

기자 성과 측정 시스템 MVP

서울경제 편집국

시스템 구축 요청서 및 기술 문서

목차

- MVP 목표
- 시스템 아키텍처 (AWS)
- 데이터 소스 및 XML 구조
- 정량평가 배점 체계
- 상대평가 전환 로직
- 화면 구성
- 개발 범위 (Phase 1/2)
- 기술 스택
- 기대 효과
- 로컬 테스트 방법 (MVP 프로토타입)
- IT지원부 요청사항
- 대시보드 미리보기

1. MVP 목표

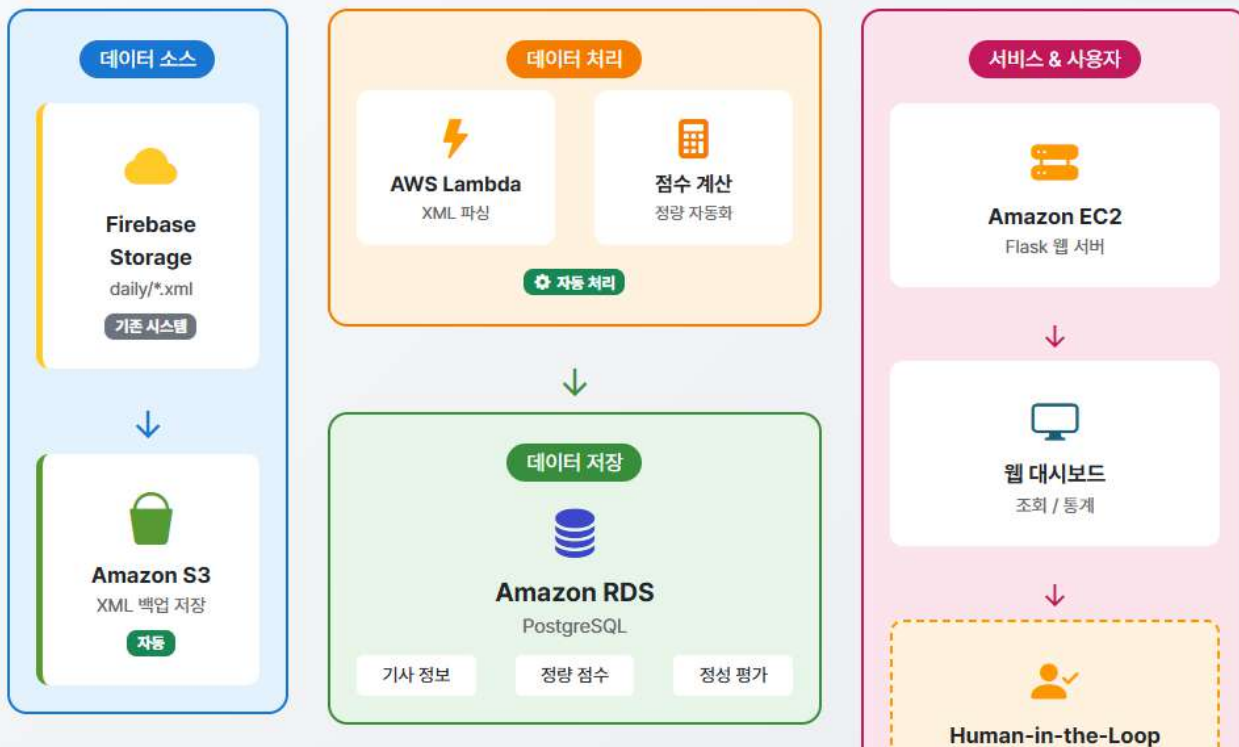
MVP 구축: XML 데이터 자동 파싱, 정량 점수 자동 계산, 정성평가 입력 시스템

시스템명	기자 성과 측정 시스템
목적	기자별 기사 성과를 정량화하여 객관적인 평가 기준 마련
대상 사용자	편집국 부장급 이상 (평가자), 기자 (조회)
데이터 소스	Firebase Storage - daily 폴더 XML 데이터 (기존 시스템에서 매일 자동 생성)

2. 시스템 아키텍처 (AWS)

서울경제 기자 성과 측정 시스템 아키텍처

■ 자동 처리
■ 수동 입력



① XML 데이터 (매일 자동 생성)

② 파싱 & 정량점수 계산

③ DB 저장

④ 조회 & 정성평가 입력

자동 정량 데이터

- 기자명 (writer)
- 면 정보 (pageNumber)
- 글자수 (content 길이)
- 원문 링크 (link)
- 기사 작성 건수
- 조회수 (시스템 연동 시)

수동 정성 데이터

- 단독기사 여부 (데스크 판단)
- 면배치 (툼/사이드/하단)
- 기획기사 여부 (데스크 판단)
- 서경기자상 수상
- 정보보고 점수
- 기명 온라인 점수

3. 데이터 소스 및 XML 구조

```
# Firebase Storage (기존 시스템) - bucket: sedaily-firebase-storage - path: /daily/YYYY-MM-DD.xml # 매일 자동 생성 # XML 데이터 구조 <article> <title>기사 제목</title> <writer>기자명</writer> <pageNumber>1</pageNumber> # 면 정보 <content>본문 내용...</content> # 글자수 계산용 <link>https://...</link> # 원문 링크 <pubDate>2024-10-27</pubDate> </article>
```

테스트 데이터 (샘플)

87

총 기사 수

42

참여 기사

5 / 12 / 70

1면 / 경제면 / 기타

4. 정량평가 배점 체계

구분	비중	수집	점수 체계
조회수	15점	자동	시스템 연동 (추후)
기사작성 건수	20점	자동	XML 파싱으로 자동 집계
단독기사 (면배치)	20점	수동	5단계 (20/15/10/5/0점) - 2개 매체 이상 보도 시 단독 인정 - 면배치: 톱/사이드(좌)/사이드(우)/하단/기타
기획기사	20점	수동	5단계 (20/15/10/5/0점) - 스트+종합면 기준 - 1회 인정 시 5점, 파급력에 따라 가점
서경기자상	5점	수동	금상 5점, 우수상 3점
정보보고	10점	수동	6단계 (10/8/6/4/2/0점) - 주간 평가 → 월평균 - 핵심 정보보고 시 2점 가점
기명 온라인	5점	수동	3단계 (5/3/1점) - 연재물 기본 1점 - 구독자/퀄리티 따라 가점

합계	100점	
----	------	--

💡 향후 고려항목: 출입처 인사 DB 입력

5. 상대평가 전환 로직

```
# 상대평가 변환 공식
def convert_to_relative(scores):
    mean = average(scores) # 평균
    std = stdev(scores) # 표준편차
    target_mean = 85 # 목표 평균
    target_std = 7.5 # 목표 표준편차
    for score in scores:
        z_score = (score - mean) / std
        relative = target_mean + (z_score * target_std)
    return relative_scores
```

변환 예시

순번	이름	입력점수(절대)	최종점수(상대)
1	강OO	100	99.01
2	김OO	97.5	96.27
3	조OO	92	90.26
4	김OO	90	88.07
5	이OO	88.5	86.43
평균		87.19	85.00
표준편차		6.86	7.5

6. 화면 구성

화면 1: 기사 목록 및 평가 입력

면	제목	기자	글자수	단독/배치	기획	정보보고	저장
1	日, 70세 군무 비결...	양종곤	1,842	단독 ▼	선트 ▼	ε ▼	✓

화면 2: 기자별 성과 요약

순위	기자명	기사수	조회수	단독	기획	서경상	정보보고	기명온라인	절대	상대
1	양준호	5	12,450	15	10	5	8	3	92	96.27

7. 개발 범위

Phase 1 (MVP)

- Firebase XML 파싱
- 기사 목록 조회
- 정성평가 입력/저장
- 기자별 점수 집계
- 상대평가 변환

Phase 2 (확장)

- 조회수 시스템 연동
- 서경기자상 DB 연동
- 엑셀 다운로드
- 기간별 리포트
- 출입처 인사 DB

8. 기술 스택 (AWS)


구분	AWS 서비스	용도
----	---------	----

데이터 소스	Firebase Storage → Amazon S3	XML 데이터 수집 및 백업
데이터 처리	AWS Lambda	XML 파싱, 정량 점수 계산
데이터베이스	Amazon RDS (PostgreSQL)	기사/평가 데이터 저장
웹 서버	Amazon EC2 / Elastic Beanstalk	Flask/FastAPI 백엔드
프론트엔드	Amazon S3 + CloudFront	정적 웹 호스팅


9. 기대 효과



업무 시간 절감
주 5시간+
수작업 집계 → 자동화



데이터 정확성
100%
수기 오류 제거



실시간 조회
즉시
기자별/기간별 현황

10. 로컬 테스트 방법 (MVP 프로토타입)

파일 구조

```
├─ run.bat # 실행 스크립트 (더블클릭) ─ requirements.txt # Python 패키지 ─ src/ ── app.py # Flask 백엔드 서버 ── sample_data/ ── 2024-10-27.xml # 샘플 XML 데이터 ── dashboard/ ── app.html # 실제 동작하는 대시보드 ── proposal.html # 이 문서
```

실행 방법

```
# 방법 1: 배치 파일 실행 (Windows) run.bat 더블클릭 # 방법 2: 직접 실행 pip install flask flask-cors python src/app.py # 브라우저 접속 http://localhost:5000
```

주요 기능

XML 새로고침	버튼 클릭 시 src/sample_data/*.xml 파일에서 데이터 다시 로드
기사 목록	XML에서 파싱한 기사 목록 표시 (제목, 기자, 면, 글자수)
평가 입력	단독/배치, 기획, 정보보고 점수 입력 및 저장
기자별 통계	정량/정성 점수 자동 집계, 상대평가 변환

XML 데이터 형식

```
<!-- src/sample_data/2024-10-27.xml --> <articles> <article> <title>日, 70세 근무 비결 '계속고용'에 있다</title> <writer>양종곤 기자</writer> <pageNumber>1</pageNumber> <pubDate>2024-10-27T06:00:00</pubDate> <link>https://www.sedaily.com/NewsView/...</link> <content>기사 본문 내용...</content> </article> ...</articles>
```

① 테스트 방법: Firebase에서 XML 파일을 다운로드하여 src/sample_data/ 폴더에 넣고 "XML 새로고침" 버튼을 클릭하면 실제 데이터로 테스트할 수 있습니다.

11. IT지원부 요청사항

항목	상세 내용
1. Firebase 접근	- daily 폴더 XML 읽기 권한 - 서비스 계정 키(JSON) 발급 - 또는 API 형태로 제공
2. 조회수 연동	- 기사별 조회수 API - 또는 일별 조회수 데이터 제공
3. 웹 서버	- 내부망 웹 서버 배포 - 또는 AWS 클라우드 호스팅

12. 대시보드 미리보기 (샘플)

서울경제 기자 성과 측정

2024-10-27

 87 총 기사 수	 42 참여 기자	 1,842 평균 글자수	 12 평가 완료
--	---	---	---

기간

기자명

2024-10-01



2024-10-31



전체



조회



XML 새로고침

XML 새로고침: src/sample_data/*.xml 파일에서 데이터를 다시 로드합니다.

기자별 성과 요약

순위	기자명	기사수	총글자수	1면	단독	기획	정량	정성	총점
1	양준호	5	12,450	1	15	10	45	47	92
2	송이라	4	9,820	0	10	15	38	42	80
3	윤지영	2	4,230	0	5	5	28	30	58

기사 목록 및 평가

면	제목	기자	글자수	단독/배치	기획	정보보고	저장
1	日, 70세 근무 비결 '계속고용'에 있다	양종곤	1,842	특	선택	8점	✓
1	마지막 여론조사도 동률...美대선 불복리스크	윤홍우	1,523	사이드(좌)	선택	점수	📄
2	'취임 2주년' 이재용, 정의선·日도요타 회동	노우리	2,156	선택	기획 5점	6점	✓
14	폰없이 식사에만 집중...LGU+ '노 폰 다이닝'	윤지영	892	선택	선택	점수	📄

엑셀 다운로드

인쇄

전체 저장

인쇄 / PDF 저장