 기능 구현

 More info

AI,
Back-end



박 예 은

“믿음을 실현한다.”

- AI
한국어 용언 함수 구현
속담 및 속어 크롤링 및 분류
VBA기반 자동화 프로그램 구현

- Front-end / Back-end
topik2 쓰기 페이지 구현

 More info

Front-end,
Back-end



기 찬 웅

“다이아몬드는 영원하다.”

- AI
워드임베딩
한국어 단어, 문장 크롤링
문제 자동생성 알고리즘 (MASKED)

- Front-end/Back-end
기능 전반 구현

 More info

Front-end,
Back-end



이 정 민

“당신의 손가락으로 걸게 하라.”

- AI
Embedding + Transformer
기반 문장생성
듣기 파일 텍스트화

- Front-end/Back-end
기능 전반 구현

 More info

02. 기획 및 개발 준비(시장 선정 및 플랫폼 기획)



학생이 스스로 판단할 수 있는 **학습관리 시스템**

Elixir 만의 전문적인 통계 시스템으로 학습자의 부족한 부분을 유형별 확인



서로 지식을 공유할 수 있는 **소통의 장**

커뮤니티 사이트를 통한 지식 공유의 장 구축



원하는 유형의 문제를 **반복 학습**

자신의 **부족한 부분**을 반복 학습하여 효율적인 실력 향상



자신의 실력을 뽐내는 **순위 경쟁**

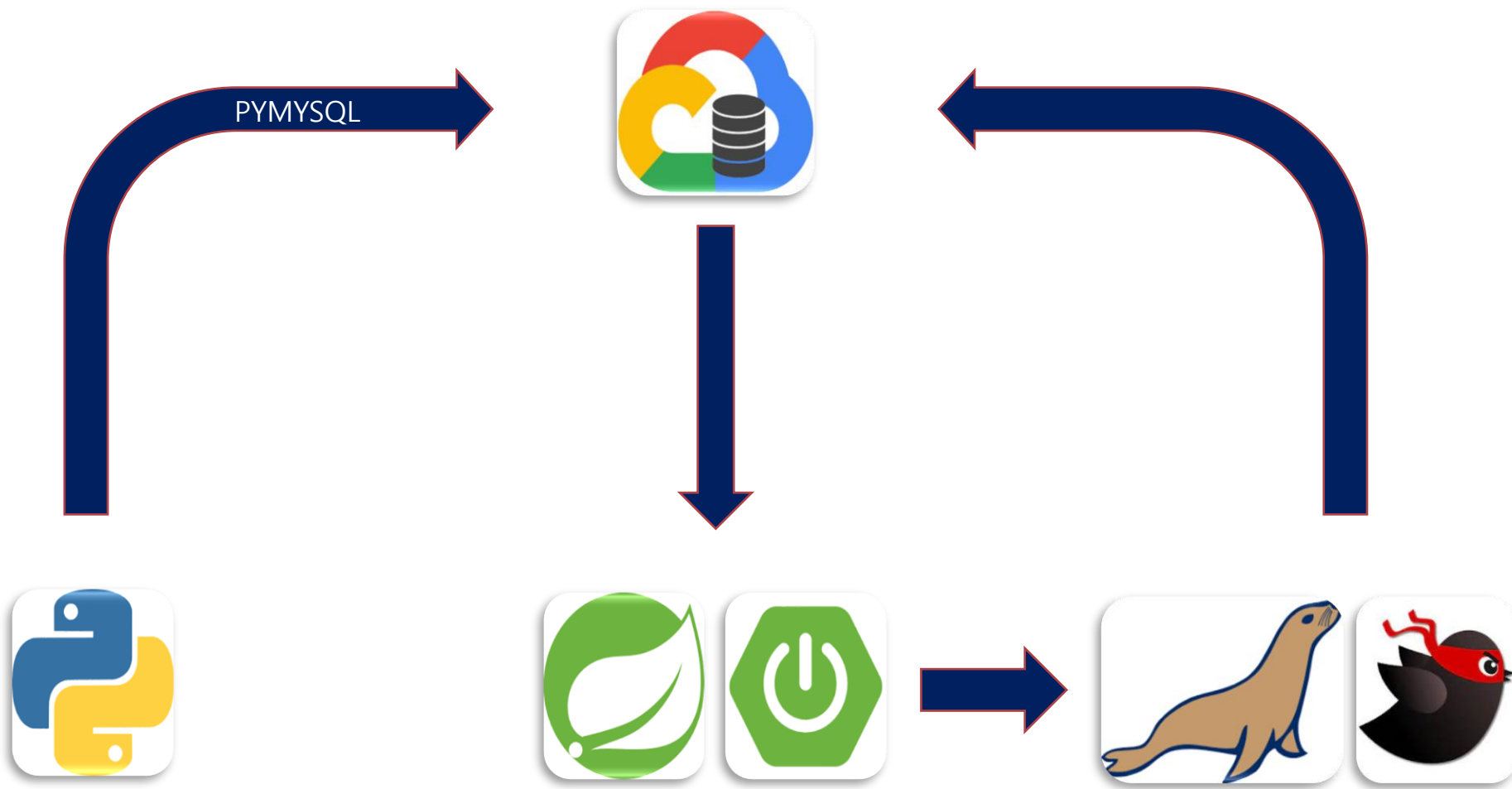
경쟁심리를 통한 플랫폼 사용 유도 및 실력 향상

[illegible]

02. 기획 및 개발 준비(개발환경)



03. 개발 과정 (Front-end / Back-end Overview)



version

Python : 3.8

Java : 11.0.16

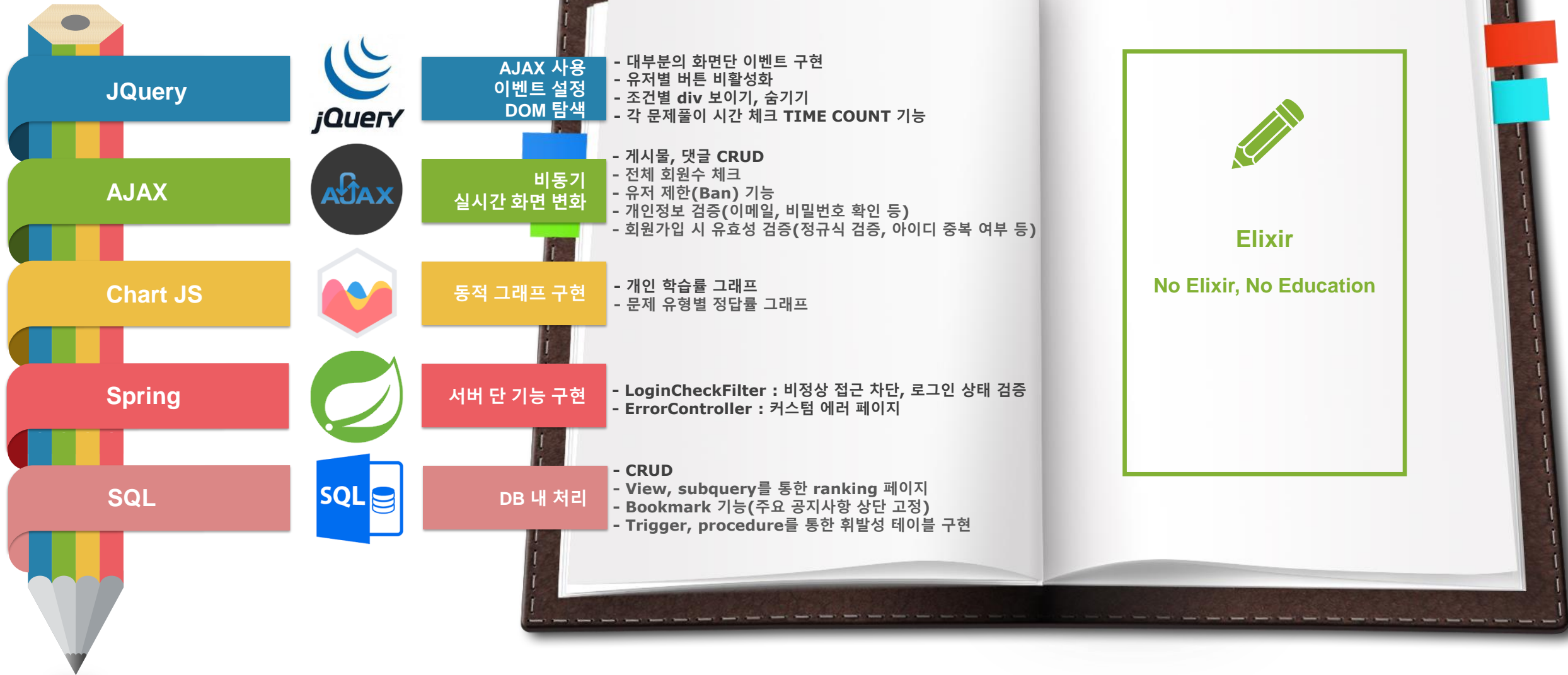
Spring : 2.7.5

MariaDB : 10.6.11

Mybatis : 3.4.6

CentOS : 7

03. 개발 과정(Back-end Overview)



03. 개발 과정(Back-end : JQuery)

- 로그인에 따른 보이기/숨기기

로그인 전

로그인 후

- 각종 이벤트 (클릭 시 색 변화 등)

문제를 풀 때

- 문제 체크 시 색 변화(파랑)
- 풀 문제 색 변화(초록)

- 각종 이벤트 (클릭 시 색 변화 등)

본인 게시물 O : 수정/삭제 보이기

본인 게시물 X : 수정/삭제 숨기기

03. 개발 과정(Back-end : AJAX)

• 개인정보 검증 : 비밀번호 찾기



비밀번호를 잊어버리셨나요?

찾으시려는 비밀번호의 전자우편을 입력하세요.

전자우편 주소 *

aaa@test.com



존재하지 않는 아이디입니다.

비밀번호 변경



비밀번호를 잊어버리셨나요?

찾으시려는 비밀번호의 전자우편을 입력하세요.

전자우편 주소 *

test@test.com



변경하실 비밀번호를 입력하세요.

새로운 비밀번호 *

새로운 비밀번호 확인 *

비밀번호 변경

• 개인정보 검증 : 회원 탈퇴

회원 탈퇴

- 회원 탈퇴시 저장된 학습데이터는 모두 삭제됩니다
- 회원 탈퇴시 서비스를 더 이상 이용하실 수 없습니다.
- 회원 탈퇴시 사용하셨던 전자우편으로 재가입이 불가능 합니다.

☒ 동의합니다

회원 탈퇴를 원하시면 비밀번호를 입력하세요

.....

기존 비밀번호와 같지 않습니다.

회원 탈퇴

회원 탈퇴

- 회원 탈퇴시 저장된 학습데이터는 모두 삭제됩니다
- 회원 탈퇴시 서비스를 더 이상 이용하실 수 없습니다.
- 회원 탈퇴시 사용하셨던 전자우편으로 재가입이 불가능 합니다.

☒ 동의합니다

회원 탈퇴를 원하시면 비밀번호를 입력하세요

.....

회원 탈퇴

03. 개발 과정(Back-end : AJAX)

• 회원가입 시 유효성 검증1

이름 *

더 작은

별명

아카데미

전자 우편 *

test1

이메일 형식에 맞게 작성해 주세요.

국적 *

나라를 선택해주세요 !

암호 *

.....

8자이상 영문 대 소문자, 숫자, 특수문자를 사용하세요.

암호 확인 *

코스 확인

TOPIK I
TOPIK II

Sign Up

이미 계정이 있나요 ? 여기서 로그인 하세요

• 회원가입 시 유효성 검증2

이름 *

더 작은

별명

아카데미

전자 우편 *

test@test.com

누군가 이 아이디를 사용하고 있어요.

국적 *

나라를 선택해주세요 !

암호 *

.....

암호 확인 *

...

코스 확인

TOPIK I
TOPIK II

Sign Up

이미 계정이 있나요 ? 여기서 로그인 하세요

• 회원가입 시 유효성 검증3

이름 *

더 작은

별명

아카데미

전자 우편 *

the@test.com

사용 가능한 아이디예요.

국적 *

대한민국

암호 *

.....

암호 확인 *

.....

코스 확인

TOPIK I
TOPIK II

Sign Up

03. 개발 과정(Back-end : AJAX)

• 유저 Ban 기능

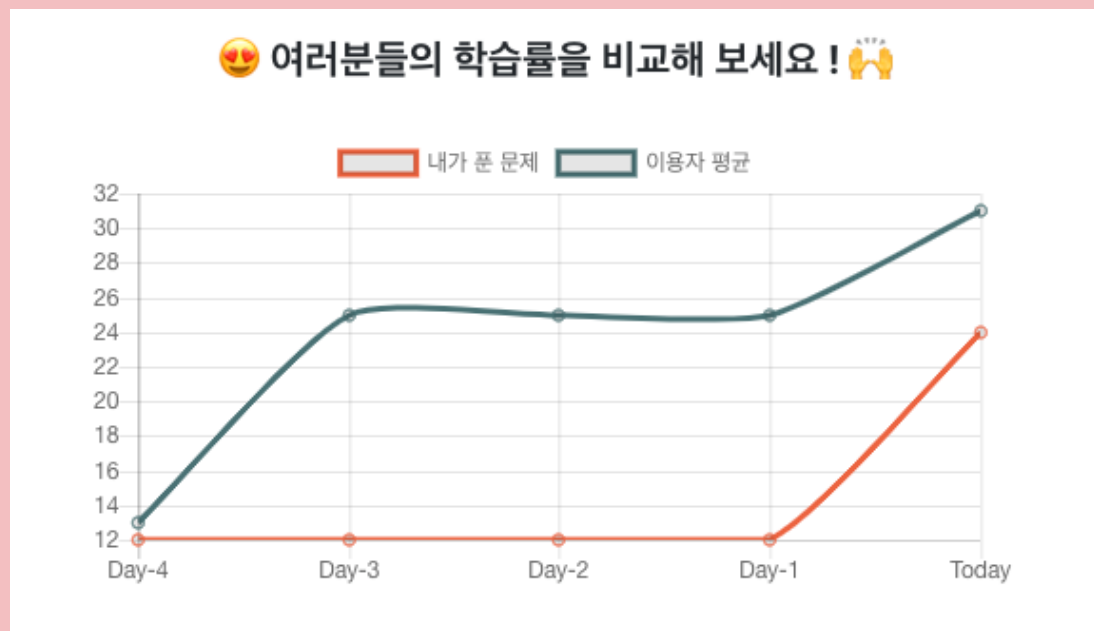
QQQ@1234.COM	ㅇㅇㅇㅇ	바바바	0	0	D+8885 (2022-12-26)	B	C	Q	BC	A	X	View
test@test.com	찬웅	찬웅	0	15	D+8886 (2022-12-25)	B	C	Q	BC	A	X	View
admin@cems.com	admin	admin	0	0	D+8886 (2022-12-25)	B	C	Q	BC	A	X	View

• Ban 유저 게시물 차단 상태

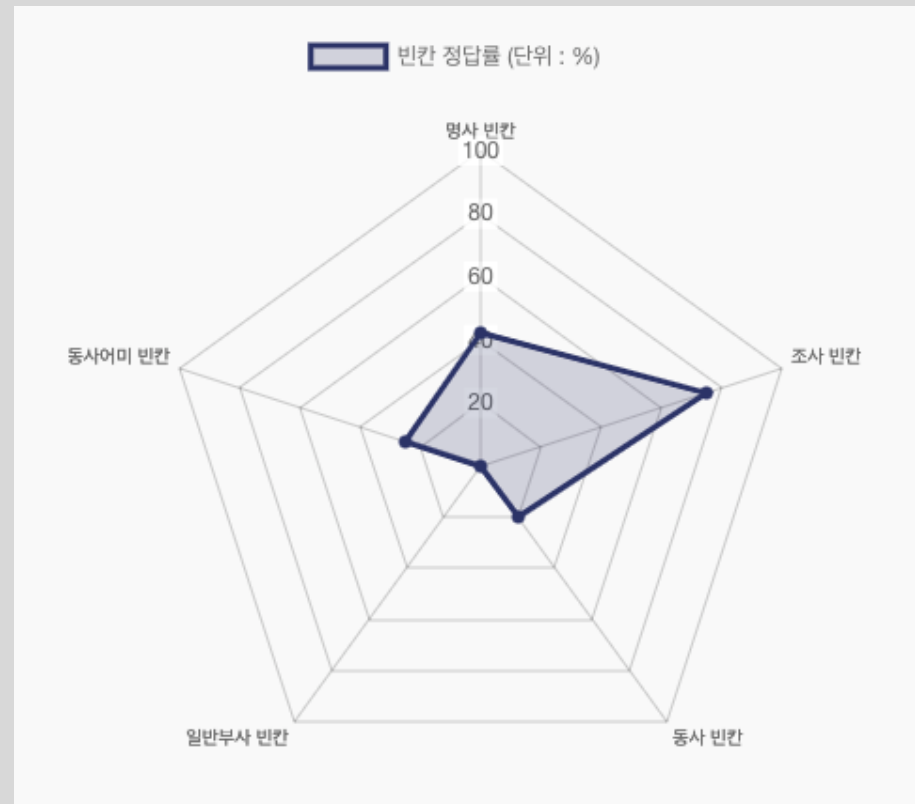
제목		작성자	작성일	조회수
공지	게시판 욕설시 이용 제한 하겠습니다.	admin	2022-12-29 07:25:10	7
1	빈칸 문제를 모르겠는데 알려주실분!!!	응웅	2023-01-11 02:24:15	2
2	이 게시물은 운영자에 의해 차단된 게시물 입니다.	찬웅	2022-12-28 09:26:06	5

03. 개발 과정(Back-end : Chart JS)

- Linear graph : 이용자 평균 vs 개인 평균



- Rader chart : 문제 유형별 정답률 확인



03. 개발 과정(Back-end : Spring)

• LoginCheckFilter

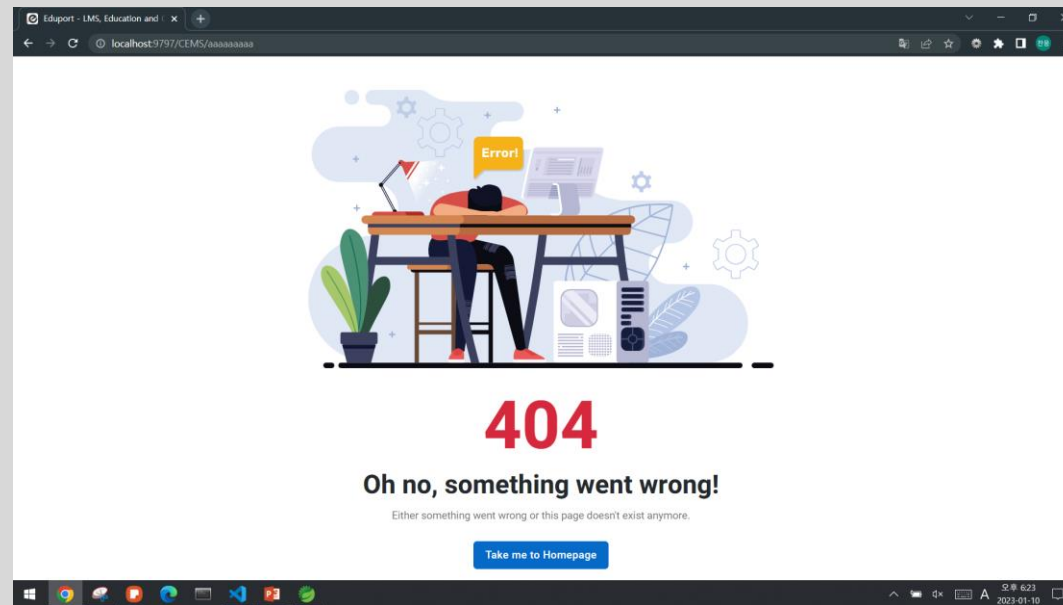
정상적인 접근 시

```
INFO 9876 --- [nio-9797-exec-2] c.c.application.filter.LoginCheckFilter : 인증 체크 필터 시작  
INFO 9876 --- [nio-9797-exec-2] c.c.application.filter.LoginCheckFilter : 인증 체크 필터 종료
```

비정상적인 접근 시 => 로그인 페이지로 이동

```
INFO 9876 --- [nio-9797-exec-9] c.c.application.filter.LoginCheckFilter : 인증 체크 필터 시작  
INFO 9876 --- [nio-9797-exec-9] c.c.application.filter.LoginCheckFilter : 비정상적인 접근  
INFO 9876 --- [nio-9797-exec-9] c.c.application.filter.LoginCheckFilter : 인증 체크 필터 종료
```

• ErrorController




03. 개발 과정(Back-end : SQL)

- View, Subquery를 통한 랭킹 시스템 구현

유저 랭킹

TOP 2

✓ 3




찬찬
@찬찬

Afghanistan

TOP 1

✓ 15




찬웅
@찬웅

Afghanistan

TOP 3


✓ 0



테스트
@업니다

Korea, Republic of

모든 유저


순위	회원	맞은 개수	나라
1	 찬웅 찬웅	15 <small>↑ today+</small>	Afghanistan
1	 찬웅 찬웅	15 <small>↑ today+</small>	Afghanistan
2	 찬찬 찬찬	3 <small>↑ today+</small>	Afghanistan
3	 테스트 테스트	0 <small>↑ today+</small>	Korea, Republic of

- Bookmark 기능 : 주요 공지사항의 상단에 고정

공지사항

Home

조회 검색 정렬 글쓰기


제목	작성자	작성일	조회수
 게시판 육설시 이용 제한 하였습니다.	admin	2022-12-29 07:25:10	7

1 2 >

자유 게시판

Home

조회 검색 정렬 글쓰기

제목	작성자	작성일	조회수
 공지 게시판 육설시 이용 제한 하였습니다.	admin	2022-12-29 07:25:10	7
1 반란 음모를 모른것인데 알려주실분??	웅웅	2023-01-11 02:24:15	2
2 이 게시물은 운영자에 의해 차단된 게시물 입니다.	찬웅	2022-12-28 09:26:06	5

1 2 >

04. 결과(개체명 분석기)

• 타 기관 개체명 분석기

제29회 한글 및 한국어 정보처리 학술대회 논문집 (2017년)

한국어 특질을 고려한 단어 벡터의
Bi-LSTM 기반 개체명 모델 적용

남석원¹, 함영균, 최기진
한국과학기술원
obiwan96@kaist.ac.kr, hahmyg@kaist.ac.kr, kschoi@kaist.ac.kr

Application of Word Vector with Korean Specific Feature
to Bi-LSTM model for Named Entity Recognition

Sukhyun Nam¹, Younggyun Hahm, Key-Sun Choi
KAIST

KAIST
원리 : BiLSTM-CRF
2016 국어 정보 처리 경진대회
dataset
학습데이터 : 3,555문장
실험데이터 : 501문장
5개 개체명 분석 표지
F1 score : 84.19%

제31회 한글 및 한국어 정보처리 학술대회 논문집 (2019년)

CNN-CRFs를 이용한 한국어 개체명 인식기

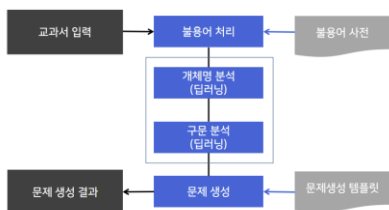
유연수¹, 박혁로
전남대학교 전자정보공학부
powinz001@naver.com, hyukro@jnu.ac.kr

Korean Named-entity Recognition Using CNN-CRFs

Yeon-Soo You¹, Hyuk-Ro Park
Department of Electronics & Computer Engineering, Chonnam National University

전남대학교
원리 : CNN-CRF
해양대학교 배포 개체명 인식용 말뭉치
총 2,000개 문장
10개 개체명 분석 표지
F1 score : 83.6%

최종 문제 생성 과정



NHN
원리 : BiLSTM-CRF
세종코퍼스, 기 구축 개체명 사전
학습데이터 91,525개 문장
F1 score : 94.35%

• Elixir 개체명 분석기

F1-score: 비공개

	precision	recall	f1-score	support
CG_BAY	0.56	0.20	0.30	45
CG_CONTINENT	0.86	0.16	0.27	76
CG_ISLAND	0.65	0.38	0.48	267
CG_MOUNTAIN	0.55	0.39	0.45	317
CG_OCEAN	0.35	0.12	0.18	57
CG_RIVER	0.66	0.43	0.52	222
CP_CAPITALCITY	0.54	0.45	0.49	176

원리 : 비공개
특징 : 비공개

잘 정제된 깨끗한 data(연구목적)

Vs

특수문자 및 인터넷 용어 등이 포함된 data(실제 활용 목적)

04. 결과(개체명 분석기)

• 오답 예시 1

단어	실제값	예측값
이	: 0	0
부분	: 0	0
을	: 0	0
,	: 0	0
비로소	: 0	0
제도	: 0	0
가	: 0	0
제대로	: 0	0
갖춰졌기	: 0	0
때문	: 0	0
에	: 0	0
은나라	: 0	DT_DYNASTY
의	: 0	0
소박한	: 0	0
문화	: 0	0
가	: 0	0
주나라	: DT_DYNASTY 0	
에	: 0	0
와서	: 0	0
훨씬	: 0	0
세련되어졌다	: 0	0
.	: 0	0

기대 답안

주나라 : DT_DYNASTY

예측 답안

은나라 : DT_DYNASTY

• 오답 예시2

단어	실제값	예측값
이비에스	: OGG_EDUCATION 0	
스쿨	: 0	0
리포터	: 0	0
김민주	: 0	PS_NAME
입니다	: 0	0
.	: 0	0

기대 답안

이비에스 : OGG_EDUCATION

예측 답안

김민주 : PS_NAME

예측 답안이 기대 답안과는 다르지만 오답이 아니다.
실제 정확도는 F1 score **비공개** 를 상당히 뛰어넘을 것이라 예상

04. 결과(구문 분석기)

- Non-Multiprocessing

100% | 150082/150082 [2:23:48<00:00, 17.39it/s]

Total 2시간 23분 48초(5,028초)

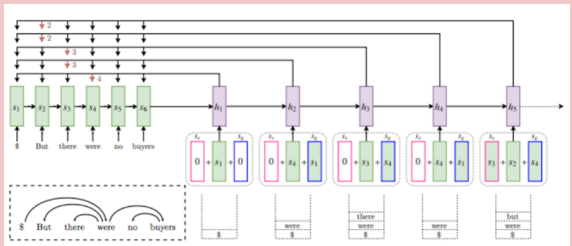
- Multiprocessing

9380/9380 [00:53<00:00, 176.30it/s]	9380/9380 [01:12<00:00, 129.54it/s]
9380/9380 [01:28<00:00, 106.28it/s]	9380/9380 [01:11<00:00, 130.90it/s]
9380/9380 [01:23<00:00, 111.70it/s]	9380/9380 [01:18<00:00, 120.11it/s]
9380/9380 [01:34<00:00, 99.68it/s]	9380/9380 [01:22<00:00, 114.31it/s]
9380/9380 [01:41<00:00, 92.82it/s]	9380/9380 [01:21<00:00, 115.47it/s]
9380/9380 [01:33<00:00, 100.24it/s]	9380/9380 [01:22<00:00, 113.81it/s]
9380/9380 [01:30<00:00, 103.58it/s]	9380/9380 [01:18<00:00, 119.12it/s]
9381/9381 [01:18<00:00, 120.08it/s]	9381/9381 [01:19<00:00, 117.93it/s]

Total 16 processes : 약 7분(420초)

Multiprocessing으로 약 12배 단축

- NHN 구문 분석기



원리 : Stack Pointer Network
 세종코퍼스, 기 구축 개체명 사전
 학습데이터 43,569개 문장
 UAS : 83.86%, LAS : 81.73%

- Elixir 구문 분석기



원리 : 비공개
 F1 score : 비공개

말뭉치 규모 3배 이상. 성능 차이 비교필요

04. 결과(오답 생성기)

• 초급 수준

비공개 오답 생성

```
1 txt = 비공개
2 txt1 = 비공개
```

✓ 0.3s

```
1 print(create_wrong.n_wrong(txt))
2 print(create_wrong.n_wrong(txt1))
```

✓ 0.4s

[*주변* *국회* *교회* *노조* *국가* *기업*]
[*친구에서* *지역에서* *노조에서* *교사에서* *대학에서* *고교에서*]

비공개 관련 오답 생성

```
1 txt = 비공개
2 txt1 = 비공개
```

✓ 0.2s

```
1 print(create_wrong.ve_wrong(txt))
2 print(create_wrong.ve_wrong(txt1))
```

✓ 2.7s

[*압니다* *잡니다* *좋습니다* *좋아합니다* *만듭니다* *가져갑니다*]
[*다지고* *가져가고* *넘어가고* *들어오고* *해보고* *위하고*]

비공개 오답 생성

```
1 txt = 비공개
2 txt1 = 비공개
```

✓ 0.3s

```
1 print(create_wrong.j_wrong(txt))
2 print(create_wrong.j_wrong(txt1))
```

✓ 2.8s

[*에서* *부터* *한테* *에게서* *로써* *로부터*]
[*친구으로* *친구에게* *친구한테서* *친구부터* *친구로부터* *친구한테*]

비공개 오답 생성

```
1 txt = 비공개
2 txt1 = 비공개
```

✓ 0.2s

```
1 print(create_wrong.m_wrong(txt))
2 print(create_wrong.m_wrong(txt1))
```

✓ 1.6s

[*이른바* *그러면* *오히려* *그러나* *더구나* *하지만*]
[*주로* *모두* *잠시* *이내* *과연* *도로*]

빈칸 유형 중심 & 4지선다인 TOPIK에는 Elixir 구문분석기 활용이 용이

04. 결과(오답 생성기)

- 중급 이상 수준

비공개 오답 생성

```
1 txt = 비공개
```

✓ 0.4s

```
1 print(create_wrong.r_wrong(txt))
2 # ['내려갈 수', '머우러질 수', '돌아올 수', '다가갈 수', '이루어질 수', '움직일 수']
```

비공개 오답 생성

```
1 txt = 비공개
```

✓ 0.2s

```
1 print(create_wrong.r_wrong(txt))
2 # ['남자가 만든', '국가가 만든', '국회가 만든', '사전가 만든', '이전가 만든', '후보가 만든']
3 # ['친구가 찾아간', '친구가 되살린', '친구가 들여다본', '친구가 들어온', '친구가 넘어간', '친구가 가져간']
```

비공개 오답 생성

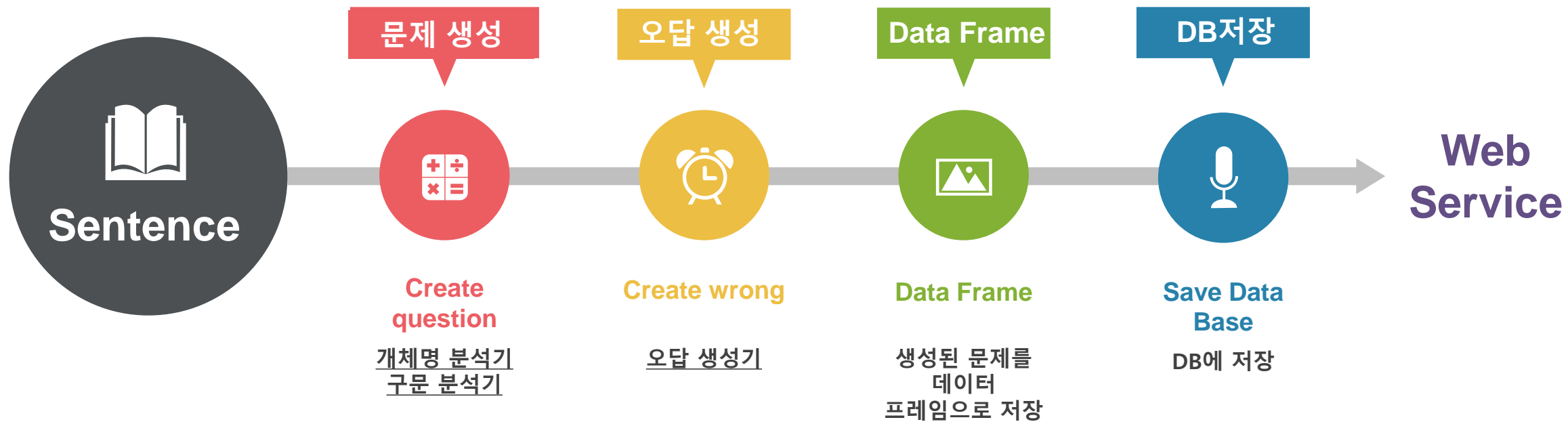
```
1 txt = 비공개
```

✓ 0.2s

```
1 print(create_wrong.r_wrong(txt))
2 # ['오래 정리', '오래 유지', '오래 지급', '오래 공개', '오래 보관', '오래 부과']
3 # ['미어 사용', '통상 사용', '미내 사용', '새로 사용', '이미 사용', '점점 사용']
```

빈칸 유형 중심 & 4지선다인 TOPIK에는 Elixir 구문분석기 활용이 용이

04. 결과(AI 문제 생성기)



Elixir TOPIK Tools

Elixir는 인공지능 기술을 이용하여 다양한 유형의 새로운 문제를 생성합니다.

04. 결과(맞춤법 검사기)

- 쓰기 유형 검사 활용

1.

()에는 답습니다. 겨울에는 춥습니다.

쓰기

0자 / 300자

텍스트를 입력하세요.

다음

문제 번호

1	2	3	4	5	6	7
8	9	10	11	12	13	14
15						

This course includes

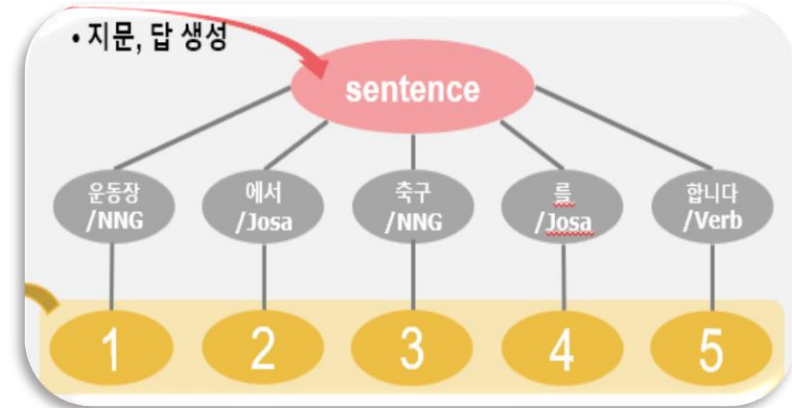
Lectures	30
Duration	4h 50m
Skills	Beginner
Language	English
Deadline	Nov 30 2021
Certificate	Yes

04. 결과(개체명 분석기 & AI 문제 생성기)

F1-score: 비공개

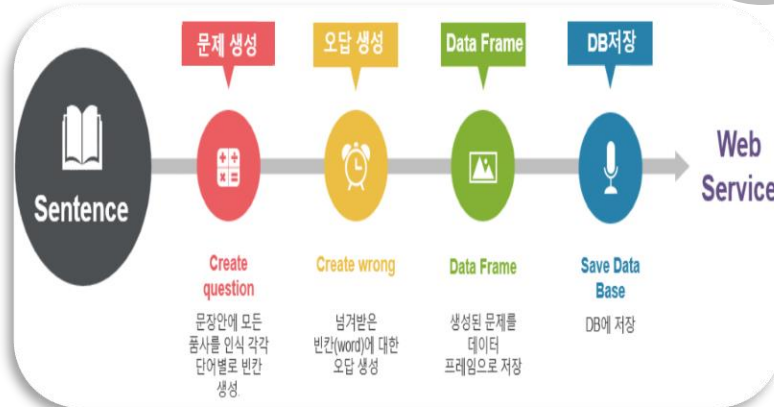
	precision	recall	f1-score	support
NP	0.84	0.72	0.77	3102
NP_AJT	0.00	0.00	0.00	14
NP_CMP	0.59	0.44	0.50	262
NP_CNJ	0.00	0.00	0.00	5
NP_MOD	0.94	0.65	0.77	1756
NP_OBJ	0.68	0.31	0.43	48
NP_SBJ	0.00	0.00	0.00	5

Elixir 개체명*구문 분석기



Elixir AI 문제 생성기

Elixir Technology



Elixir TOPIK Tools

비공개

```

1  # coding: utf-8
2  from TOPIK1.TOPIK1_1_3439_def import TOPIK1_1_3439_def
3
4  text = '밥을 먹습니다. 맛있습니다.'
5  a = TOPIK1_1_3439_def
6  paragraph = a.Josa(text)
7  print(paragraph)

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL

```

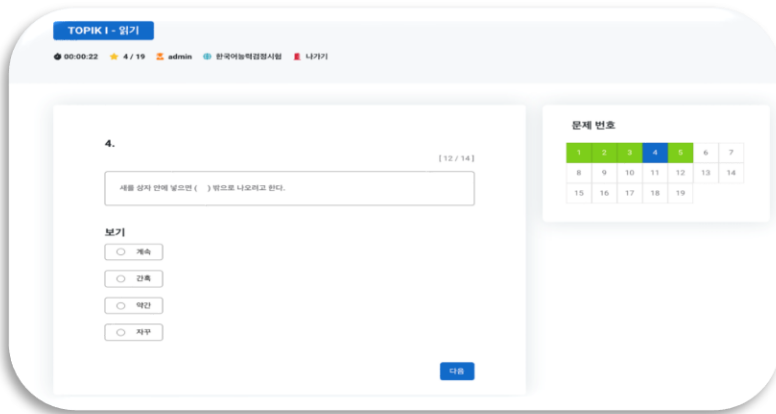
File "c:\dev\python-dev\AI-dev\workspace\jython\test
paragraph = a.Josa(text)
NameError: name 'text' is not defined

(nlp-dev) C:\dev\python-dev\AI-dev\workspace> c: && co
ns\ms-python.python-2022.20.1\pythonFiles\lib\python\
밥( ) 먹습니다. 맛있습니다.

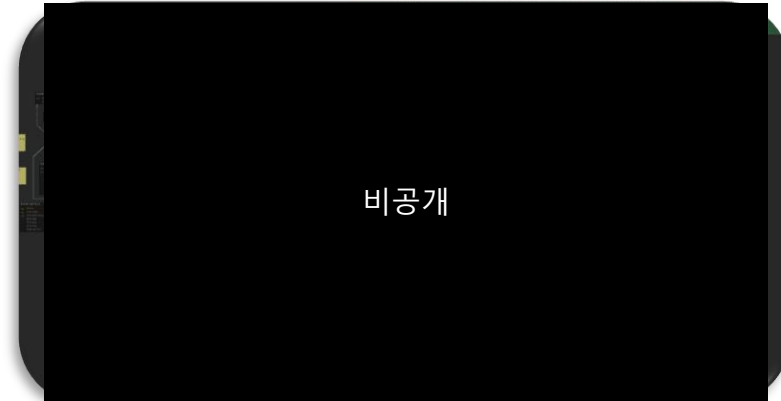
```

Elixir Open API

04. 결과(Front-end / Back-end)

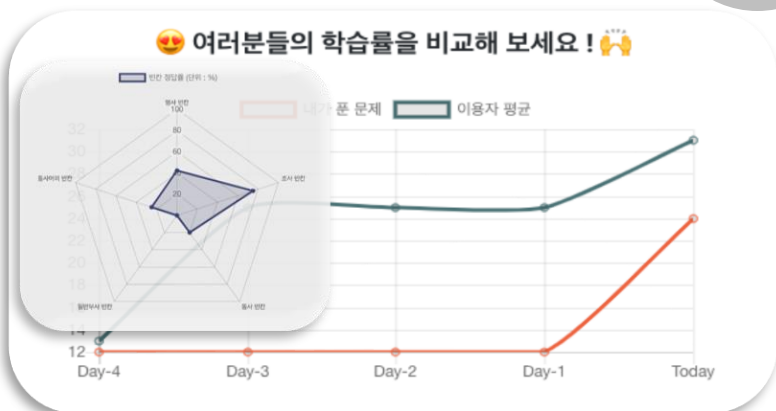


문제 유형별 학습 시스템

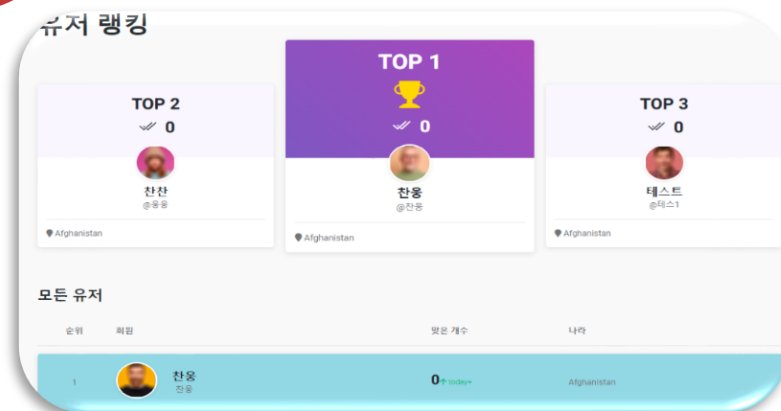


Database 구축

Elixir Platform



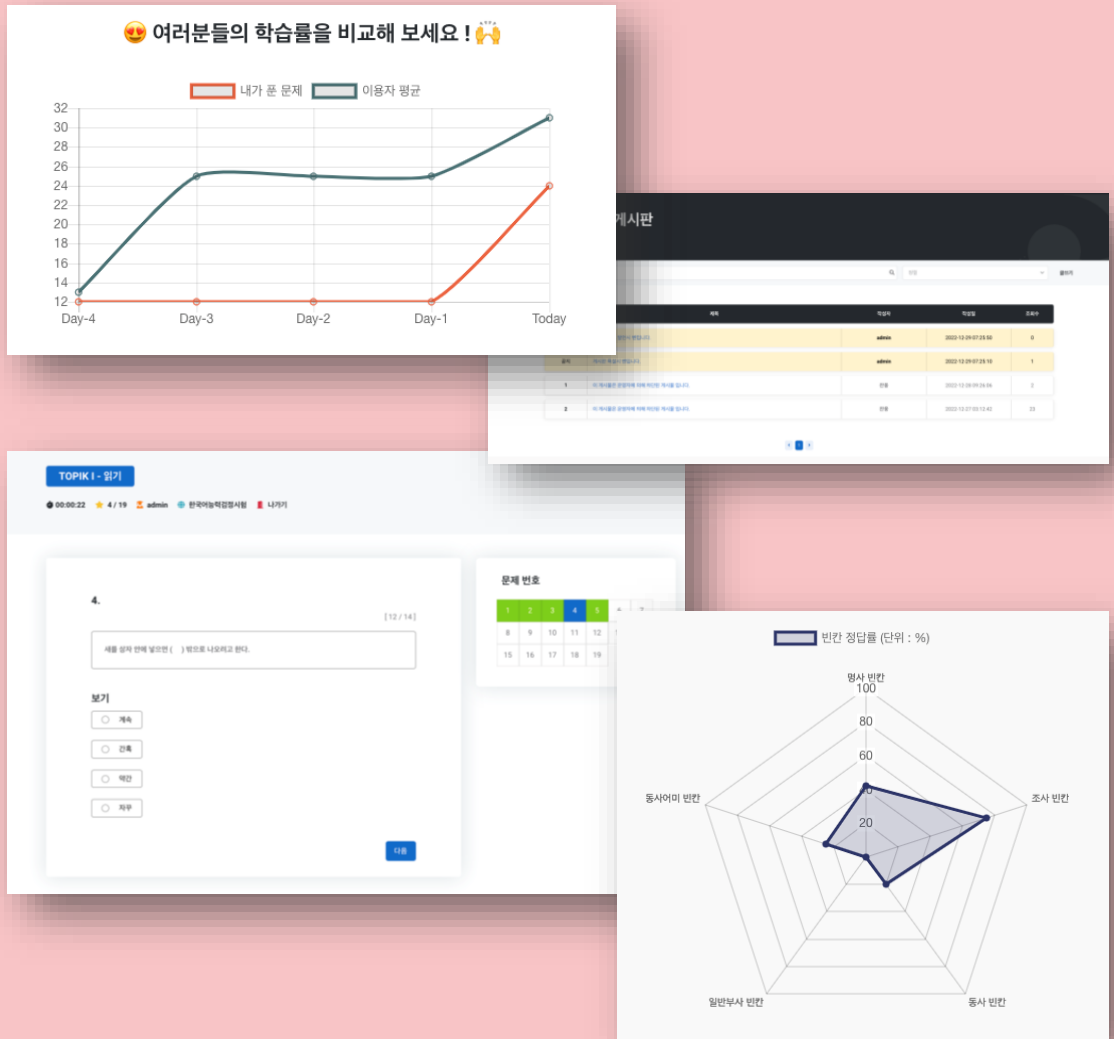
학습관리 시스템



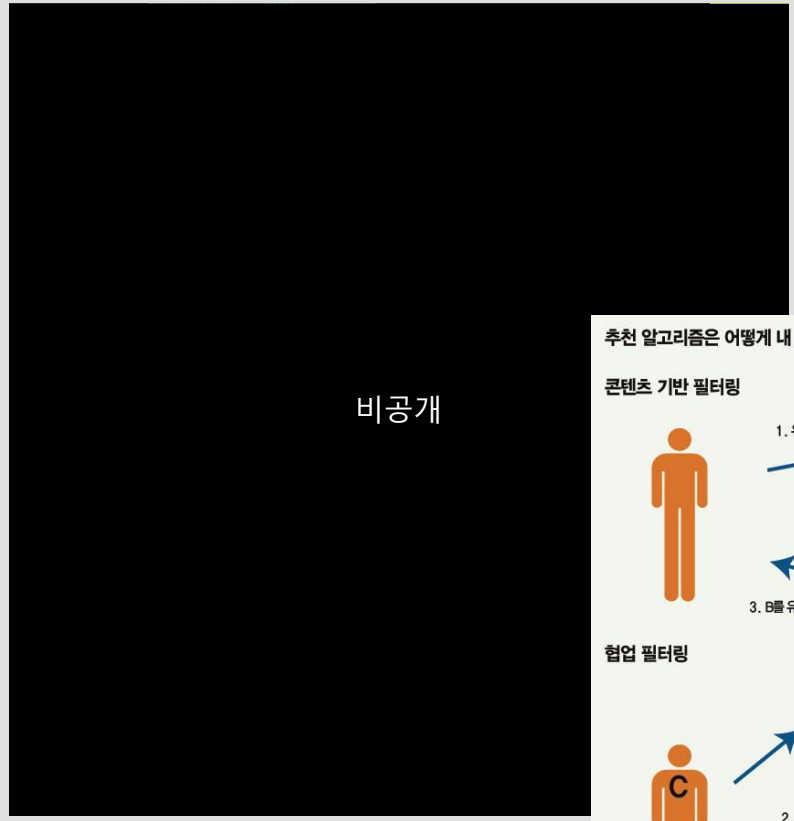
순위 경쟁 시스템

04. 결과(Database)

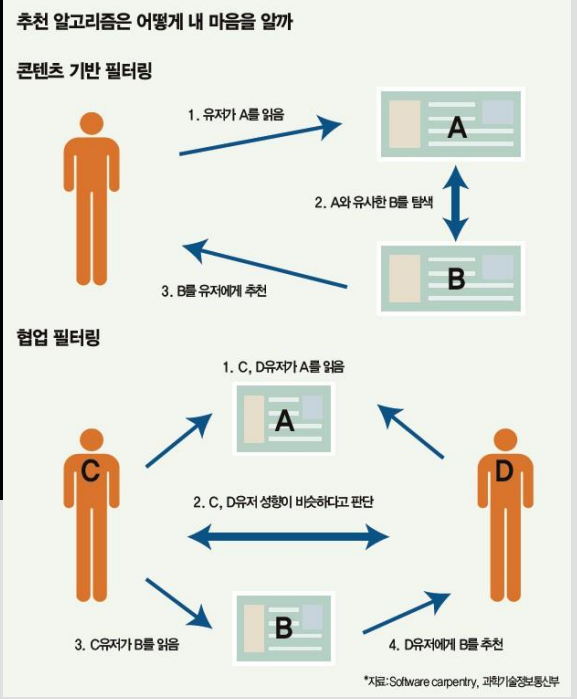
Front-end/Back-end와 연동 및 구현 완료



문제추천시스템 개발을 위한 db 준비 완료



비공개




04. 결과(Overview)

파이프라인					
분야	구분	연구	개발	연구내용	협력기관(예상)
개체명 분석기	AI(자연어 처리)	모델 제작 완료		개체명 인식 모델 제작(완료) 성능 보완(진행중)	대학 산학협력단
구문 분석기	AI(자연어 처리)	모델 제작 완료 (서비스 가능 단계)		구문 분석 모델 제작	대학 산학협력단
TOPIK AI 문제생성기	AI(자연어 처리)	Class 구현 완료 (서비스 가능 단계)		문제 자동 생성 알고리즘(완료) 성능 보완(진행중)	대학 산학협력단, 국내외 기관
문제추천시스템	AI(자연어 처리)	DB구축 완료		유저의 부족한 유형을 파악한 AI 문제추천시스템	대학 산학협력단
Elixir platform	Front-end/Back-end	필수기능 구현 완료		AI 교육 플랫폼(완료) 성능 보완 및 기능 추가(진행중)	대학 산학협력단, 한국어학당, 국내외 기관


“No Elixir, No Education”

Elixir platform




AI 교육 platform
Database에 따라 다양한 분야에 적용

Elixir Open API



TOPIK 자동 문제생성 알고리즘

Elixir 문제추천시스템



학습자의 부족한 부분을 파악하여
효율적인 문제풀이 학습이 이루어
지도록 함

THANK YOU

“No Elixir, No Education”

