**채용 공고 메일링 시스템 기술 보고서**

2024.11.21  
 디지털러닝팀 기술파트

**개요**

회사 주관 국비 교육 프로그램(K-Digital Training)(이하 KDT) 프로그램에 참여하는 교육생들에게 적합한 채용 공고를 제공하기 위해, 과정별 담당 매니저들은 많은 시간과 노력을 들여 직접 채용 사이트를 방문하여 데이터를 수집하고 있다. 그러나 **1만건 이상의 데이터 속에서 과정에 맞는 공고를 선별하는 작업은 매니저들에게 상당한 업무 피로를 초래**할 수 있다.

이에 따라, 채용 공고 사이트에서의 **데이터 수집 과정을 자동화**하고, 수집된 정보를 **KDT 교육생 및 수료생에게 이메일로 전송하는 시스템**을 구축하고자 한다. 이를 통해 KDT를 운영하는 매니저의 업무를 경감하고, 프로그램의 수료생들에게 보다 적합한 **사후 관리 서비스를 제공**함으로써 수료생의 직무 적합성을 높이고 성공적인 취업을 지원할 수 있을 것이다.

**요약**

본 프로젝트에서는 교육생들에게 정확한 채용공고 정보를 효율적으로 제공하기 위하여 **채용 사이트를 두 곳에서 네 곳으로 확대**하고, 스프레드 시트에서 작업하던 수강생 정보를 **학습 관리 시스템과 연동**하였다. 또한, 교육생들의 사후관리를 위해 **과정 시작일과 종료일을 기준으로 메일 자동화**를 구축하였으며 **메일을 발송하기 전 학생의 교육 상태 여부와 채용공고의 일부 URL을 다시 검토**하여 마감일이 지난 공고는 제외하였다. 마지막으로 사용자 행동 패턴을 파악하기 위해 메일에 **수신거부 설문 조사, 피드백 설문 조사 기능**을 추가하고, **AWS SES(Send Email System)와 CloudWatch 프로그램을 이용하여 메일 로그 추적**이 가능하도록 구현하였다.

**R&R**

| 역할 | 담당자 |
| --- | --- |
| 채용 사이트 별 데이터 수집 | 전체 |
| 메일 트래킹 & 대시보드 & 기능 구현 | 김\*\* |
| 메일링 시스템 구축 & html | 박\*\* |
| DB & 쿼리 & 백업 관리 | 장\*\* |

**데이터 수집 및 자동화**

* **수집 Tool** : Selenium, BeautifulSoup
* **담당 채용 공고 사이트**

1. 김\*\*: 인크루트, 점핏
2. 박\*\*: 원티드
3. 장\*\*: 사람인

* **자동화 : Linux crontab**

1. 각 사이트 마다 실행되는 날짜와 시간을 다르게 설정.

| 채용 사이트 | 요일 | 시간 |
| --- | --- | --- |
| 인크루트 | 월, 목 | 오전 1시 |
| 점핏 | 화, 금 | 오전 1시 |
| 원티드 | 수, 토 | 오전 1시 |
| 사람인 | 목, 일 | 오전 3시 |

1. 메일 발송 **주 2회** (화요일, 금요일 17시)
2. 공고 중복 방지를 위해 사용하는 **json 파일 두 달 간격 1일 자정에 삭제**

**메일 보내기**

* **공고 데이터 검토**

1. 채용공고 정보를 설정한 조건에 따라 필터링(자세한 조건은 별첨 참고)
2. **과정 조건 검사** : 각 과정 **시작일 기준 4.5(4개월 15일)개월 후 부터 종료일 기준 6개월 이전**까지 메일 자동 발송, 그 이후 메일 발송 거부
3. **학생 상태 조건 검사**: **취업, 중도 포기, 수강철회**는 보내지 않고 **교육 중, 수료, 조기 수료**는 메일 보냄
4. **공고 검사** : **마감된 url 공고 확인**을 위해 각 사이트 별로 10개씩 검사
5. **과정 별 공고 개수 수정: 빅데이터/풀스택은 사람인 10개, 인크루트, 원티드, 점핏은 5개를 선정하며 PM은 공고 사이트 상관없이 12개 선정.**

* **이메일 발송**

1. 메일 부분에 수집한 데이터(채용공고명, 회사명, 모집기간, 채용사이트)를 표시하고 수신거부, 설문조사 링크는 마지막 부분에 삽입
2. **피드백 설문조사** 링크 삽입
3. **수신거부 설문조사** 링크 삽입
4. **수신거부 버튼 클릭** 시 **구글 설문조사 폼**으로 이동
5. 폼 작성 완료 후 **구글 스프레드 시트에 이름/메일이 입력**됨
6. 메일 발송 전 스프레드 시트를 확인해 메일 수신 여부 결정
7. **안내문구 출력:** 총 공고수가 20개 미만 시 **“최신 공고 제공으로 채용공고가 20개 미만일 수 있습니다.”** 라는 문구 삽입
8. **최종 이메일 송부**
9. 회사 메일은 트레킹을 위한 구성세트 추가 시 차단되어 **chunjae 메일은 구성세트를 사용하지 않고, 그 외는 사용하여 메일 트레킹 유지**
10. json 파일을 불러와 직전에 보낸 공고를 제외하여 채용공고 송부 및 보낸 공고를 json 파일에 저장

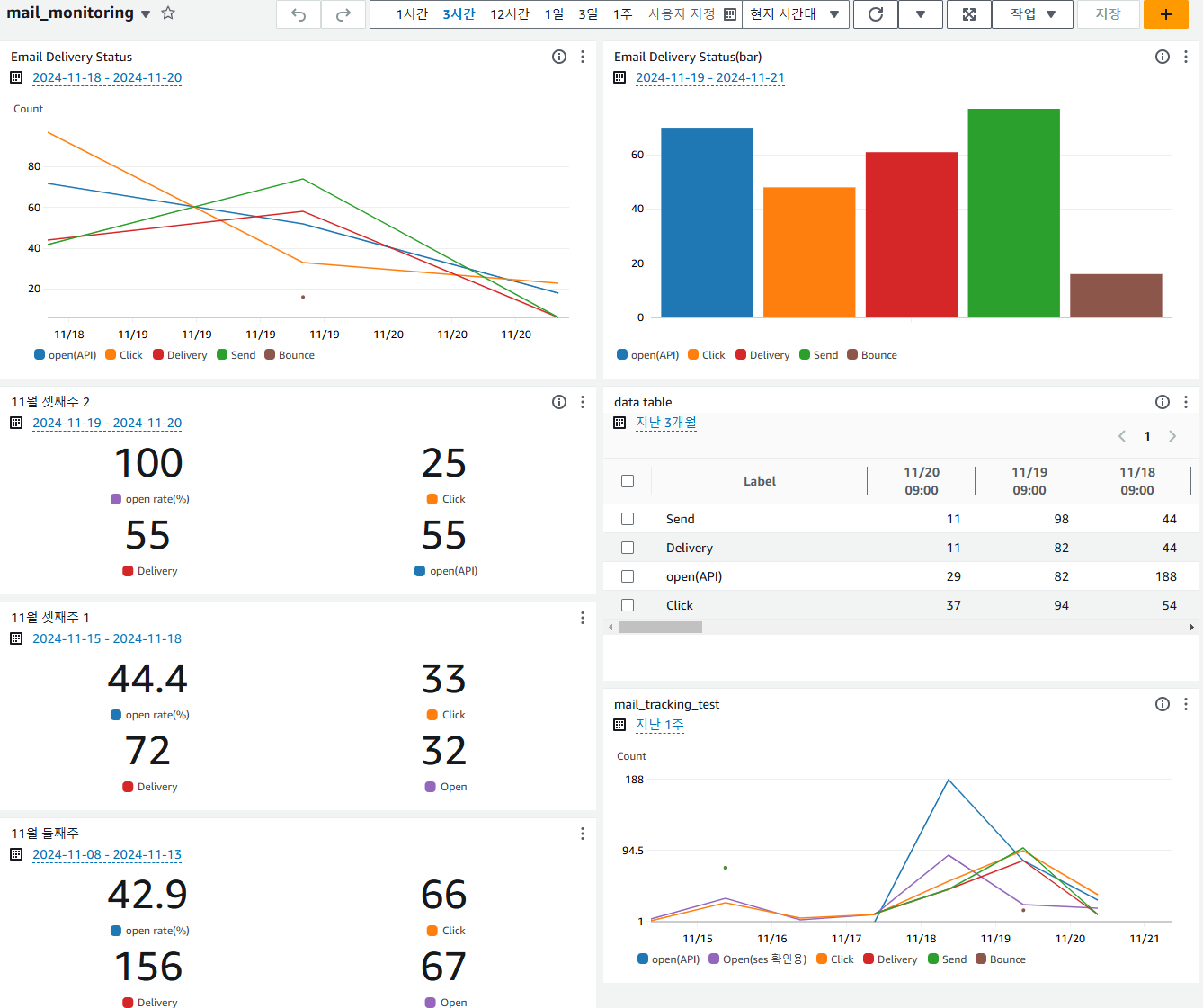
**메일 트레킹(AWS CloudWatch)**

* **AWS SES, Cloudwatch 연결**

1. 메일 로그 추적을 위한 SES 구성 세트 설정(observe - mail\_monitoring)
2. 대시보드 구축을 위한 이벤트 대상 CloudWatch 설정
3. SES에서 제공하는 open 이용 시 네이버에서 차단되어 로그가 찍히지 않는 현상 발생 -> **API Gateway와 lambda함수를 연결하여 천재교육 로고 이미지가 로드될 때마다 요청**을 보내 메일 열람 로그를 찍을 수 있게 구현 -> Cloudwatch와 연결결

* **대시보드**

| **항목** | **추적 방식** | **기타 설명** | **천재 메일 로그** |
| --- | --- | --- | --- |
| **open(API)** | Lambda + API Gateway | 모든 메일 열람 시 카운팅 (이미지 요청 시마다 로그 기록) | O |
| **click** | aws.track.me  (URL 변경 추적) | 메일 내 링크 클릭 여부 추적 | X |
| **bounce** | SES 로그 | 반송 여부 확인 (1~2회 재발송 시도 후 실패한 메일) | X |
| **send** | SES 로그 | 발송 성공 여부 확인 | X |
| **delivery** | SES 로그 | 메일함 도착 여부 확인 | X |

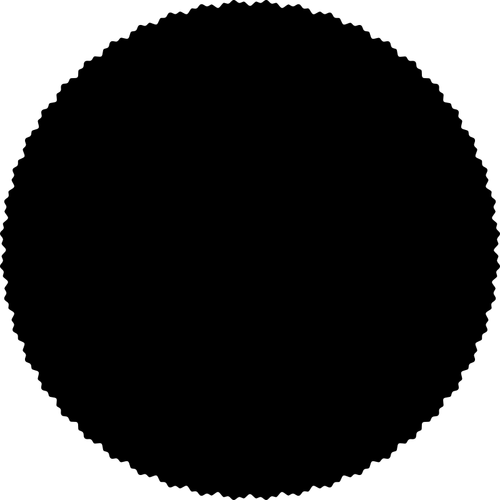


<대시보드 기능 구현>

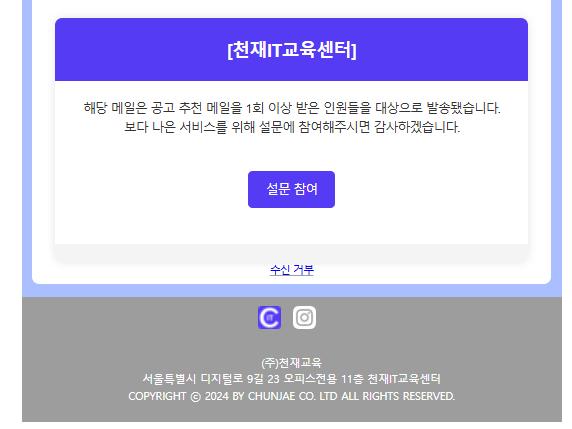
**s3백업**

1. **s3백업 주기**는 **일주일(매주 일요일 12시)**로 crontab을 통해 **자동 백업**이 이루어지도록 설정.
2. **대상**: **크롤링 코드**와 **쿼리 파일**, DB에 적재된 **데이터**
   1. 인크루트, 점핏, 원티드, 사람인 **py파일(코드) 및 쿼리(sql) 파일을** s3에 적재.
   2. 크롤링을 통해 수집되어 DB에 적재된 채용공고 **데이터는 sql 파일로 변환**하여 s3에 적재.
3. **s3**에 적재된 후 **3개월이 지난 데이터**에 대해서는 s3내 수명 주기 규칙을 통해 **자동**으로 **삭제**되도록 설정.

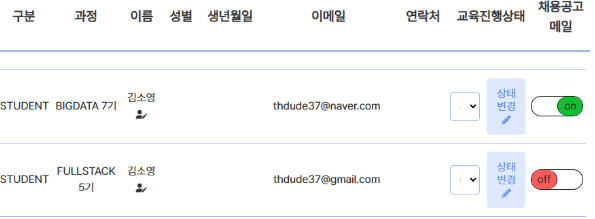
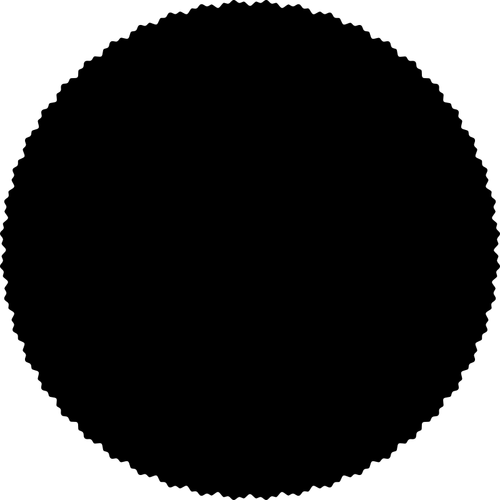
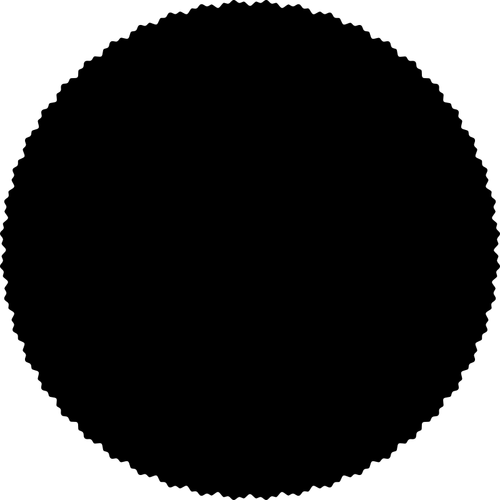
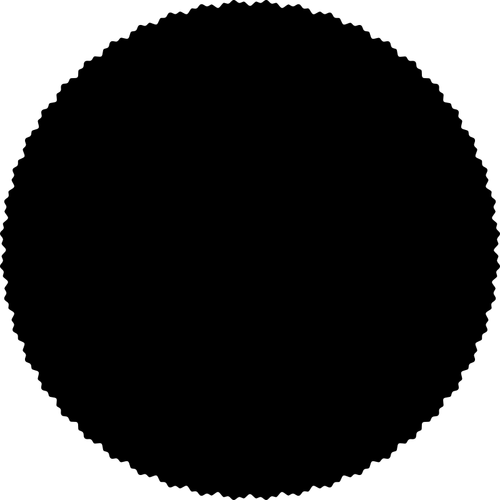
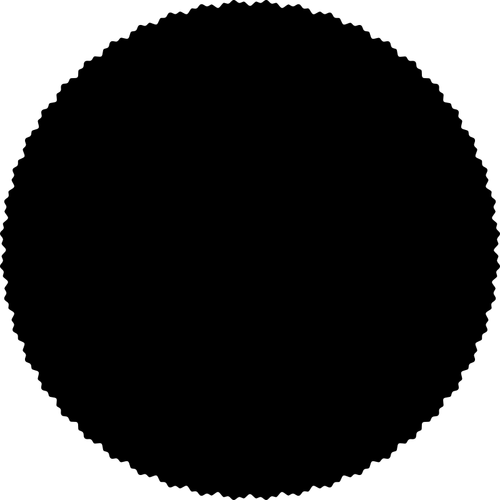
**채용공고 메일 결과화면**



<채용공고 화면(클릭 시 해당 사이트로 넘어감)>



<설문조사 및 수신거부>



<학습관리 시스템과 연동하여 메일 수신 on/off 기능 구현>

**별첨**

**공고 명 키워드**

| 과정 명 | 키워드 |
| --- | --- |
| 빅데이터 | 데이터 사이언티스트, 데이터 엔지니어, 데이터 분석가, 데이터사이언티스트, 데이터엔지니어, 데이터분석가, 데이터 분석, 데이터분석, 데이터 정제, 데이터정제, 데이터 처리, 데이터처리, ai 기획, AI 기획, ai기획, 데이터 관리, 데이터 분석 매니저, 머신러닝, AI 엔지니어, 인공지능 엔지니어, 데이터 시각화, tableau, Tableau, 데이터 마이닝, 비즈니스 인텔리전스, ETL 개발, SQL, R 분석, Hadoop, 에듀테크 콘텐츠 개발, 딥러닝/머신러닝, Data Scientist, Data Engineer, Data Analyst, Machine Learning, AI Engineer, Data Visualization, Business Intelligence, ETL Developer, Big Data Engineer, Data Management, Data Mining, 딥러닝, 자연어 처리, Deep Learning Engineer, Natural Language Processing, NLP, DBA, 빅데이터, AI 모델, DB관리, db관리, 사이언티스트, 인공지능, 데이터분석, 데이터처리, 데이터관리, 데이터마이닝, LLM, Data Architect, 데이터 리터러시 |
| 풀스택 | 프론트엔드, 프론트앤드, 백엔드, 백앤드, 웹 서비스, 모바일 서비스, 풀스택, Java, 웹 개발자, 소프트웨어 엔지니어, 소프트웨어 개발자, 시스템 엔지니어, 시스템 개발자, 모바일 개발자, 앱 개발자, API 개발자, 클라우드 엔지니어, DevOps 엔지니어, DevOps 개발자, 서버 개발자, 네트워크 엔지니어, 네트워크 개발자, Front-End Developer, Front-End, Back-End, Back-End Developer, Full Stack Developer, Software Engineer, Mobile Developer, Cloud Engineer, API Developer, DevOps Engineer, Server Developer, Web Designer, 웹 퍼블리셔, Vue.js, Node.js, Frontend Engineer, Backend Engineer, React Developer, Node.js Developer, Vue.js Developer, CI/CD |
| PM | 웹 기획, 앱 기획, 서비스 기획, 콘텐츠 기획, 전략 기획, 프로덕트 매니저저 |

**제외 키워드**

| 과정 명 | 제외 키워드 |
| --- | --- |
| 빅데이터 | 석사, 박사, 코드잇, 교육생, 국비, 청년수당, 제약, 건설, 연구, 기계, 행정, 호텔, 천재교육, 천재교과서, SeSAC, APM, apm, edi, EDI, 마케터, 양성 |
| 풀스택 | 석사, 박사, 코드잇, 교육생, 국비, 청년수당, 제약, 건설, 연구, 기계, 행정, 호텔, 천재교육, 천재교과서, SeSAC, APM, apm, edi, EDI, **UX, UI**, 마케터, 양성 |
| PM | 창업, 번역, 석사, 박사, 코드잇, 교육생, 국비, 청년수당, 제약, 건설, 연구, 기계, 행정, 호텔, 천재교육, 천재교과서, SeSAC, APM, apm, edi, EDI, **개발자, 시공, 유튜브, 양성** |