Software Test Plan

for

Epoch Pulse.

Version 1.0

May 12, 2026

Contents

1 Test Plan Identifier 4

2 Introduction 4
2.1 Objectives 4
2.2 Scope 4
2.3 References 4

3 Test Items 5

4 Features to Be Tested 5

5 Features Not to Be Tested 6

6 Approach 6
6.1 Test Levels and Techniques 6
6.2 Test Data Strategy 7
6.3 Coverage Method 7

7 Item Pass/Fail Criteria 7
7.1 Project Success Criteria 8
7.2 Test Item Criteria 8
7.3 Basis for Criteria 9

8 Suspension Criteria and Resumption Requirements 9
8.1 Suspension Criteria 9
8.2 Resumption Requirements 9

9 Test Deliverables 10

10 Testing Tasks 10

11 Environmental Needs 11
11.1 Hardware 11
11.2 Software 11
11.3 Network and Configuration 11

12 Responsibilities 11

13 Staffing and Training Needs 12

14 Schedule 12

15 Risks and Contingencies 12

16 Approvals 13

17 Appendix A: Test Cases (Functional) 14
17.1 [TC-F-001] 다중 심볼 가격 수집 및 정규화 14
17.2 [TC-F-002] 가격 이벤트 DB 저장 및 중복 Upsert 14
17.3 [TC-F-003] Robust Z-score 이상 탐지 14
17.4 [TC-F-004] EMA 기반 이동 평균 이탈 탐지 14
17.5 [TC-F-005] 동시 탐지 시 단일 이벤트 선택 15
17.6 [TC-F-006] 이상 이벤트 저장 및 클러스터링 15
17.7 [TC-F-007] 이벤트 기반 뉴스 수집 범위 및 후보 제한 15
17.8 [TC-F-008] 뉴스 진행 상태 선발행 및 LLM 요약 성공 15

17.9 [TC-F-009] LLM 요약 실패 시 Fallback 처리 16
17.10 [TC-F-010] 뉴스 중복 식별 16
17.11 [TC-F-011] NLI 입력 생성 및 심볼 명칭 정규화 16
17.12 [TC-F-012] NLI 관련성 필터링 및 상태 발행 16
17.13 [TC-F-013] ABSA 입력 생성 및 감성 분석 17
17.14 [TC-F-014] ABSA 완료 및 최종 상태 발행 17
17.15 [TC-F-015] 영향도 점수 계산 17
17.16 [TC-F-016] 뉴스 목록 상태 및 영향도 정렬 17
17.17 [TC-F-017] 이벤트 요약 생성 18
17.18 [TC-F-018] 이벤트 요약 표시 및 생성 시점 18
17.19 [TC-F-019] 최종 상태 저장 규칙 18
17.20 [TC-F-020] Realtime Gateway REST 및 Socket.IO 제공 18
17.21 [TC-F-021] 인터페이스 및 런타임 구성 검증 19

18 Appendix B: Test Scenarios (Quality / Non-functional) 20

18.1 [TS-P-001] 활성 심볼 및 가격 이벤트 처리량 20
18.2 [TS-P-002] 실시간 브로드캐스트 지연 20
18.3 [TS-P-003] REST 조회 및 1년치 시계열 조회 성능 20
18.4 [TS-P-004] 뉴스 수집 및 분석 지연 20
18.5 [TS-P-005] 외부 Processor 통신 지연 21
18.6 [TS-U-001] UI 상호작용 및 사용성 검증 21
18.7 [TS-R-001] 중복 메시지 및 재시작 신뢰성 21
18.8 [TS-R-002] 외부 API 장애 및 fallback 가용성 21
18.9 [TS-R-003] 서비스 장애 후 30초 내 복구 22
18.10 [TS-M-001] 유지보수성 및 어댑터 구조 검증 22
18.11 [TS-S-001] 보안 설정 검증 22
18.12 [TS-PT-001] Docker 이식성 검증 22
18.13 [TS-D-001] 설계 제약 및 데이터 계약 검증 23
18.14 [TS-Q-001] ABSA 모델 품질 벤치마크 23

1 Test Plan Identifier

본 문서의 식별 번호는 `EP-STP-1.0`이다. 본 문서는 Epoch Pulse. Version 1.0의 시스템 테스트 계획을 정의하며, IEEE Std 829-1998의 Test Plan 문서 구조를 따른다.

본 문서에서 사용하는 기준 요구사항 명세서는 `Epoch Pulse. Software Requirements Specification, Version 1.0`이다.

2 Introduction

2.1 Objectives

본 테스트 계획의 목적은 Epoch Pulse.의 기능 요구사항과 품질 요구사항이 SRS에 명시된 수락 기준을 만족하는지 검증하기 위한 범위, 접근 방법, 자원, 일정, 책임, 산출물, 중단 및 재개 기준을 정의하는 것이다.

Epoch Pulse.는 Kafka 기반 Choreography 마이크로서비스 아키텍처를 사용하여 실시간 주가 이상 탐지, 관련 뉴스 수집 및 요약, NLI 관련성 검증, ABSA 감성 분석, 이벤트 요약, 웹 대시보드 표시를 수행한다. 따라서 본 테스트는 개별 모듈의 계산 정확성뿐 아니라 서비스 간 메시지 정합성, 최종 상태 저장 규칙, 실시간 브로드캐스트 지연, 모델 품질, 장애 복구를 함께 검증한다.

2.2 Scope

본 계획은 단위 테스트, 통합 테스트, 시스템 테스트, 성능 테스트, 장애 복구 테스트, 모델 품질 검증, 설정 및 보안 검토, 사용성 검증을 포함한다. 요구사항 커버리지는 Appendix A의 기능 테스트 케이스와 Appendix B의 품질 테스트 시나리오에서 `연결 요구사항` 필드로 표시한다.

기능 및 품질 요구사항 전체 커버리지 증명은 각 테스트 케이스와 시나리오의 요구사항 연결 항목으로 확인한다.

2.3 References

1. IEEE Std 829-1998, IEEE Standard for Software Test Documentation.
2. Epoch Pulse. Software Requirements Specification, Version 1.0.
3. Iglewicz, B. and Hoaglin, D.C. (1993), “How to Detect and Handle Outliers”, ASQC Quality Press.
4. Apache Kafka Documentation.
5. PostgreSQL and TimescaleDB Documentation.
6. Socket.IO Documentation.

3 Test Items

본 테스트의 대상 항목은 Epoch Pulse.를 구성하는 소프트웨어 항목과 주요 데이터 계약이다. 각 소프트웨어 항목의 버전은 Epoch Pulse. Version 1.0 기준으로 하며, 컨테이너 또는 독립 실행 가능한 서비스 단위로 시험한다.

Test Item	**Description**
Anomaly Sentinel	활성 심볼 가격 이벤트 수집, 정규화, Robust Z-score 및 EMA 기반 이상 탐지, `stock-prices` 및 `anomaly-event` 발행.
News Crawler Summarizer	이상 이벤트 기준 전후 20분 뉴스 수집, 최대 20건 후보 제한, LLM 요약, fallback content 처리, `raw-news` 및 `nli-in` 발행.
NLI Processor	DeBERTa 기반 자연어 추론 모델 로드, 심볼 관련성 검증, `NLI_PASSED` 또는 `NLI_DROPPED` 상태 발행.
ABSA Processor	DeBERTa 기반 ABSA 모델 로드, 심볼 감성 분석, `DONE` 최종 상태 발행.
Storage Writer	가격 데이터, 이상 이벤트, 최종 뉴스 분석 결과를 TimescaleDB에 upsert하고 중간 상태 저장을 배제.
Realtime Gateway	REST API 제공, Kafka 이벤트를 Socket.IO 이벤트로 변환, `price:tick`, `anomaly:detected`, `news:progress`, `news:final`, `event:summary`, `symbol:clear` 제공.
Front-end	React 기반 단일 웹 대시보드, 가격 차트, 이상치 클러스터, 뉴스 진행 상태, 영향도 점수, 이벤트 요약 표시.
Database Schema	`stock_prices`, `anomalies`, `news_analysis` 테이블 키, JSONB payload, 인덱스, 보존 정책.
Kafka Data Contract	`stock-prices`, `anomaly-event`, `raw-news`, `nli-in`, `nli-out-absa-in`, `absa-out` 토픽 및 JSON 메시지 형식.

4 Features to Be Tested

본 테스트는 SRS의 다음 요구사항 그룹과 조합 기능을 시험한다.

- **External Interface**: UI-1..UI-10, HW-1..HW-4, SI-1..SI-11, CI-1..CI-4.
- **Functional**: FR-1.1..FR-1.5 가격 수집 및 정규화, FR-2.1..FR-2.8 이상 탐지 및 클러스터링, FR-3.1..FR-3.10 뉴스 수집 및 요약, FR-4.1..FR-4.10 NLI 관련성 필터링, FR-5.1..FR-5.11 ABSA 감성 분석, FR-6.1..FR-6.3 영향도 점수 및 뉴스 정렬, FR-7.1..FR-7.5 이벤트 요약, FR-8.1..FR-8.4 저장 및 서빙.
- **Performance**: PR-1..PR-12의 활성 심볼 수, 가격 이벤트 처리량, p95 지연 시간, REST 조회, 이벤트 요약 지연, 외부 Processor RTT, 모델 추론 시간.
- **Database**: DB-1..DB-6의 키, JSONB 저장, 최종 상태 제약, 인덱스, 1년 보존.
- **Design Constraints**: DC-1..DC-20의 Kafka Choreography flow, Docker 배포, 어댑터 경계, 모델 로드 전략, 웹 대시보드 페이지 구조, 단일 사용자 운영, 토픽 입출력, Offset 커밋, UTC 타임스탬프, 토픽 파티션 상한.

- **Software System Attributes**: QA-1..QA-20의 신뢰성, 가용성, 유지보수성, 사용성, 보안, 이식성, ABSA 모델 품질.

5 Features Not to Be Tested

다음 항목은 본 테스트 계획에서 제외한다. 제외 항목은 Epoch Pulse.의 제어 범위 밖이거나 SRS의 제외 범위에 속한다.

Excluded Feature	Reason
외부 시장 데이터 API, 뉴스 API, LLM API의 내부 구현	외부 공급자의 책임 범위이며 Epoch Pulse.는 어댑터, 호출, 오류 처리, fallback 동작만 검증한다.
Kafka Broker 및 TimescaleDB 내부 엔진 성능	상용/오픈소스 인프라 자체의 내부 알고리즘은 시험 대상이 아니며, 시스템 통합 관점의 메시지 흐름과 저장 결과만 검증한다.
주식 주문, 매매 체결, 포트폴리오 리밸런싱, 투자 자문 자동 실행	SRS 제외 범위에 해당한다.
다중 사용자 계정, 로그인, 권한 관리, 개인화 설정	SRS는 단일 사용자 내부 운영을 기준으로 하며 인증/인가 레이어를 포함하지 않는다.
하드웨어 및 운영체제 자체 보안 취약점 진단	Epoch Pulse. 소프트웨어 요구사항 범위를 벗어난 플랫폼 보안 평가이다.
법적 투자 자문 적합성 판단	시스템 출력은 투자 참고 정보이며 법적 자문 또는 매매 권고가 아니다.

6 Approach

6.1 Test Levels and Techniques

본 테스트는 다음 수준과 기법을 사용한다.

- **Unit Testing**: 이상 탐지 공식, 식별자 생성, 시간 범위 필터, Impact Score, 이벤트 요약 공식, 상태 정규화와 같은 순수 로직을 Pytest 또는 동등한 단위 테스트로 검증한다.
- **Integration Testing**: Kafka 토픽 간 JSON 메시지 스키마, NLI/ABSA 파이프라인 상태 전이, Storage Writer upsert, Gateway REST 및 Socket.IO 변환을 검증한다.
- **System Testing**: Mock 시장 데이터와 Mock 뉴스/LLM 응답을 사용하여 `가격 이벤트 -> 이상 이벤트 -> 뉴스 수집 -> 요약 -> NLI -> ABSA -> 저장 -> UI 표시`의 end-to-end 흐름을 검증한다.
- **Performance Testing**: 가격 이벤트 처리량, REST 조회, Socket.IO 브로드캐스트, 이벤트 요약 지연, Processor RTT, 모델 추론 시간을 p95 기준으로 측정한다.
- **Recovery Testing**: 외부 API timeout, LLM 실패, 컨테이너 재시작, Kafka/DB 일시 장애 이후 중복 저장 및 데이터 유실 여부를 검증한다.

- **Model Benchmark Testing**: 금융 뉴스 도메인 테스트셋을 사용해 ABSA 모델의 Accuracy, Macro F1, Weighted F1, Base Model 대비 향상률을 검증한다.
- **Configuration Review**: Docker Compose, Kafka topic 설정, 환경 변수, 모델 파일 경로, 타임스탬프 포맷, 포트와 파티션 상한을 정적 검토한다.
- **Usability Inspection**: 웹 대시보드에서 사용자가 메인 페이지의 전체 요약과 최근 이상 이벤트, 심볼별 상세 페이지의 차트, 이상치, 뉴스 진행 상태, NLI/ABSA 결과, 이벤트 요약을 구분해서 확인할 수 있는지 검토한다.

6.2 Test Data Strategy

테스트 데이터는 입력 데이터, 기대 결과, 외부 의존성 Mock 조건을 분리하여 관리한다. LLM/API 실패와 같은 장애 조건은 뉴스 데이터 자체에 포함하지 않고 Mock 응답 프로파일로 정의한다.

- **Market Price Baseline Data**: 심볼별 1분 주기의 정상 가격 틱 데이터. 이상 탐지가 발생하지 않아야 하는 음성 대조군과 EMA/Z-score 계산을 위한 warm-up 데이터로 사용한다.
- **Market Price Anomaly Data**: Robust Z-score 3.0 초과, EMA 20 대비 3.0% 초과, 그리고 두 알고리즘이 동시에 이상을 탐지하는 가격 틱 데이터. 동시 탐지 데이터는 FR-2.3의 단일 이벤트 선택 로직을 검증하기 위해 필요하다.
- **Idempotency and Ordering Data**: 동일 `symbol`과 동일 `timestamp`에 대해 payload가 다른 가격 이벤트 및 재전송 이벤트. FR-1.5의 upsert, UI-3의 중복 병합, QA-1의 재시작 후 중복 방지를 검증하기 위해 사용한다.
- **News Candidate Data**: `title`, `content`, `source`, `url`, `published timestamp`를 포함한 금융 뉴스 레코드. 이벤트 시각 전후 20분 내부/외부 기사, 관련/무관 기사, 동일 URL 및 동일 세션 중복 기사를 포함한다.
- **ABSA Evaluation Data**: 대상 심볼 또는 aspect가 명확한 금융 뉴스 데이터. ABSA 품질 검증을 위해 기대 label이 부여되어 있어야 한다.
- **Load and Retention Data**: 120개 활성 심볼 부하 데이터와 심볼당 1년치 약 98,280건의 시계열 데이터. PR-1, PR-2, PR-10의 처리량과 조회 성능을 검증하는 데 사용한다.
- **External Dependency Mock Profiles**: 뉴스 API 및 LLM API의 정상 응답, 지연 응답, timeout, 500 error 응답 프로파일. FR-3.6, QA-3, PR-9의 fallback과 가용성 검증에 사용한다.

6.3 Coverage Method

Appendix A는 기능 요구사항과 외부 인터페이스, 데이터 저장, 실시간 서빙 요구사항을 검증한다.
Appendix B는 성능, 신뢰성, 가용성, 유지보수성, 사용성, 보안, 이식성, 모델 품질, 설계 제약을 검증한다.

7 Item Pass/Fail Criteria

본 섹션은 IEEE Std 829-1998의 Item Pass/Fail Criteria에 따라 각 테스트 항목의 통과/실패 판단 기준을 정의한다. 프로젝트 성공 기준은 개별 항목 기준을 종합한 최종 수락 기준으로 사용한다.

7.1 Project Success Criteria

Quality Item	Pass Criteria	Fail Criteria
기능 완전성	Appendix A의 기능 테스트 케이스 실행률 100%, 치명적 결함 0건	치명적 결함 1건 이상 미해결, 또는 핵심 end-to-end 파이프라인이 완료되지 않음.
성능	Base Node 기준 20개 심볼과 확장 구성 120개 심볼 기준을 만족하고, PR-3..PR-12의 p95 지연 및 처리량 기준을 모두 만족.	SRS의 p95 지연 또는 처리량 기준 중 하나 이상을 초과.
데이터 무결성	중복 가격/이상/뉴스 이벤트가 upsert로 정규화되고, 최종 뉴스 상태만 DB에 저장되며, 데이터 유실 0건.	동일 key 중복 저장, 최종 상태 누락, 중간 상태 DB 저장, 또는 데이터 유실 발생.
신뢰성 및 가용성	외부 LLM API 장애 시 fallback으로 파이프라인 지속, 서비스 장애 후 30초 이내 가용 상태 복구.	외부 LLM API 장애가 가격 수집/이상 탐지/기존 조회를 차단하거나, 복구 기준을 초과.
AI 모델 품질	ABSA 모델이 금융 뉴스 테스트셋 기준 Accuracy 90% 이상, Macro F1 0.85 이상, Weighted F1 0.90 이상, Base 대비 Accuracy +5%p 이상.	모델 품질 지표 중 하나 이상 미달.
사용성	웹 대시보드의 메인 페이지에서 전체 요약과 최근 이상 이벤트를, 심볼별 상세 페이지에서 차트, 대표 이상치, 뉴스 진행 상태, NLI/ABSA 결과, 이벤트 요약을 명확히 확인 가능.	사용자가 분석 상태 또는 결과 의미를 구분하기 어려운 UI 결함 존재.

7.2 Test Item Criteria

Test Item	Pass/Fail Criteria
Anomaly Sentinel	가격 이벤트 정규화, `stock-prices` 발행, DB 저장, Robust Z-score/EMA 임계값 판정, 단일 이상 이벤트 선택, 클러스터링과 대표 이상치 선정이 SRS와 일치하면 Pass.
News Crawler Summarizer	이벤트 시각 전후 20분 필터, 최대 20건 제한, `COLLECTED` 선발행, LLM 성공/실패 처리, 성능 기준 내 후보 뉴스 처리, 메타데이터 기록이 모두 동작하면 Pass.
NLI Processor	NLI 입력 문장쌍, 심볼 명칭 정규화, entailment 0.5 기준 분기, `NLI_PASSED`/`NLI_DROPPED` 발행, duration/timestamp 기록이 일치하면 Pass.
ABSA Processor	ABSA 입력 형식, 심볼 명칭 정규화, Negative/Neutral/Positive 라벨, `DONE` 발행, 결과 필드와 메타데이터 기록이 일치하면 Pass.
Storage Writer	가격/이상 이벤트 저장, `DONE`/`DROPPED` 최종 상태 저장, 중간 상태 미저장, upsert 중복 방지가 검증되면 Pass.

Realtime Gateway	REST 조회와 Socket.IO 이벤트 변환이 SRS의 엔드포인트, 이벤트명, JSON 포맷, 지연 기준을 만족하면 Pass.
Front-end	메인 페이지 요약 및 최근 이상 이벤트 표시, 심볼 검색과 상세 페이지 이동, 상세 페이지 차트 deduplication, 이상치 마커, 뉴스 상태 구분, 영향도 정렬, 이벤트 요약 표시가 요구사항과 일치하면 Pass.
Database Schema	기본 키, 외래 키, JSONB 필드, 인덱스, 상태 제약, 1년 보존 기준이 충족되면 Pass.
Kafka Data Contract	토픽명, 확장 설정, 메시지 필드, ISO8601 UTC 타임스탬프, 상태 전이가 SRS와 일치하면 Pass.

7.3 Basis for Criteria

- **Robust Z-score**: SRS의 임계값 3.0은 실시간 탐지 민감도를 확보하기 위한 기준이며, 최소 10개 누적 데이터 이후에만 적용한다.
- **EMA**: 기간 20의 EMA를 기준으로 3.0% 초과 이탈을 이상으로 판정한다.
- **NLI**: 0.5는 entailment confidence의 수락 기준이며, 이 값 이상이면 관련 뉴스로 통과시킨다.
- **ABSA**: 금융 뉴스 도메인 품질 보장을 위해 Accuracy, Macro F1, Weighted F1 및 Base 대비 향상률을 함께 사용한다.
- **Visualization**: 동일 분 단위 가격 데이터는 최신값으로 병합되어 차트 라이브러리 오류와 왜곡을 방지해야 한다.

8 Suspension Criteria and Resumption Requirements

8.1 Suspension Criteria

다음 조건 중 하나가 발생하면 해당 범위의 테스트를 중단한다.

- Kafka Broker가 동작하지 않아 핵심 토픽 발행/소비를 검증할 수 없는 경우.
- TimescaleDB 접속 불가 또는 스키마 불일치로 저장 검증을 수행할 수 없는 경우.
- NLI 또는 ABSA 모델 파일 및 토크나이저 로드 실패로 검증이 불가능한 경우.
- 테스트 데이터 기준선이 손상되어 예상 결과를 신뢰할 수 없는 경우.
- 치명적 결함으로 인해 이후 테스트 결과가 무의미해지는 경우.

8.2 Resumption Requirements

테스트를 재개하려면 다음 조건을 만족해야 한다.

- 중단 원인과 수정 내역이 Defect Log 또는 테스트 로그에 기록되어야 한다.
- 수정된 항목에 대해 단위 또는 설정 검증이 먼저 통과해야 한다.

- 중단 전에 실패한 테스트와 그 영향을 받는 회귀 테스트를 재실행해야 한다.
- 데이터 손상 가능성이 있는 경우 테스트 DB와 Kafka topic 상태를 초기화하고 기준 데이터를 재적재해야 한다.

9 Test Deliverables

본 테스트 활동의 산출물은 다음과 같다.

1. Software Test Plan (STP): 본 문서.
2. Test Design Specification: 본 문서의 Approach 및 Appendix A/B에 포함된 설계 수준 테스트 식별 정보.
3. Test Case Specification: Appendix A의 기능 테스트 케이스.
4. Test Scenario Specification: Appendix B의 품질/비기능 테스트 시나리오.
5. Test Procedure Specification: 테스트 실행 단계가 별도 자동화 스크립트 또는 실행 체크리스트로 분리될 경우 산출한다.
6. Test Log: 테스트 실행 일시, 환경, 입력 데이터, 결과, 수행자를 기록한 로그.
7. Test Incident Report 또는 Defect Log: 결함 ID, 심각도, 재현 절차, 영향 요구사항, 수정 상태를 기록한 이슈 목록.
8. Test Summary Report (TSR): 최종 실행 결과, 미해결 결함, 잔여 리스크, 승인 판단을 요약한 보고서.
9. Model Benchmark Report: ABSA 모델 품질 지표와 테스트셋 정보를 포함한 보고서.
10. Performance Test Report: p95 지연 시간, 처리량, 오류율, 시스템 자원 사용량을 포함한 보고서.

10 Testing Tasks

ID	Task
T-01	테스트 기준 SRS와 STP 버전을 확정하고 테스트 데이터 세트를 준비한다.
T-02	컨테이너 기반 테스트 환경을 기동하고 Kafka, DB, Gateway, Front-end 상태를 확인한다.
T-03	Mock 시장 데이터로 baseline, anomaly, idempotency 가격 이벤트를 주입한다.
T-04	뉴스 후보 데이터와 외부 의존성 Mock profile을 사용하여 수집, 요약, fallback, 후보 뉴스 처리 흐름을 검증한다.
T-05	NLI/ABSA 모델 로드, 추론, 상태 전이, 메타데이터 기록을 검증한다.
T-06	Storage Writer의 DB upsert, 최종 상태 저장, 스키마 제약을 검증한다.
T-07	Gateway REST API와 Socket.IO 브로드캐스트를 검증한다.
T-08	Front-end 대시보드 표시, 정렬, 상태 구분, 이벤트 요약 UI를 검증한다.
T-09	성능, 장애 복구, 보안 설정, 이식성 검증을 수행한다.
T-10	결함을 분류하고 수정 후 회귀 테스트와 Test Summary Report를 작성한다.

11 Environmental Needs

11.1 Hardware

- Base Node: Intel Processor N100급 4코어/4스레드 CPU, 16GB RAM, 100GB 이상의 SSD 저장소.
- Optional External Node: NLI/ABSA Processor 외부 배치를 위한 x86-64 서버 또는 PC.

11.2 Software

- OS: Docker 27.x와 Docker Compose 2.x를 실행할 수 있는 x86-64 운영체제.
- Runtime: Python 3.10+, Node.js 18+, React 18.3.1.
- Middleware: Apache Kafka 3.9.2 (KRaft Mode).
- Database: PostgreSQL 17 기반 TimescaleDB.
- Test Tools: Pytest, Kafka CLI 또는 동등 도구, Docker CLI, Chrome DevTools, Scikit-learn metric 도구, Mock API Server.

11.3 Network and Configuration

- 내부 Kafka 통신: TCP 9092.
- 외부 Processor 연결: TCP 20000(LAN), TCP 20001(Tailscale).
- Realtime Gateway: TCP 20004.
- 각 마이크로서비스는 환경 변수 또는 설정 파일로 필요한 Kafka, DB, 외부 API 키 등을 주입받아야 한다.

12 Responsibilities

Role	Name	Responsibility
팀장	김태수	테스트 범위 승인, 일정 조정, 최종 릴리즈 판단, 테스트 실행 관리 및 총괄. Anomaly Sentinel, Storage Writer, Realtime Gateway, Front-end, DevOps 관련 결함 수정 및 테스트 지원.
팀원	박종윤	ABSA 모델 파인튜닝, NLI/ABSA 모델 추론 결과 검증, 모델 벤치마크 테스트 및 결함 수정 지원.
팀원	레이셸른	뉴스 수집 및 LLM 요약 파이프라인(News Crawler Summarizer) 관련 테스트 및 결함 수정 지원.

13 Staffing and Training Needs

테스트 수행 인원은 Kafka 이벤트 기반 시스템, Docker Compose 운영, PostgreSQL/TimescaleDB, React 대시보드, Python 기반 ML 추론, 기본 성능 측정 도구를 이해해야 한다.

필요 교육은 다음과 같다.

- IEEE Std 829-1998 기반 테스트 문서 구조 이해.
- SRS 요구사항 ID 및 Epoch Pulse. 데이터 흐름 이해.
- Kafka topic, partition, consumer group, offset commit 기본 교육.
- NLI/ABSA 모델 입력/출력과 성능 지표 교육.
- 테스트 로그와 Defect Log 작성 규칙 교육.

14 Schedule

Milestone	Activity	Exit Criteria
M-01	테스트 계획 및 기준선 확정	STP 승인, 테스트 데이터 준비 완료
M-02	단위 및 컴포넌트 테스트	계산 로직, 상태 전이, 스키마 검증 통과
M-03	통합 테스트	Kafka 파이프라인, DB 저장, Gateway 연동 통과
M-04	시스템 테스트	end-to-end 사용자 시나리오 통과
M-05	성능 및 장애 복구 테스트	PR/QA 기준 통과
M-06	모델 벤치마크	QA-17..QA-20 기준 통과
M-07	회귀 테스트 및 승인	Critical 결함 0건, TSR 승인

15 Risks and Contingencies

Risk	Contingency
외부 뉴스 API 또는 LLM API 지연으로 PR-6/PR-7 측정이 불안정함.	Mock API Server와 기록된 응답 세트를 사용해 시스템 내부 지연과 외부 의존 지연을 분리 측정한다.
NLI/ABSA CPU-only 추론 시간이 환경에 따라 변동됨.	Base Node 기준 측정과 외부 확장 기준 측정을 분리하고, 모델 warm-up 후 p95를 산출한다.

Kafka 또는 DB 장애가 테스트 데이터 오염을 유발함.	장애 테스트 후 topic과 DB를 초기화하고 기준 데이터를 재적재한다.
외부 의존성을 Mock으로 대체한 시험 결과가 실제 운영 응답 특성을 완전히 대표하지 못할 수 있음.	Mock 기반 내부 지연과 실제 외부 의존 지연을 분리하여 보고하고, 운영 전 실제 공급자 연동 확인을 별도 수행한다.
테스트셋 편향으로 ABSA 모델 품질 지표가 실제 사용 품질을 대표하지 못할 수 있음.	금융 뉴스 도메인의 Positive/Neutral/Negative 분포를 점검하고, Macro F1과 Weighted F1을 함께 보고한다.

16 Approvals

Role	Name	Date
팀장	김태수	2026-05-12
팀원	박종윤	2026-05-12
팀원	레이셸른	2026-05-12

17 Appendix A: Test Cases (Functional)

본 부록은 기능 요구사항, 외부 인터페이스, 데이터 저장 및 서빙 요구사항을 검증하는 테스트 케이스 명세이다. 각 항목은 IEEE Std 829-1998의 Test Case Specification 요소 중 식별자, 테스트 항목, 입력, 예상 출력, 환경 및 특수 절차를 포함하도록 작성한다.

17.1 [TC-F-001] 다중 심볼 가격 수집 및 정규화

- 연결 요구사항: FR-1.1, FR-1.2, FR-1.3, SI-1, CI-4
- 테스트 유형: 기능 / 통합
- 전제 조건: 20개 활성 심볼 설정, Anomaly Sentinel 및 Kafka 실행
- 입력값: 각 심볼에 대해 1분 주기의 가격 이벤트(`symbol`, `event_time`, `price`, `previous_price`, `change_percent`, `volume`)
- 예상 결과: 모든 이벤트가 JSON UTF-8 형식으로 정규화되어 `stock-prices` 토픽에 발행된다.
- **Pass/Fail**: 누락 이벤트 0건, 필수 필드 누락 0건, ISO8601 UTC 타임스탬프 준수 시 Pass.

17.2 [TC-F-002] 가격 이벤트 DB 저장 및 중복 Upsert

- 연결 요구사항: FR-1.4, FR-1.5, DB-1, DB-5, DC-16
- 테스트 유형: 기능 / 데이터베이스
- 전제 조건: Storage Writer와 TimescaleDB 실행
- 입력값: 동일 `symbol`, 동일 `timestamp`를 가진 가격 이벤트 2건
- 예상 결과: `stock_prices`에는 1건만 저장되고 payload는 최신 이벤트 값으로 갱신된다.
- **Pass/Fail**: `(symbol, timestamp)` 기준 중복 행이 없고 최신 payload가 저장되면 Pass.

17.3 [TC-F-003] Robust Z-score 이상 탐지

- 연결 요구사항: FR-2.1, FR-2.4, FR-2.5
- 테스트 유형: 기능 / 단위 + 통합
- 전제 조건: 대상 심볼에 최소 10개 누적 가격 데이터 존재
- 입력값: 변동률 Robust Z-score가 3.0을 초과하는 가격 이벤트
- 예상 결과: UUID, 심볼, 이벤트 시각, 탐지 시각, 가격, 변동률, 알고리즘, 점수, 메시지를 포함한 `anomaly-event`가 발행된다.
- **Pass/Fail**: 임계값 초과 시 이벤트 1건이 발행되고 임계값 이하 데이터에서 발행되지 않으면 Pass.

17.4 [TC-F-004] EMA 기반 이동 평균 이탈 탐지

- 연결 요구사항: FR-2.2, FR-2.4, FR-2.5
- 테스트 유형: 기능 / 단위 + 통합

- 전제 조건: EMA(기간 20)를 계산할 수 있는 가격 이력 존재
- 입력값: EMA 대비 3.0%를 초과해 이탈하는 가격 이벤트
- 예상 결과: EMA 기반 이상 이벤트가 `anomaly-event`로 발행된다.
- **Pass/Fail**: EMA 이탈률이 기준 초과일 때만 이벤트가 발행되면 Pass.

17.5 [TC-F-005] 동시 탐지 시 단일 이벤트 선택

- 연결 요구사항: FR-2.3, FR-2.4, FR-2.5
- 테스트 유형: 기능 / 단위
- 전제 조건: Z-score와 EMA가 모두 이상으로 판정되는 입력 준비
- 입력값: Z-score와 EMA의 이상 점수가 서로 다른 가격 이벤트
- 예상 결과: 절대값이 더 높은 score를 가진 알고리즘 결과가 최종 채택되고 단일 `anomaly-event`만 발행된다.
- **Pass/Fail**: 중복 이상 이벤트 없이 더 높은 score의 결과 1건만 남으면 Pass.

17.6 [TC-F-006] 이상 이벤트 저장 및 클러스터링

- 연결 요구사항: FR-2.6, FR-2.7, FR-2.8, UI-4, UI-5, DB-2
- 테스트 유형: 기능 / 통합 + UI
- 전제 조건: 같은 심볼의 이상 이벤트 여러 건 준비
- 입력값: 발생 간격 15분 이내이고 방향이 같은 이상 이벤트와, 방향 또는 시간이 다른 이벤트
- 예상 결과: 동일 방향 15분 이내 이벤트는 같은 클러스터로 묶이고 score 절대값 최대 이벤트가 대표 이상치로 표시된다.
- **Pass/Fail**: 클러스터 수, 대표 이상치, 차트 마커 위계가 기대값과 일치하면 Pass.

17.7 [TC-F-007] 이벤트 기반 뉴스 수집 범위 및 후보 제한

- 연결 요구사항: FR-3.1, FR-3.2, FR-3.3, SI-2
- 테스트 유형: 기능 / 통합
- 전제 조건: `anomaly-event` 발생, 뉴스 데이터셋 또는 Mock News API 준비
- 입력값: 이벤트 시각 전후 20분 안팎의 뉴스 30건
- 예상 결과: 전후 20분 범위 내 뉴스만 유지되고 분석 후보는 최대 20건으로 제한된다.
- **Pass/Fail**: 범위 밖 뉴스 0건 전달, 후보 수 20건 이하이면 Pass.

17.8 [TC-F-008] 뉴스 진행 상태 선발행 및 LLM 요약 성공

- 연결 요구사항: FR-3.4, FR-3.5, FR-3.7, FR-3.9, FR-3.10, SI-3
- 테스트 유형: 기능 / 통합
- 전제 조건: LLM Mock API 정상 응답
- 입력값: 후보 뉴스 20건

- 예상 결과: 수집 직후 `COLLECTED` 상태가 `raw-news`로 발행되고, 요약 완료 후 `SUMMARIZED` 상태가 `nli-in`으로 발행되며 duration과 summarized_at이 기록된다.
- **Pass/Fail**: 최대 20건 후보 뉴스 처리, 상태 발행 순서, 메타데이터 기록이 모두 일치하면 Pass.

17.9 [TC-F-009] LLM 요약 실패 시 Fallback 처리

- 연결 요구사항: FR-3.6, QA-3, PR-9
- 테스트 유형: 기능 / 장애 복구
- 전제 조건: LLM Mock API timeout 또는 500 error 설정
- 입력값: 후보 뉴스 원문 또는 snippet
- 예상 결과: 요약 실패 후에도 원문 또는 snippet이 `nli-in`으로 발행되어 NLI/ABSA 파이프라인이 계속 진행된다.
- **Pass/Fail**: 파이프라인 중단 없이 fallback content가 후속 분석으로 전달되면 Pass.

17.10 [TC-F-010] 뉴스 중복 식별

- 연결 요구사항: FR-3.8
- 테스트 유형: 기능 / 단위
- 전제 조건: 동일 기사와 서로 다른 기사 데이터를 준비
- 입력값: 동일 기사 2건, 다른 기사 1건
- 예상 결과: 동일 기사(url기준)는 같은 뉴스로 식별되고 다른 기사는 별도 뉴스로 식별된다.
- **Pass/Fail**: 동일 세션 내 중복 뉴스가 중복 분석되지 않으면 Pass.

17.11 [TC-F-011] NLI 입력 생성 및 심볼 명칭 정규화

- 연결 요구사항: FR-4.1, FR-4.2, FR-4.3, FR-4.4, DC-17
- 테스트 유형: 기능 / 단위
- 전제 조건: 표시명과 분석용 명칭이 정의된 심볼 데이터 준비
- 입력값: 뉴스 content와 대상 심볼
- 예상 결과: NLI 입력이 `content`와 `This news is about {symbol}.` 형식의 문장쌍으로 구성되고 대상 심볼 명칭이 설정된 기준에 따라 정규화된다.
- **Pass/Fail**: 입력 문장쌍과 정규화 결과가 기대값과 일치하면 Pass.

17.12 [TC-F-012] NLI 관련성 필터링 및 상태 발행

- 연결 요구사항: FR-4.5, FR-4.6, FR-4.7, FR-4.8, FR-4.9, FR-4.10, DC-9
- 테스트 유형: 기능 / 모델 파이프라인
- 전제 조건: NLI Processor 실행, `nli-in` 입력 준비
- 입력값: entailment score 0.5 이상 뉴스와 0.5 미만 뉴스

- 예상 결과: 0.5 이상은 `NLI_PASSED`로 `nli-out-absa-in`에 발행되고, 0.5 미만은 `NLI_DROPPED`로 `absa-out`에 발행된다.
- **Pass/Fail**: label, score, all_scores, duration, processed_at이 포함되고 분기가 정확하면 Pass.

17.13 [TC-F-013] ABSA 입력 생성 및 감성 분석

- 연결 요구사항: FR-5.1, FR-5.2, FR-5.3, FR-5.4, FR-5.5, FR-5.6, DC-10
- 테스트 유형: 기능 / 모델 파이프라인
- 전제 조건: ABSA Processor 실행, `NLI_PASSED` 뉴스 준비
- 입력값: 뉴스 content와 symbol
- 예상 결과: `[CLS] content [SEP] symbol [SEP]` 형식 입력으로 Negative, Neutral, Positive 중 하나의 label과 score, all_scores가 산출된다.
- **Pass/Fail**: 입력 형식, 심볼 명칭 정규화, 감성 결과 필드가 SRS와 일치하면 Pass.

17.14 [TC-F-014] ABSA 완료 및 최종 상태 발행

- 연결 요구사항: FR-5.7, FR-5.8, FR-5.9, FR-5.10, FR-5.11
- 테스트 유형: 기능 / 통합
- 전제 조건: ABSA 분석 성공
- 입력값: `NLI_PASSED` 뉴스
- 예상 결과: `absa_result`가 채워지고 `DONE` 상태의 메시지가 `absa-out`으로 발행된다.
- **Pass/Fail**: `DONE` 상태, `absa_result`, `absa_duration`, `absa_processed_at`이 모두 존재하면 Pass.

17.15 [TC-F-015] 영향도 점수 계산

- 연결 요구사항: FR-6.1, FR-6.2
- 테스트 유형: 기능 / 단위
- 전제 조건: NLI score, ABSA score, 감성 label, 가격 방향 데이터 준비
- 입력값: 방향 일치, 방향 상충, Neutral 뉴스 각각의 점수 세트
- 예상 결과: `Impact = min(1.0, (EntailmentScore * 0.4) + (SentimentScore * 0.6 * AlignmentWeight))` 공식과 Alignment Weight(1.0, 0.5, 0.1)가 적용된다.
- **Pass/Fail**: 계산 결과가 허용 오차 내에서 기대값과 일치하면 Pass.

17.16 [TC-F-016] 뉴스 목록 상태 및 영향도 정렬

- 연결 요구사항: FR-6.3, UI-6, UI-7, UI-8, UI-9
- 테스트 유형: 기능 / UI
- 전제 조건: 다양한 상태와 점수를 가진 뉴스 목록과 이상치 클러스터 마커 데이터 준비
- 입력값: `DONE`, `COLLECTED`, `SUMMARIZED`, `NLI_PASSED`, `DROPPED`, `NLI_DROPPED` 상태 뉴스, 마커 클릭

- 예상 결과: 마커 클릭 시 해당 클러스터 시간 범위로 뉴스 목록이 필터링되며, 우선순위에 따라 정렬되고 같은 상태 내에서는 영향도 점수 내림차순으로 정렬된다.
- **Pass/Fail**: 화면 표시 순서와 상태 구분이 SRS와 일치하면 Pass.

17.17 [TC-F-017] 이벤트 요약 생성

- 연결 요구사항: FR-7.1, FR-7.2, FR-7.3
- 테스트 유형: 기능 / 통합
- 전제 조건: 대표 이상치에 연결된 DONE 뉴스 1건 이상 존재
- 입력값: DONE 뉴스들의 ABSA 점수, NLI entailment 점수, 영향도 점수
- 예상 결과: 종합 감성 점수, 감성 label, 감성 일관성, 신호 신뢰도, 핵심 근거 상위 3개 뉴스가 생성된다.
- **Pass/Fail**: SRS Appendix B의 공식과 기대 결과가 일치하면 Pass.

17.18 [TC-F-018] 이벤트 요약 표시 및 생성 시점

- 연결 요구사항: FR-7.4, FR-7.5, UI-10, CI-3
- 테스트 유형: 기능 / UI + 통합
- 전제 조건: 후보 뉴스 완료 조건 또는 180초 timeout 조건 준비
- 입력값: 모든 후보 뉴스 최종 상태 완료 케이스, 일부 미완료 후 180초 경과 케이스
- 예상 결과: 이벤트 요약이 대표 이상치 내역 상단에 표시되고 `event:summary` 이벤트로 제공된다.
- **Pass/Fail**: 완료 또는 timeout 조건에서 정확히 한 번 생성 및 표시되면 Pass.

17.19 [TC-F-019] 최종 상태 저장 규칙

- 연결 요구사항: FR-8.1, FR-8.2, DB-3, DB-4, QA-2
- 테스트 유형: 기능 / 데이터베이스
- 전제 조건: Storage Writer 실행
- 입력값: `COLLECTED`, `SUMMARIZED`, `NLI_PASSED`, `NLI_DROPPED`, `DONE` 상태 메시지
- 예상 결과: DB에는 `DONE` 또는 `DROPPED` 최종 상태만 저장되고 `NLI_DROPPED`는 `DROPPED`로 정규화된다.
- **Pass/Fail**: 중간 상태 DB 행 0건, 최종 상태 저장 정상, 탈락 뉴스 보존 시 Pass.

17.20 [TC-F-020] Realtime Gateway REST 및 Socket.IO 제공

- 연결 요구사항: FR-8.3, FR-8.4, CI-2, CI-3, UI-1, UI-2
- 테스트 유형: 기능 / 통합 + UI
- 전제 조건: DB에 가격, 이상 이벤트, 뉴스 분석 데이터 존재
- 입력값: 심볼 목록, 가격 히스토리, 이상 탐지, 뉴스, 대시보드 요약 REST 호출 및 Kafka 실시간 이벤트
- 예상 결과: REST 응답은 JSON 형식으로 반환되고 Kafka 이벤트는 Socket.IO 이벤트로 브로드캐스트되며, 메인 페이지 요약과 심볼별 상세 페이지 데이터가 각각 표시된다.

- **Pass/Fail**: 모든 지정 REST/API와 이벤트명이 정상 동작하고 메인 페이지 및 심볼별 상세 페이지에 요구 정보가 표시되면 Pass.

17.21 [TC-F-021] 인터페이스 및 런타임 구성 검증

- 연결 요구사항: HW-1..HW-4, SI-4..SI-11, CI-1, CI-4, DC-1, DC-2, DC-7, DC-8, DC-18, DC-20
- 테스트 유형: 설정 검토 / 통합
- 전제 조건: 배포 설정 파일과 실행 환경 준비
- 입력값: Docker Compose, Kafka 설정, 포트 설정, Processor 배치 설정
- 예상 결과: x86-64 PC, Docker, Kafka 3.9.2, PostgreSQL 17/TimescaleDB, Python 3.10+, React 18.3.1, 포트 및 파티션 상한이 SRS와 일치한다.
- **Pass/Fail**: 설정 불일치가 없고 기본 및 외부 Processor 통신 구성이 검증되면 Pass.

18 Appendix B: Test Scenarios (Quality / Non-functional)

본 부록은 SRS의 성능 요구사항, 품질 속성, 설계 제약을 검증하는 시나리오 명세이다.

18.1 [TS-P-001] 활성 심볼 및 가격 이벤트 처리량

- 연결 요구사항: PR-1, PR-2
- 테스트 도구: Kafka Producer 부하 스크립트, 시스템 로그, DB 카운트 쿼리
- 시나리오: Base Node 기준 20개 활성 심볼에 대해 20건/분 가격 이벤트를 30분간 주입하고, 최대 확장 구성 시 120개 활성 심볼에 대해 120건/분을 주입한다.
- 측정 항목: Kafka 발행 수, DB 저장 수, 누락률, 오류율
- **Pass** 기준: 허용 심볼 수와 이벤트 처리량을 만족하고 누락률 0%, 처리 오류 0건.

18.2 [TS-P-002] 실시간 브로드캐스트 지연

- 연결 요구사항: PR-3, PR-4, CI-3
- 테스트 도구: Gateway timestamp log, Browser performance log
- 시나리오: 가격 이벤트와 이상 이벤트를 각각 발생시키고 Kafka 수신 시각부터 대시보드 표시 시각까지 측정한다.
- 측정 항목: `price:tick` p95 지연, `anomaly:detected` p95 지연
- **Pass** 기준: `price:tick` p95 500ms 이하, `anomaly:detected` p95 1초 이하.

18.3 [TS-P-003] REST 조회 및 1년치 시계열 조회 성능

- 연결 요구사항: PR-5, PR-10, DB-5, DB-6
- 테스트 도구: HTTP benchmark tool, SQL timing, TimescaleDB query log
- 시나리오: 심볼당 약 98,280건의 1년치 가격 데이터를 적재한 후 단일 심볼 상세 REST 조회와 특정 1시간 범위 조회를 반복한다.
- 측정 항목: REST p95 응답시간, DB query p95 응답시간
- **Pass** 기준: 단일 심볼 초기 REST 조회 p95 500ms 이하, 1시간 범위 DB 조회 p95 1초 이하.

18.4 [TS-P-004] 뉴스 수집 및 분석 지연

- 연결 요구사항: PR-6, PR-7, PR-12
- 테스트 도구: Trace log, Mock News API, Mock LLM API, Processor log
- 시나리오: 이상 이벤트 1건에 대해 후보 뉴스 20건을 처리하고 `COLLECTED`, NLI/ABSA 완료, 이벤트 요약 생성까지 추적한다.
- 측정 항목: `COLLECTED` 표시 지연, 이벤트 요약 지연, NLI/ABSA 단일 건 추론 시간

- **Pass** 기준: `COLLECTED` p95 3초 이하, Base Node 이벤트 요약 p95 180초 이하, 확장 구성 이벤트 요약 p95 30초 이하, 단일 모델 추론 시간 3초 이하.

18.5 [TS-P-005] 외부 **Processor** 통신 지연

- 연결 요구사항: PR-11, CI-1, HW-4, DC-18
- 테스트 도구: RTT 측정 도구, Kafka consumer lag metric
- 시나리오: Tailscale 또는 LAN으로 외부 NLI/ABSA Processor를 연결하고 `nli-in`, `nli-out-absa-in` 메시지 왕복 지연을 측정한다.
- 측정 항목: 네트워크 RTT p95, Kafka consumer lag
- **Pass** 기준: RTT p95 500ms 이하, 메시지 lag가 지속적으로 증가하지 않음.

18.6 [TS-U-001] UI 상호작용 및 사용성 검증

- 연결 요구사항: PR-8, QA-9, QA-10, QA-11, QA-12, QA-13, UI-1..UI-10
- 테스트 도구: Chrome DevTools, UI inspection checklist
- 시나리오: 사용자가 웹 대시보드의 메인 페이지에서 활성 심볼 수, 분석된 뉴스 수, 최근 발견된 이상 이벤트 최대 10건을 확인하고, 특정 심볼을 검색해 심볼별 상세 페이지로 이동한 뒤 가격 차트, 대표 이상치, 뉴스 진행 상태, NLI 탈락 뉴스, ABSA 완료 뉴스, 감성 라벨, 클러스터 필터링, 이벤트 요약을 확인한다.
- 측정 항목: INP p95, 정보 구분 가능 여부, 표시 누락 여부
- **Pass** 기준: INP p95 200ms 이하, 요구된 모든 정보가 메인 페이지와 심볼별 상세 페이지에서 명확히 구분되어 표시됨.

18.7 [TS-R-001] 중복 메시지 및 재시작 신뢰성

- 연결 요구사항: QA-1, DC-15, FR-1.5, FR-8.2
- 테스트 도구: Docker CLI, Kafka replay, DB count query
- 시나리오: 같은 key의 가격/이상/뉴스 메시지를 재전송하고, 처리 중 컨테이너를 재시작한 뒤 같은 메시지를 재소비하게 한다.
- 측정 항목: 중복 행 수, upsert 결과, offset commit 이후 재처리 여부
- **Pass** 기준: 중복 저장 0건, 최신 값 upsert, 최종 상태 저장 규칙 유지.

18.8 [TS-R-002] 외부 **API** 장애 및 **fallback** 가용성

- 연결 요구사항: QA-3, QA-5, PR-9, FR-3.6
- 테스트 도구: Mock API Server, Trace log
- 시나리오: 뉴스 API 또는 LLM API에 timeout과 500 error를 주입한다.
- 측정 항목: fallback content 사용 여부, 가격 수집/이상 탐지/기존 조회 지속 여부, 메시지 스트림 차단 여부

- **Pass** 기준: fallback으로 NLI/ABSA 파이프라인이 계속 진행되고 핵심 서비스 가용성이 저해되지 않음.

18.9 [TS-R-003] 서비스 장애 후 30초 내 복구

- 연결 요구사항: QA-4
- 테스트 도구: Docker Chaos, health check log
- 시나리오: 각 마이크로서비스 컨테이너를 강제 종료하고 재시작 정책에 따라 복구되는지 관찰한다.
- 측정 항목: 장애 감지 시각, health check 정상화 시각
- **Pass** 기준: 각 서비스가 30초 이내 정상 가용 상태로 복구됨.

18.10 [TS-M-001] 유지보수성 및 어댑터 구조 검증

- 연결 요구사항: QA-6, QA-7, QA-8, DC-3, DC-4, DC-11, DC-12, DC-13, DC-14
- 테스트 도구: Source inspection, configuration review
- 시나리오: Kafka/DB/API 키 설정 분리, 시장 데이터/뉴스/LLM 어댑터 경계, 모델 디렉터리 분리, 서비스 시작 시 1회 모델 로드 여부를 검토한다.
- 측정 항목: 하드코딩 설정 수, 어댑터 교체 영향 범위, 모델 reload 횟수
- **Pass** 기준: 설정은 코드와 분리되고, 공급자 교체가 메소드/클래스 수준에서 가능하며, 모델은 디렉터리에서 최초 실행 시 1회 로드됨.

18.11 [TS-S-001] 보안 설정 검증

- 연결 요구사항: QA-14, QA-15, DC-6
- 테스트 도구: Secret scanner, configuration review
- 시나리오: 소스 코드와 설정 파일에서 외부 API 키 및 DB 비밀번호 하드코딩 여부를 검색하고, 인증/인가 레이어가 단일 사용자 내부 운영 가정과 일치하는지 검토한다.
- 측정 항목: 하드코딩 secret 검출 수, 공개 인터넷 다중 사용자 인증 요구 미포함 여부
- **Pass** 기준: 하드코딩 secret 0건, 단일 사용자 내부 운영 기준과 문서/구현이 일치함.

18.12 [TS-PT-001] Docker 이식성 검증

- 연결 요구사항: QA-16, SI-11, DC-2
- 테스트 도구: Docker CLI, Docker Compose
- 시나리오: 깨끗한 Docker 환경에서 모든 마이크로서비스를 컨테이너 단위로 빌드 및 실행한다.
- 측정 항목: 빌드 성공 여부, 컨테이너 health 상태, OS 종속 오류
- **Pass** 기준: 단일 Docker Compose 환경에서 모든 서비스가 정상 기동하고 OS별 수동 설치 의존성이 없음.

18.13 [TS-D-001] 설계 제약 및 데이터 계약 검증

- 연결 요구사항: DC-1, DC-5, DC-7, DC-8, DC-9, DC-10, DC-15, DC-16, DC-17, DC-19, DC-20, CI-4
- 테스트 도구: Architecture review, Kafka topic inspection, log inspection
- 시나리오: Kafka Choreography flow, 웹 대시보드 페이지 구조, Base Node 단일 인스턴스, 확장 가능한 Processor 구조, 토픽 입출력, Offset commit, UTC 타임스탬프를 검토한다.
- 측정 항목: 토픽/파티션 설정, 서비스 인스턴스 수, timestamp 포맷, producer/consumer 설정
- **Pass** 기준: SRS의 설계 제약 및 데이터 계약과 불일치가 없음.

18.14 [TS-Q-001] ABSA 모델 품질 벤치마크

- 연결 요구사항: QA-17, QA-18, QA-19, QA-20
- 테스트 도구: PyTorch, Scikit-learn metrics
- 시나리오: 금융 뉴스 도메인 테스트셋을 사용하여 파인튜닝 전 Base Model과 최종 ABSA 모델의 예측 결과를 비교한다.
- 측정 항목: Accuracy, Macro F1, Weighted F1, Base 대비 Accuracy 향상률
- **Pass** 기준: Accuracy 90% 이상, Macro F1 0.85 이상, Weighted F1 0.90 이상, Base 대비 Accuracy 향상률 5%p 이상.