


# AI 기반 Observability Platform 백서: 'APM and Observability'에서 'Observability Platforms'로

“옵저버빌리티 패러다임이 2010년대 APM/통합관제 분리 → 2020~22년 3축  
옵저버빌리티 → 2025~ Agentic AIOps로 이동, 평가 좌표가 "데이터 수집  
폭"에서 "Time-to-Resolution"으로 전환됨 ”

이 백서는 Datadog·Dynatrace·Splunk·Elastic·IBM Instana 등 글로벌 11개  
벤더와 국내 6사를 BYO LLM·MCP·Self-hosted 3축으로 정량 비교해, 한국  
시장에서 실제 도입 가능한 5강 후보군과 도입 불가 영역을 명확히 분리해냅니다.



 [hello@cncf.co.kr](mailto:hello@cncf.co.kr)

 02-469-5426

 [www.cncf.co.kr](http://www.cncf.co.kr)

# Contents

- 1.1 모니터링에서 옴저버빌리티로, 옴저버빌리티에서 Agentic AIOps로 — 2010년  
대~2026년 시장 변천사 . . . . . 8
  - 1.1.1 APM·통합관제의 분리 시대(2010년대) → 3축 옴저버빌리티(2020~2022)  
→ Agentic AIOps(2025~) . . . . . 8
  - 1.1.2 4가지 가속 페달 — 텔레메트리 폭증·MTTR 압력·SRE 채용난·GenAI 비용  
급락 . . . . . 10
  - 1.1.3 Gartner 2025의 “Observability Platforms” 카테고리 통합과 MCP 표  
준화의 결정적 변곡점 . . . . . 12
- 1.2 트러블슈팅 즉각 종결 패러다임 — “데이터 수집 폭”에서 “Time-to-Resolution”  
으로 . . . . . 14
  - 1.2.1 레거시 옴저버빌리티의 5대 한계 — 알람 폭증·RCA 가설 인간 의  
존·Top·Bottom 단절·NL 부재·다양한 모니터링 함정 . . . . . 14
  - 1.2.2 “다양한 모니터링”의 함정 — APM/Logs/Metrics/Network/Synthetic  
5개 SKU·4개 콘솔 분리 패턴 . . . . . 16
  - 1.2.3 Time-to-Resolution의 측정 가능 정의 — conclusive investiga-  
tion·MTTR·NL→결론까지의 클릭 수 . . . . . 17
- 1.3 IT 의사결정자에게 주는 1장의 결론 — 평가 좌표를 “Time-to-Resolution +  
Agentic AI 자율화 깊이”로 재설정하라 . . . . . 19
  - 1.3.1 평가 좌표 재설정 가이드 — 데이터 수집 폭(넓음) → Time-to-Resolution  
+ Agentic AI 자율화 깊이(새 좌표) . . . . . 19
  - 1.3.2 2장 예고 — 평가 좌표를 정확히 사용하기 위한 7개 필수 개념과 의사결정  
프레임 . . . . . 20
- 2장: 핵심 개념과 용어 — Agentic AI / Causal AI / MCP / BYO LLM / Top-  
down·Bottom-up . . . . . 21**
  - 2.1 IT 의사결정자가 30분 안에 흡수해야 할 7개 필수 개념 . . . . . 21
    - 2.1.1 Observability·Monitoring·AIOps — 데이터·분석·실행 3계층 구분 . . . 22

- 2.1.2 Agentic AI vs Causal AI — 폐쇄 루프 자율화 vs 인과 추론 . . . . . 23
- 2.1.3 MCP·BYO LLM·Top-down·Bottom-up — 한국 도입의 4가지 결정 변수 25
- 2.2 혼동 쉬운 개념 쌍 — 의사결정 오류를 만드는 3가지 함정 . . . . . 27
  - 2.2.1 AIOps vs Observability — “AIOps 도입했으니 Observability는 별도 안 사도 된다” 의 오해 . . . . . 27
  - 2.2.2 LLM Observability vs Observability with LLM — 모니터링 대상이 LLM인가, 도구가 LLM인가 . . . . . 28
  - 2.2.3 Monitoring as Code vs Configuration as Code — 모니터링 정책의 Git 관리 영역 분리 . . . . . 30
- 2.3 의사결정 프레임 적용 — 7개 개념을 RFP·Vendor Scorecard·법무 검토에 어떻게 배치하는가 . . . . . 32
  - 2.3.1 7개 개념 → RFP 평가 항목 매핑 — 30분 회의용 평가지 . . . . . 32
  - 2.3.2 BYO LLM·MCP·Self-hosted·Causal·Agentic 게이팅 — 5축 의사결정 매트릭스의 도입 . . . . . 33
- 3장: 글로벌 11개 벤더 AI 기능 심층 비교 — Datadog·Dynatrace·New Relic·Splunk·IBM Instana 등 . . . . . 34**
  - 3.1 11개 벤더 AI 기능 매트릭스 — 핵심 기능 / 기반 모델 / GA 시점 / 과금 4축 비교 34
    - 3.1.1 SaaS 5사 — Datadog Bits AI SRE·Dynatrace Davis·New Relic NRAI·Splunk AI Assistant·Cisco AppD . . . . . 34
    - 3.1.2 엔터프라이즈·하이브리드 4사 — IBM Instana(watsonx)·Elastic·Grafana·BMC HelixGPT . . . . . 36
    - 3.1.3 ITSM·AIOps 통합 진영 2사 — ServiceNow Now Assist·LogicMonitor Edwin AI . . . . . 38
  - 3.2 데이터 주권·폐쇄망 적합도 매트릭스 — Self-hosted / Air-gap / BYO LLM 3축 39
    - 3.2.1 11개 벤더 × Self-hosted × Air-gap × BYO LLM 적합도 매트릭스 . 39
    - 3.2.2 한국 진입 가능 5강 — Splunk Enterprise / IBM Instana / Elastic / BMC HelixGPT / Dynatrace Managed의 정렬 . . . . . 41
  - 3.3 BYO LLM의 부상 — 2025~2026 글로벌 패러다임 전환 . . . . . 42

- 3.3.1 BYO LLM 채택 진영(Elastic/BMC/Grafana/IBM) vs 고정 LLM 진영 (ServiceNow/Dynatrace/Datadog) . . . . . 42
- 3.3.2 BYO LLM이 한국 규제·비용·sLLM 결합에 동시 작동하는 3가지 이유 . . 43
- 3.4 NL → Causal RCA → Agentic 워크플로의 단일 흐름을 잇는 글로벌 3사 — Dynatrace·Datadog·IBM Instana . . . . . 45
- 3.4.1 단일 흐름 4계층 — NL→쿼리·Causal RCA·Agentic 워크플로·MCP 외부 LLM의 직선 결합 . . . . . 45
- 3.4.2 Dynatrace·Datadog·IBM Instana 3사 평가 — 한국 적합도와의 충돌 진단 47
- 4장: 5대 CSP 자체 통합 관제 AI 전략 — AWS/Azure/GCP/OCI/Alibaba 48**
- 4.1 5대 CSP AI 운영 에이전트 비교 — DevOps Agent·SRE Agent·Gemini Cloud Assist·OCI GenAI·ARMS Copilot . . . . . 48
- 4.1.1 AWS DevOps Agent·Azure SRE Agent·GCP Gemini Cloud Assist 의 GA 현황과 agent-second 과금 . . . . . 49
- 4.1.2 OCI GenAI/Agent Hub·Alibaba ARMS Copilot — Oracle DB 워크로드 강점과 동남아 한정성 . . . . . 51
- 4.2 Microsoft Azure Arc — 멀티클라우드 단일 콘솔의 가장 공격적인 카드 . . . . . 52
- 4.2.1 Azure Arc·SRE Agent·AIOps Observability Agent의 멀티클라우드 단일 콘솔 비전 . . . . . 53
- 4.2.2 분산 추적·코드 레벨 RCA의 한계 — 독립 오피버빌리티 벤더와의 깊이 차이 54
- 4.3 한국 의사결정 시사점 — CSP-Native vs 독립 벤더의 균형, 한국 리전 가용성의 그림자 . . . . . 56
- 4.3.1 한국 리전 AI 기능 가용성 매트릭스 — CloudWatch Investigations·Korea Central·asia-northeast3 . . . . . 56
- 4.3.2 단일/멀티클라우드 비율로 본 분기 가이드 — CSP-Native 우선 vs 독립 벤더 우선 . . . . . 57
- 5장: DB 모니터링 AI 분화 — 생성형 자체 탑재 vs ML+MCP 외부 연동 59**
- 5.1 DB 6개 제품 분화 매트릭스 — SolarWinds DPA·Quest Foglight·Percona PMM·Datadog DBM·Oracle EM 24ai·MongoDB Atlas . . . . . 59

- 5.1.1 6개 제품 × AI 기능 × 모델 기반 × GA 매트릭스 . . . . . 59
- 5.1.2 분화 패턴 — 생성형 자체 탑재 진영(4개) vs ML+MCP 외부 연동 진영(2개) 61
- 5.2 생성형 자체 탑재 진영 분석 — Oracle Cohere·Quest 생성형·Datadog Bits AI·SolarWinds AI . . . . . 62
  - 5.2.1 Oracle EM 24ai + Cohere LLM on OCI·Select AI Agents — DB 도메인 통합 깊이의 표준 . . . . . 62
  - 5.2.2 Quest Foglight 8.1·SolarWinds DPA·Datadog DBM — 모델 비공개 의 한국 도입 함의 . . . . . 63
- 5.3 ML+MCP 외부 연동 진영 분석 — Percona PMM·MongoDB Atlas의 한국 금융 권 적합도 . . . . . 64
  - 5.3.1 Percona PMM 3.5 + MCP — OSS 진영의 BYO LLM 표준화 . . . . . 64
  - 5.3.2 MongoDB Atlas Performance Advisor + MCP — 클라우드 매니지드 진영의 BYO LLM 표준화 . . . . . 65
  - 5.3.3 한국 금융권·공공 도입 권장 시나리오 — Percona PMM/Atlas+MCP를 1차 도입 후보로 삼는 이유 . . . . . 66
- 6장: 리서치 기관 보고서가 말하는 AI 평가 기준의 진화 — Gartner / IDC / Forrester 67**
  - 6.1 Gartner MQ for Observability Platforms 2025 + Hype Cycle 2025 — 카테 고리 통합과 GenAI→Agentic AI 정점 이동 . . . . . 67
    - 6.1.1 MQ 2025 Leader 8사 — Datadog/Dynatrace/Splunk(Cisco)/Elastic/Grafana/IBM/Chronosphere/New Relic의 의미 . . . . . 67
    - 6.1.2 Hype Cycle 2025 — GenAI 환멸기 진입·Agentic AI 정점 이동의 의미 69
  - 6.2 Forrester Wave: AIOps Platforms Q2 2025 — Dynatrace·ScienceLogic·Datadog Leader . . . . . 70
    - 6.2.1 Leader 3사 — Dynatrace(Current Offering 최고)·ScienceLogic(Strategy 최고)·Datadog . . . . . 71
    - 6.2.2 Forrester 핵심 인용 — “GenAI 평준화·Agentic 통합 깊이 차별화” . . . 72
  - 6.3 IDC MarketScape 2025·2026 + 평가 기준의 진화 (2024→2025→2026) . . . 73

- 6.3.1 IDC MarketScape Worldwide AIOps 2026 + Worldwide Observability Platforms 2025의 Leader . . . . . 74
- 6.3.2 평가 기준의 진화 — 2024 ‘AI 어시스턴트’ → 2025 ‘Agentic 폐쇄 루프’ → 2026 ‘Outcome-centric, governed automation’ . . . . . 75
- 7장: 한국 시장의 4대 구조적 제약과 글로벌 벤더의 한계** . . . . . 77
  - 7.1 한국 4대 규제 제약 — 망분리·PIPA·국정원 N²SF·CSAP . . . . . 77
    - 7.1.1 망분리 — 금융위 SaaS 개선 로드맵·국정원 N²SF의 SaaS-AI 진입 차단 . 77
    - 7.1.2 PIPA 4단계 분류·2025-08 PIPC 생성형 AI 안내서 — 텔레메트리 PII 외부 LLM 전송의 별도 적법 근거 . . . . . 78
    - 7.1.3 CSAP 2026 변화 — 보안적합성 검증 도입과 “상·중” 등급의 외산 차단 지속 79
  - 7.2 글로벌 AI 모니터링 기능의 한국 도입 가능성 분류 — 즉시 가능 / 한국형 변형 / 도입 불가 . . . . . 80
    - 7.2.1 즉시 도입 가능 — Dynatrace Davis Predictive·Splunk MLTK·IBM Instana watsonx 자체 호스팅 . . . . . 80
    - 7.2.2 한국형 변형 필요 — Elastic·Grafana·BMC HelixGPT의 BYO LLM에 한국어 sLLM 결합 . . . . . 81
    - 7.2.3 도입 사실상 불가 — Datadog Bits AI·PagerDuty SRE Agent·AWS DevOps Agent의 외부 LLM 고정 . . . . . 82
  - 7.3 글로벌 벤더의 한국 도입 좌초 패턴 — 망분리·PIPA·콘솔 분리의 3중 게이트 . . . 84
    - 7.3.1 망분리·PIPA에서 좌초된 PoC 사례 패턴 — 6개월 PoC가 거버넌스 검토에서 무산된 구조 . . . . . 84
    - 7.3.2 SKU 분리·콘솔 분리의 사용자 경험 회귀 — “기능은 많은데 한국어 NL이 어색하고 SRE가 같은 SPL/DQL을 다시 쓴다” . . . . . 85
  - 7.4 “다양한 모니터링” 함정의 한국 시장 검증 — 단일 트러블슈팅 흐름이 한국형 답이라는 결론 . . . . . 86
    - 7.4.1 핵심 주제 #2 가설 검증 결론 — “다양한 모니터링” 백화점 모델의 한국 시장 부적합 . . . . . 87
    - 7.4.2 한국형 답의 4축 — Top-down + Bottom-up + 단일 콘솔 + 한국어 NL 87

<b>8장: 국내 모니터링 벤더 AI 도입 현황 — 출처 기반 실측</b>	<b>88</b>
8.1 국내 6개 벤더 AI 도입 현황 매트릭스 — 제니퍼소프트·와탭·엑셈·브레인즈컴퍼니·LG CNS·삼성SDS . . . . .	89
8.1.1 제품 벤더 4사 — JenniferAI · WhaTap OpsLake · eXemble · Zenius AI Agent . . . . .	89
8.1.2 SI 진영 2사 — LG CNS AIOps · 삼성SDS Fabrix/Brity Copilot . . . . .	91
8.1.3 6개 벤더 공통 패턴 — 단일 흐름·한국어·폐쇄망의 동시 추구 . . . . .	92
8.2 한국 시장 AI 도입의 공백 좌표 — 4개 미충족 영역 . . . . .	93
8.2.1 4개 공백 좌표 — BYO LLM 동시충족·MCP 표준 채택·Self-hosted 자체 ML RCA·OTel GenAI 의미 규약 . . . . .	93
8.2.2 4개 공백 좌표의 RFP 입력화 — 즉시 평가 가능한 항목으로 분해 . . . . .	94
8.3 한국 벤더 vs 글로벌 벤더 매핑 — 즉시 가능 / 한국형 변형 / 도입 불가 영역의 벤더별 위치 . . . . .	94
8.4 2026~2028 한국 벤더 AI 로드맵 시사점 — 골든 타임 24개월 가설 . . . . .	95
8.4.1 골든 타임 24개월 가설의 근거 — 글로벌 벤더의 한국 거버넌스 통과 부재	95
8.4.2 의사결정자 시사점 — 9장 7개 체크리스트로의 연결 . . . . .	96
<b>9장: 결론 및 권장사항 — 2026~2028 의사결정 체크리스트</b>	<b>96</b>
9.1 2026~2028 시점별 변곡점 — Agentic AIOps GA 대전환 / LLM 오피저버빌리티 융합 / 자율 remediation 일부 실용화 . . . . .	96
9.1.1 2026 — Agentic AIOps Preview→GA 대전환 / MCP 사실상 표준 . . . . .	96
9.1.2 2027 — LLM 오피저버빌리티 + APM 본격 융합 / OTel GenAI 안정화 / 한국 SaaS-AI 단계적 허용 본격화 . . . . .	97
9.1.3 2028 — 자율 remediation 일부 도메인 실용화 / 멀티모달 통합 오피저버빌리티 초기 진입 . . . . .	98
9.2 5대 전략적 합의 — BYO LLM+MCP·SKU 통합 콘솔·Agentic 단계적 도달·CSP-Native vs 독립·국산 골든 타임 . . . . .	100
9.2.1 합의 1·2·3 — 벤더 종속 회피의 새 표준(BYO LLM + MCP) / SKU 통합 콘솔 / Agentic 자율 운영의 단계적 도달 . . . . .	100

9.2.2 합의 4·5 — CSP-Native vs 독립 벤더 균형 동요 / 국산 옴저버빌리티의  
 골든 타임 . . . . . 101

9.3 의사결정 체크리스트 — 7개 평가 항목으로 도입 결정하기 . . . . . 101

9.3.1 의사결정 체크리스트 7개 항목 — BYO LLM · MCP · Self-hosted · 자체  
 ML RCA · 인간 승인 게이팅 · 한국 거버넌스 · OTel GenAI . . . . . 102

9.3.2 체크리스트 사용법 — RFP · Vendor Scorecard · 법무 검토 3개 단계 적용 103

부록: 전체 출처 목록 . . . . . 105

**Appendix . . . . . 115**

References . . . . . 115

Glossary . . . . . 125

Endnotes . . . . . 127

## 1.1 모니터링에서 옹저버빌리티로, 옹저버빌리티에서 Agentic AIOps로 — 2010년대~2026년 시장 변천사

AI 시대의 옹저버빌리티는 단순한 모니터링의 확장판이 아닙니다. 2010년대의 APM(Application Performance Monitoring)과 전통적 통합관제에서 출발한 시장은, 2020년대 초반 메트릭·로그·트레이스의 3축 옹저버빌리티로 진화했고, 2025년을 기점으로 Agentic AIOps라는 새로운 패러다임으로 이동하고 있습니다. 이 변화의 배경에는 클라우드 네이티브, 마이크로서비스, 그리고 AI 기반 자동화의 급속한 확산이 자리잡고 있습니다. 본 절에서는 시장의 3단계 변천사, 전환을 가속화한 4가지 구조적 동인, 그리고 Gartner와 Anthropic이 주도한 카테고리 통합 및 표준화의 변곡점을 심층적으로 분석합니다.

### 1.1.1 APM·통합관제의 분리 시대(2010년대) → 3축 옹저버빌리티(2020~2022) → Agentic AIOps(2025~)

2010년대 이후 IT 인프라 운영 시장은 빠르게 변화해왔습니다. 초기에는 각기 분리된 도구와 프로세스가 주류를 이루었으나, 클라우드 네이티브와 AI 기술의 발전으로 인해 통합과 자동화가 시장의 핵심 화두로 부상하게 되었습니다. 본 절에서는 2010년대의 APM과 통합관제의 분리, 2020년대 초반 3축 옹저버빌리티의 표준화, 그리고 2025년 이후 Agentic AIOps의 부상까지, 시장의 주요 변곡점과 기술적 진화 과정을 구체적으로 살펴봅니다.

#### 2010년대 APM과 통합관제의 분리

2010년대 초반, IT 인프라 모니터링 시장은 APM(대표적으로 New Relic, Dynatrace, AppDynamics)과 SNMP/EMS 기반 통합관제 시스템이 명확히 분리되어 있었습니다. APM은 애플리케이션 레벨의 성능 추적에 집중했고, 통합관제는 네트워크·서버·스토리지 등 인프라 레이어의 상태를 모니터링했습니다. 이 시기에는 각 도구가 제공하는 데이터의 범위와 분석 깊이에 한계가 있었으며, 장애 원인 추적(RCA) 역시 별도의 수작업 프로세스로 진행되는 경우가 많았습니다.

이러한 분리 구조는 운영 효율성 저하, 데이터 사일로, 그리고 장애 대응 속도의 한계 등 여러 문제를 야기하였습니다. 예를 들어, 애플리케이션 장애와 인프라 장애가 동시에 발생할 경우, 운영자는 두 개 이상의 시스템을 오가며 원인을 추적해야 했고, 이로 인해 MTTR이 길어지는 현상이 빈번했습니다. 또한, 각 도구의 데이터 포맷과 분석 방식이 상이하여 통합적인 시야 확보가

어려웠습니다.

### 2020~2022년: 3축 옴저버빌리티의 보편화

클라우드 네이티브와 마이크로서비스 아키텍처의 확산으로, 단일 애플리케이션·인프라 모니터링만으로는 복잡한 장애의 근본 원인을 파악하기 어려워졌습니다. 이에 따라 메트릭, 로그, 트레이스의 3축 데이터를 통합적으로 수집·분석하는 ‘옴저버빌리티’가 시장의 표준으로 자리잡았습니다. OpenTelemetry 프로젝트(2021년 CNCF Incubating 등록)는 이 3축 표준화의 결정적 계기가 되었으며, 주요 벤더들은 자사 제품에 3축 통합 기능을 빠르게 도입했습니다.

이 시기에는 데이터의 폭과 깊이가 크게 확장되었고, 운영자들은 다양한 소스의 정보를 한눈에 파악할 수 있게 되었습니다. 그러나 여전히 데이터의 해석과 장애 원인 분석은 인간의 경험에 크게 의존하였으며, 복잡한 쿼리 언어와 다수의 콘솔을 사용하는 불편함이 남아 있었습니다.

### 2025년 이후: Agentic AIOps의 부상

2025년을 전후로, 옴저버빌리티의 초점은 ‘데이터 수집 폭’에서 ‘트러블슈팅의 즉각 종결 (Time-to-Resolution 단축)’로 이동하고 있습니다. AI 기반 자동화, 특히 도구 호출·메모리·계획을 결합한 Agentic AIOps가 새로운 좌표로 부상했습니다. 이 단계에서는 LLM(대형 언어 모델) 기반 자연어 질의, 인과 추론(Causal AI), 자동화된 remediation까지 하나의 폐쇄 루프(closed loop)로 구현하는 것이 핵심 경쟁력이 되고 있습니다.

Agentic AIOps는 단순한 데이터 수집과 시각화를 넘어, 실제 장애 상황에서 원인 분석과 조치까지의 전 과정을 자동화하는 데 초점을 맞춥니다. LLM 기반 자연어 인터페이스와 MCP와 같은 표준 프로토콜의 도입으로, 운영자는 복잡한 쿼리 언어를 몰라도 자연어로 문제를 질의하고, AI가 자동으로 데이터를 분석하여 결론과 조치 방안을 제시할 수 있게 되었습니다. 이로써 운영 효율성과 장애 대응 속도가 비약적으로 향상되고 있습니다.

### 시장 변천사 요약 표

시기	대표 기술	대표 벤더	평가 키워드
2010년대	APM, SNMP/EMS	New Relic, Dynatrace, AppDynamics	분리된 데이터, 수작업 RCA
2020~2022	3축 옴저버빌리티	Datadog, Splunk, Elastic	메트릭·로그·트레이스 통합, OpenTelemetry
2025~	Agentic AIOps	Datadog, Dynatrace, IBM Instana	AI 기반 자동화, 폐쇄 루프, Time-to-Resolution

당신의 조직은 지금 어느 단계에 위치해 있습니까?

## 1.1.2 4가지 가속 페달 — 텔레메트리 폭증·MTTR 압력·SRE 채용난·GenAI 비용 급락

옵저버빌리티 시장이 빠르게 Agentic AIOps로 전환되는 데에는 단순한 기술 발전만이 아니라, 네 가지 구조적 동인이 복합적으로 작용하고 있습니다. 텔레메트리 데이터의 폭증, 장애 복구 시간(MTTR) 단축에 대한 압력, SRE 인력의 만성적 부족, 그리고 LLM 기반 GenAI 도입 비용의 급락이 바로 그 요인입니다. 이 네 가지는 글로벌뿐 아니라 국내 시장에서도 동일하게 체감되고 있으며, 실제로 각 동인이 어떻게 시장 변화를 촉진하는지 구체적으로 살펴보겠습니다.

### 텔레메트리 데이터 폭증

클라우드 네이티브, 컨테이너, 마이크로서비스 환경이 확산되면서, 텔레메트리 데이터의 양이 기하급수적으로 증가했습니다. 하루 수십만~수백만 건의 이벤트와 로그가 생성되며, 이는 인간 운영자의 분석 한계를 초과합니다. 실제로 글로벌 대형 서비스에서는 일평균 알람 수가 SRE 1인당 100~500건에 달하고, 이 중 70% 이상이 false positive로 판명됩니다.

이러한 데이터 폭증은 기존 수작업 분석 방식의 한계를 명확히 드러내고 있습니다. 운영자는 알람의 진위를 일일이 판별해야 하며, 중요한 장애 신호를 놓칠 위험이 커집니다. 이로 인해 자동화 도구와 AI 기반 분석의 도입이 필수로 인식되고 있습니다.

### MTTR 단축 압력

디지털 비즈니스의 경쟁이 심화되면서, 장애 발생 시 서비스 복구까지 걸리는 시간(MTTR: Mean Time To Resolution)을 단축해야 한다는 압력이 커졌습니다. 평균 MTTR이 60~240분에 이르는 환경에서는, 장애가 장기화될수록 비즈니스 손실이 기하급수적으로 커집니다. 이에 따라, 장애 탐지부터 근본 원인 분석(RCA), 조치까지의 전체 사이클을 단축하는 것이 핵심 KPI로 부상했습니다.

MTTR 단축은 단순히 운영 효율성의 문제가 아니라, 기업의 매출 및 고객 신뢰도와 직결되는 중요한 경영 지표로 자리잡았습니다. 실제로 AI 기반 자동화 도입 후 MTTR이 절반 이하로 감소한 사례가 다수 보고되고 있습니다.

## SRE 채용난

SRE(Site Reliability Engineer) 인력은 글로벌 시장뿐 아니라 한국에서도 심각한 공급 부족 현상을 겪고 있습니다. 복잡한 분산 시스템을 운영·분석할 수 있는 고급 인력의 연봉은 매년 상승하고 있으며, SRE 1인당 담당 시스템의 복잡도 또한 증가하고 있습니다. 이는 AI 기반 자동화 도입의 필요성을 더욱 가속화하는 요인입니다.

특히, SRE 채용난은 중견·중소기업에서 더욱 심각하게 나타나고 있습니다. 인력 부족으로 인해 운영팀의 업무 부담이 가중되고, 자동화 도구 도입에 대한 요구가 높아지고 있습니다. AI 기반 솔루션은 인력 부족 문제를 완화하고, 운영의 표준화와 효율화를 동시에 달성할 수 있는 대안으로 주목받고 있습니다.

## GenAI 등장과 비용 급락

2023~2025년 사이, LLM(Generative AI) 기반 자연어 인터페이스의 도입 비용이 급격히 하락했습니다. OpenAI, Anthropic, Google 등 주요 LLM 서비스의 토큰 단가는 1년 새 50% 이상 하락했고, 한국어 지원도 빠르게 개선되었습니다. 이로 인해, 자연어 질의 기반 트러블슈팅, NL→쿼리 자동화, 자동 RCA 등 AI 기반 기능이 ‘옵션’이 아닌 ‘필수’로 자리잡게 되었습니다.

이러한 비용 구조 변화는 대기업뿐 아니라 중소기업, 스타트업에서도 AI 기반 옴져버빌리티 도입을 현실적인 선택지로 만들었습니다. 특히 한국어 지원의 비약적 향상은 국내 기업의 도입 장벽을 크게 낮추었습니다.

동인	정량 지표 예시	시장 변화 촉진 효과
텔레메트리 폭증	일평균 알람 100~500건/SRE	수작업 분석 불가, 자동화 필요
MTTR 압력	평균 MTTR 60~240분	장애 복구 시간 단축 요구
SRE 채용난	SRE 연봉 연 10~20% 상승	AI 자동화 도입 가속
GenAI 비용 급락	LLM 토큰 단가 연간 50% ↓	NL 인터페이스 대중화, AI 보급 확대

이 4가지 동인은 한국 시장에도 동일하게 적용되고 있습니다. 특히 SRE 채용난과 GenAI 한국어 비용 측면에서 국내 기업의 체감도가 높아지고 있습니다.

### 1.1.3 Gartner 2025의 “Observability Platforms” 카테고리 통합과 MCP 표준화의 결정적 변곡점

2025년은 옵저버빌리티 시장에서 중요한 변곡점으로 기록될 것입니다. Gartner의 카테고리 통합과 Anthropic이 주도한 MCP(Model Context Protocol) 표준화는 단순한 용어 변경이나 기술적 사양의 변화가 아니라, 옵저버빌리티 도구의 역할과 평가 기준 자체를 근본적으로 재정하는 사건입니다. 본 절에서는 Gartner의 카테고리 통합 배경과 MCP 표준화가 시장에 미치는 영향, 그리고 실제 도입 평가 시 반드시 고려해야 할 핵심 포인트를 구체적으로 설명합니다.

#### Gartner의 카테고리 통합

Gartner는 2024년까지 ‘APM and Observability’로 분리되어 있던 시장 카테고리를 2025년부터 ‘Observability Platforms’로 통합했습니다(G00813330, 2025-07-07). 이는 단순한 용어 변경이 아니라, APM, 로그, 메트릭, 트레이스, 이벤트, 프로파일 등 모든 관찰 데이터와 분석·자동화 기능을 하나의 플랫폼 내에서 통합적으로 제공해야 한다는 시장 표준의 변화를 의미합니다. 이에 따라, RFP(입찰 제안서)나 Vendor Scorecard에서 ‘APM/Logs/Metrics’ 등 개별 SKU 기준을 삭제하고, ‘단일 플랫폼’ 제공 여부와 AI 기반 자동화의 깊이를 평가 기준으로 삼아야 합니다.

이러한 변화는 벤더의 제품 전략뿐 아니라, 실제 구매자와 운영자의 평가 방식에도 큰 영향을 미치고 있습니다. 단일 플랫폼 제공이 필수 요건이 되면서, 여러 개의 모듈을 조합하는 방식에서 벗어나 통합된 사용자 경험과 자동화 수준이 핵심 경쟁력으로 부상하고 있습니다.

#### MCP(Model Context Protocol) 표준화

2023년 11월 Anthropic이 발표한 MCP(Model Context Protocol)는 LLM 에이전트가 외부 도구(옵저버빌리티, ITSM 등)와 표준화된 방식으로 데이터를 주고받을 수 있도록 정의한 프로토콜입니다. 이 MCP는 1년 6개월 만에 OpenAI Agents SDK, VS Code Agent Mode, Cursor 등 주요 LLM 에이전트 클라이언트의 표준 입력 채널로 채택되었습니다. 옵저버빌리티 도구가 LLM 에이전트의 ‘표준 입력 채널’로 재정된 이 변곡점은, 향후 1~2년간 옵저버빌리티 솔루션 평가 시 ‘MCP 호환성’이 필수 항목이 되어야 함을 의미합니다.

MCP 표준화는 LLM 에이전트와 옵저버빌리티 도구 간의 연동을 획기적으로 단순화하고, 다양한 벤더 솔루션 간의 상호운용성을 보장합니다. 실제로 MCP를 지원하지 않는 도구는 LLM 기반 트러블슈팅 자동화에서 배제될 위험이 커지고 있습니다. 따라서, 사내 RFP나 벤더 평가 시

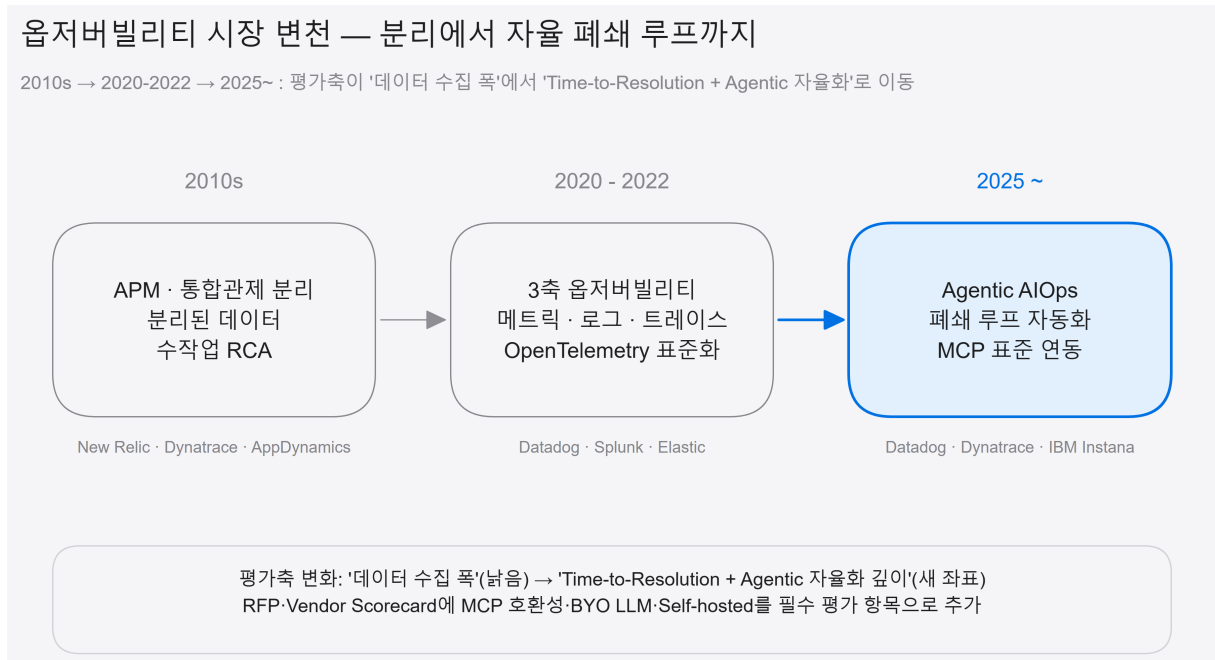
MCP 호환성 여부를 반드시 확인해야 하며, 이는 향후 2~3년간 시장에서 중요한 차별화 요소가 될 것입니다.

연도	Gartner 카테고리 정의 변화
2024	APM and Observability (분리)
2025	Observability Platforms (통합)

### MCP 채택 타임라인 예시

- 2023-11: Anthropic MCP 발표
- 2024-11~2025-05: OpenAI Agents SDK, VS Code Agent Mode, Cursor 등 주요 에이전트 클라이언트 MCP 표준 채택
- 2025-07: Gartner MQ 평가 기준에 MCP 호환성 공식 반영

사내 RFP 작성 시, “MCP 호환성 여부” 를 필수 조항으로 추가해야 합니다.



[그림 1] 옵저버빌리티 시장 변천 — 분리에서 자율 폐쇄 루프까지|739

## 1.2 트러블슈팅 즉각 종결 패러다임 — “데이터 수집 폭”에서 “Time-to-Resolution” 으로

AI 시대 옴저버빌리티의 본질은 더 이상 ‘얼마나 많은 데이터를 수집하는가’가 아니라, ‘얼마나 빨리, 정확하게 문제를 해결할 수 있는가’에 있습니다. 본 절에서는 AI 도입 이전 레거시 옴저버빌리티의 5대 한계, 글로벌 벤더의 SKU·콘솔 분리 패턴이 초래하는 사용자 경험 단절, 그리고 Time-to-Resolution을 측정 가능한 KPI로 정의하는 방법을 구체적으로 제시합니다. 이를 통해 의사결정자가 “수집 폭”이 아닌 “트러블슈팅 종결 속도”를 평가의 핵심 축으로 삼아야 함을 데이터로 증명합니다.

### 1.2.1 레거시 옴저버빌리티의 5대 한계 — 알람 폭증·RCA 가설 인간 의존·Top·Bottom 단절·NL 부재·다양한 모니터링 함정

레거시 옴저버빌리티 환경은 여러 가지 구조적 한계를 안고 있습니다. 알람의 폭증과 높은 오탐률, 근본 원인 분석(RCA)에서 인간의 직관에 대한 과도한 의존, 비즈니스와 인프라 지표 간의 단절, 자연어 인터페이스의 부재, 그리고 다양한 모듈 및 콘솔의 분리로 인한 사용자 경험 저하 등이 대표적입니다. 이러한 한계들은 AI 기반 자동화 도입 이전까지 대부분의 기업에서 공통적으로 경험하는 문제로, 실제 운영 현장에서 업무 효율성과 장애 대응 속도를 크게 저해하는 요인으로 작용해왔습니다.

#### 알람 폭증과 False Positive

레거시 옴저버빌리티 환경에서는 SRE 1인당 하루 100~500건의 알람이 발생하는 것이 일반적입니다. 이 중 70% 이상이 false positive(오탐)로, 실제로 대응이 필요한 장애와 무관한 경우가 대부분입니다. 이는 SRE팀의 피로도를 극대화하고, 중요한 장애 신호의 탐지율을 저하시킵니다.

알람 폭증은 운영팀의 업무 부담을 가중시키고, 실제로 중요한 장애 신호를 놓칠 위험성을 높입니다. 오탐이 반복되면 운영자는 알람에 무감각해지는 ‘알람 피로(Alarm Fatigue)’ 현상에 빠지게 되며, 이는 장애 대응의 신속성과 정확성 모두에 부정적인 영향을 미칩니다. 실제로 글로벌 금융, 이커머스, 게임 등 대규모 트래픽을 처리하는 산업군에서 알람 피로로 인한 장애 대응 실패 사례가 다수 보고되고 있습니다.

#### RCA(근본 원인 분석) 가설의 인간 의존

장애 발생 시, 평균 MTTR(60~240분) 중 70%가량이 ‘어디서부터 분석을 시작할지’ 결정하는 가설 단계에 소요됩니다. 이 과정은 전적으로 인간 운영자의 경험과 직관에 의존하며, 반복적인 쿼리 작성과 데이터 탐색이 필수적입니다.

이러한 인간 의존적 구조는 운영자의 역량과 경험에 따라 장애 대응 품질이 크게 달라지는 문제를 야기합니다. 신규 인력이나 경험이 부족한 운영자는 장애 원인 분석에 더 많은 시간이 소요되고, 반복적인 쿼리 작성과 데이터 탐색 과정에서 실수가 발생할 위험도 높아집니다. AI 기반 RCA 자동화가 도입되면, 이러한 가설 수립 과정을 대폭 단축하고, 운영자의 역량 편차를 최소화할 수 있습니다.

### Top-Down·Bottom-Up 단절

비즈니스 KPI(매출, 전환율 등)와 인프라 메트릭(서버 CPU, 네트워크 트래픽 등) 간의 인과 관계 분석이 단절되어 있습니다. 즉, 비즈니스 이상 현상이 감지되어도, 인프라 레이어에서 이를 추적하는 과정이 매끄럽지 않으며, 반대로 인프라 장애가 비즈니스에 미치는 영향을 신속하게 파악하기 어렵습니다.

이로 인해, 장애가 발생했을 때 비즈니스 임팩트와 인프라 원인 간의 연결고리를 찾는 데 많은 시간이 소요됩니다. 예를 들어, 결제 실패율이 급증한 경우, 해당 이슈가 네트워크, DB, 애플리케이션 중 어디에서 기인한 것인지 신속하게 파악하기 어렵습니다. AI 기반 인과 추론(Causal AI)이 도입되면, 이러한 단절을 자동으로 연결하여, 비즈니스-인프라 간의 흐름을 실시간으로 추적할 수 있습니다.

### NL(자연어) 인터페이스 부재

PromQL, SPL, DQL, NRQL 등 각 벤더별 쿼리 언어의 학습 곡선이 높아, 비전문가(비-SRE, Dev, QA 등)는 실질적으로 도구를 활용하기 어렵습니다. NL 인터페이스가 부재한 환경에서는 동일한 쿼리를 반복적으로 수작업으로 작성해야 하며, 이는 생산성 저하로 이어집니다.

특히, 다양한 부서와 직무에서 오피버빌리티 도구를 활용해야 하는 대기업 환경에서는 쿼리 언어의 진입 장벽이 협업과 정보 공유를 저해하는 요인으로 작용합니다. 자연어 인터페이스 도입 시, 누구나 쉽게 장애 현황을 질의하고, 빠르게 결론에 도달할 수 있는 환경이 조성됩니다.

### 다양한 모니터링 합정

여러 개의 SKU(예: APM, Logs, Metrics, Network, Synthetic)를 모두 도입해도, 실제 트러블슈팅 시에는 4개 이상의 콘솔을 동시에 띄워야 하는 사용자 경험 단절이 발생합니다.

이러한 콘솔 분리는 운영자의 업무 효율성을 저하시킬 뿐 아니라, 장애 원인 추적의 연속성을

방해합니다. 각 콘솔마다 데이터 포맷과 시각화 방식이 달라, 정보를 종합적으로 해석하는 데 추가적인 노력이 필요합니다. 단일 콘솔 통합과 AI 기반 자동화가 도입되면, 이러한 단절을 극복하고, 트러블슈팅의 전체 과정을 하나의 흐름으로 연결할 수 있습니다.

한계	정량 지표 예시	AI 도입 시 개선 폭
알람 폭증·False Positive	100~500건/일, 70%+	30~60% 알람 감소
RCA 가설 인간 의존	MTTR의 70% 소요	RCA 자동화로 50% 단축
Top/Bottom 단절	KPI-메트릭 연결 부재	인과 추론 자동화
NL 인터페이스 부재	쿼리 언어 학습 필요	NL 질의로 3~5배 활용도 증가
다양한 모니터링 함정	4개 콘솔 동시 사용	단일 콘솔 통합

각 한계에 대해 “우리 조직에서도 동일하게 발생하는가?” 를 체크해보시기 바랍니다.

## 1.2.2 “다양한 모니터링”의 함정 — APM/Logs/Metrics/Network/Synthetic 5개 SKU·4개 콘솔 분리 패턴

오피저빌리티 도구를 도입한 많은 기업들이 실제 트러블슈팅 현장에서 겪는 가장 큰 문제 중 하나는, 다양한 모듈(SKU)과 콘솔의 분리로 인한 사용자 경험의 단절입니다. 글로벌 벤더들은 각 기능을 별도의 SKU로 판매하는 전통적 모델을 고수해왔으며, 이는 실제 장애 대응 시 운영자에게 불필요한 복잡성과 비효율을 초래합니다. 본 절에서는 SKU·콘솔 분리 구조의 한계와, 실제 한국 시장에서 나타난 사례, 그리고 트러블슈팅 과정에서 발생하는 반복적 콘솔 전환의 문제점을 구체적으로 설명합니다.

### SKU·콘솔 분리의 구조적 패턴

글로벌 벤더들은 기능별로 모듈(SKU)을 분리해 판매하는 전통적 모델을 유지해왔습니다. 예를 들어, APM, Logs, Metrics, Network, Synthetic 모듈을 각각 별도 SKU로 제공하며, 고객은 장애 트러블슈팅 시 4개 이상의 콘솔을 오가야 합니다. 이는 “기능은 많지만, 정작 장애가 발생했을 때 원인을 한눈에 파악할 수 없다”는 사용자 경험 단절로 이어집니다.

이러한 구조는 실제 운영 현장에서 여러 가지 문제를 야기합니다. 첫째, 각 콘솔마다 로그인, 권한 설정, 데이터 시각화 방식이 달라 운영자가 동일한 장애에 대해 여러 번 정보를 재확인해야 합니다. 둘째, 장애 원인 분석 과정에서 데이터를 수동으로 복사하거나, 서로 다른 콘솔 간에 쿼리를

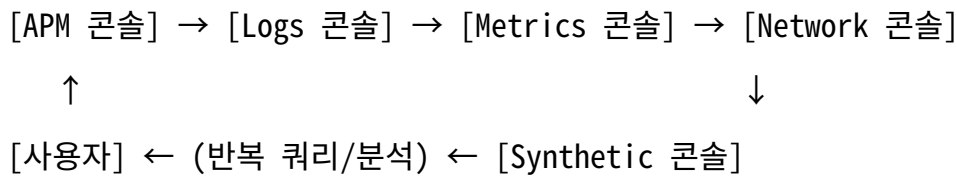
반복 작성해야 하는 비효율이 발생합니다. 셋째, 각 SKU별로 라이선스와 비용이 분리되어 있어, 전체 운영 비용이 증가하는 부작용도 나타납니다.

### 한국 시장의 실제 사례

한국의 한 대형 금융사는 “우리는 글로벌 벤더의 APM, 로그, 메트릭, 네트워크, 시뮬레이션 모듈을 모두 도입했지만, 장애가 발생하면 각 콘솔을 번갈아 띄우고, SRE가 다시 쿼리를 직접 작성해야 한다”고 토로합니다. 이처럼 SKU 수가 많다고 해서 트러블슈팅이 쉬워지는 것이 아니라, 오히려 콘솔 분리로 인해 업무 효율성이 저하되는 역설적 현상이 발생합니다.

이 사례는 단순히 한 기업의 문제가 아니라, 다수의 대기업·공공기관에서 공통적으로 보고되는 현상입니다. 실제로 장애가 발생했을 때, 운영자는 APM 콘솔에서 애플리케이션 로그를 확인하고, 별도의 로그 콘솔에서 에러 메시지를 탐색한 뒤, 다시 메트릭 콘솔로 이동해 시스템 지표를 분석해야 합니다. 이 과정에서 동일한 장애에 대해 여러 번 클릭과 쿼리 작성이 반복되며, 장애 대응 속도가 현저히 저하됩니다.

### 다이어그램: 트러블슈팅 1회당 콘솔 전환 시각화



이와 같이 트러블슈팅 한 번에 4개 이상의 콘솔을 전환해야 하는 구조적 한계는, 운영자의 업무 효율성 저하와 장애 대응 속도 저하로 직결됩니다. AI 기반 단일 콘솔 통합과 자연어 인터페이스 도입이 이러한 한계를 극복하는 핵심 솔루션으로 부상하고 있습니다.

*트러블슈팅 한 번에 4개 이상의 콘솔을 전환해야 하는 구조적 한계가 존재합니다.*

## 1.2.3 Time-to-Resolution의 측정 가능 정의 — conclusive investigation·MTTR·NL→결론까지의 클릭 수

트러블슈팅의 효율성을 객관적으로 평가하기 위해서는 Time-to-Resolution을 정량적 KPI로 정의하고, 실제 운영 환경에서 측정 가능한 지표로 관리해야 합니다. 본 절에서는 Time-to-Resolution의 세부 정의, 글로벌 벤더의 실제 도입 사례, 그리고 AI 도입 전후의 KPI 변화 측정 방법을 구체적으로 제시합니다. 이를 통해, 의사결정자가 “얼마나 많은 데이터를 수집했는가”가 아니라 “얼마나 빠르게 결론에 도달했는가”를 평가의 핵심 축으로 삼을 수 있습니다.

### 정량적 KPI로서의 Time-to-Resolution

Time-to-Resolution은 이제 추상적 개념이 아니라, 측정 가능한 KPI로 정의되어야 합니다. 대표적으로 다음 3가지 지표가 있습니다.

- **Conclusive Investigation 평균 소요 시간:** Datadog Bits AI SRE 적용 시, 장애 1건당 conclusive investigation(최종 결론 도달)까지 평균 8~12분이 단축된 사례가 보고되었습니다.
- **평균 MTTR:** 전체 장애 복구까지 걸리는 평균 시간(60~240분). AI 도입 전후의 MTTR 변화를 측정해야 합니다.
- **NL 질의→결론까지의 클릭 수·콘솔 전환 횟수:** 자연어 질의 후 최종 결론(예: 장애 원인, 조치 방안)까지 필요한 클릭 수, 콘솔 전환 횟수를 계량화합니다.

이러한 KPI는 실제 운영 환경에서 쉽게 측정할 수 있으며, PoC(개념 검증) 단계에서 벤더별 솔루션의 효과를 객관적으로 비교하는 데 활용할 수 있습니다. 예를 들어, 동일한 장애 시나리오에서 AI 기반 솔루션 도입 전후의 평균 소요 시간, 클릭 수, 콘솔 전환 횟수를 비교함으로써, 도입 효과를 수치로 증명할 수 있습니다.

또한, Time-to-Resolution을 KPI로 관리하면, 운영팀의 업무 효율성뿐 아니라, 장애 대응 프로세스의 표준화와 자동화 수준까지 체계적으로 개선할 수 있습니다. 실제로 글로벌 벤더들이 이 지표를 바탕으로 자사 솔루션의 ROI(투자 대비 효과)를 고객에게 명확히 제시하고 있습니다.

KPI	측정 방법 예시	AI 도입 전	AI 도입 후
Conclusive Investigation	장애 1건당 결론 도달까지 평균 소요 시간	20~30분	8~12분
MTTR	장애 발생~복구까지 평균 시간	120분	60분
NL→결론 클릭 수/콘솔 전환	NL 질의 후 결론까지 클릭/콘솔 전환 횟수	10회/4회	3회/1회

각 KPI는 AI 도입 전/후 동일한 방식으로 측정해야 하며, PoC 평가서에 그대로 복사해 활용할 수 있습니다.

## 1.3 IT 의사결정자에게 주는 1장의 결론 — 평가 좌표를 “Time-to-Resolution + Agentic AI 자율화 깊이” 로 재설정하라

본 장의 분석을 통해, 옴저버빌리티 도입·고도화의 평가 기준은 더 이상 ‘데이터 수집 폭’이나 ‘SKU 수’가 아니라, ‘트러블슈팅의 즉각 종결(Time-to-Resolution)’과 ‘Agentic AI 자율화의 깊이’임이 명확해졌습니다. 본 절에서는 평가 좌표 재설정 가이드와, 다음 장에서 다룰 7개 필수 개념의 예고를 통해, 의사결정자가 즉시 행동에 옮길 수 있는 실천적 결론을 제시합니다.

### 1.3.1 평가 좌표 재설정 가이드 — 데이터 수집 폭(넓음) → Time-to-Resolution + Agentic AI 자율화 깊이(새 좌표)

옴저버빌리티 도입 및 벤더 평가 시, 기존의 평가 기준은 더 이상 시장의 요구를 반영하지 못합니다. 지원 데이터 소스 수, 대시보드 템플릿 수, SKU 수 등은 기능의 폭만을 보여줄 뿐, 실제 트러블슈팅의 효율성이나 자동화 수준을 평가하는 데 한계가 있습니다. 이제는 Time-to-Resolution과 Agentic AI 자율화의 깊이를 중심으로 평가 좌표를 재설정해야 하며, 이를 위해 RFP와 Vendor Scorecard의 항목을 대대적으로 개편할 필요가 있습니다.

#### RFP·Vendor Scorecard에서 빼야 할 항목

기존 평가 항목 중 ‘지원 데이터 소스 수’, ‘대시보드 템플릿 수’, ‘SKU 수’ 등은 더 이상 핵심 가치가 아닙니다. 이들은 단순히 기능의 폭을 보여줄 뿐, 실제 트러블슈팅의 효율성이나 자동화 수준을 평가하지 못합니다.

이러한 항목은 벤더의 마케팅 자료에는 자주 등장하지만, 실제 운영 현장에서는 장애 대응 속도나 자동화 수준과 직접적인 상관관계가 낮습니다. 따라서, 평가 기준에서 과감히 제외하는 것이 바람직합니다.

#### 추가해야 할 핵심 평가 항목

- ‘NL 질의 → conclusive investigation 평균 소요 시간’ (분 단위)
- ‘Agentic 워크플로 인간 승인 게이팅 정책 명시 여부’
- ‘MCP(Model Context Protocol) 호환성’
- ‘BYO LLM(Bring Your Own LLM) 지원 여부’
- ‘Self-hosted/On-prem 옵션’

- ‘Causal AI 기반 RCA 지원 여부’

이러한 항목들은 실제 트러블슈팅의 효율성과 자동화 수준, 그리고 향후 확장성까지 평가할 수 있는 핵심 지표입니다. 특히, MCP 호환성과 BYO LLM 지원 여부는 LLM 기반 자동화의 유연성과 데이터 주권 확보 측면에서 매우 중요한 평가 요소입니다.

**빼기/더하기 비교 표**

기존 항목(빼기)	새 항목(더하기)
지원 데이터 소스 수	NL 질의→결론까지 평균 소요 시간
대시보드 템플릿 수	Agentic 워크플로 인간 승인 게이팅 정책
SKU 수(기능별 모듈 수)	MCP 호환성
BYO LLM 지원 여부	
Self-hosted/On-prem 옵션	
Causal AI 기반 RCA 지원	

다음 RFP 사이클부터는 위 표의 ‘더하기’ 항목을 반드시 평가 기준에 포함시키시기 바랍니다.

**1.3.2 2장 예고 — 평가 좌표를 정확히 사용하기 위한 7개 필수 개념과 의사결정 프레임**

1장에서 제시한 평가 좌표를 실제로 적용하기 위해서는, 옴저버빌리티와 Alops, Agentic AI, Causal AI 등 핵심 개념에 대한 명확한 이해가 선행되어야 합니다. 2장에서는 이러한 7개 필수 개념의 정의와, 의사결정자가 RFP와 Vendor Scorecard 작성 시 활용할 수 있는 평가 프레임을 구체적으로 안내할 예정입니다. 각 개념의 차이점과 실제 평가에 미치는 영향을 정확히 파악함으로써, 벤더의 마케팅 문구에 휘둘리지 않고, 실질적인 가치를 평가할 수 있는 기준을 갖추게 됩니다.

**7개 필수 개념 예고**

1. **Observability vs Monitoring:** ‘알려진 문제 추적’(Monitoring)과 ‘외부 출력으로 내부 상태 추론’(Observability)의 차이를 명확히 구분해야, RFP 작성 시 벤더의 마케팅 문구에 휘둘리지 않습니다.
2. **AIOps:** 분석 계층의 ML/LLM 자동화가 어디까지 구현되어 있는지 평가의 핵심 축입니다.

3. **Agentic AI**: 도구 호출·계획·메모리 기반의 폐쇄 루프 자율화가 실제로 구현되어 있는지 판단해야 합니다.
4. **Causal AI**: 단순 상관 분석이 아닌 인과 추론 기반 RCA가 가능한지 확인이 필요합니다.
5. **Top-down·Bottom-up**: 비즈니스 KPI에서 인프라까지, 인프라에서 비즈니스까지 양방향 추적이 단일 흐름으로 연결되는지 평가합니다.
6. **MCP(Model Context Protocol)**: 오피저버빌리티 도구가 LLM 에이전트와 표준화된 방식으로 연동되는지, MCP 호환성 여부가 필수입니다.
7. **BYO LLM**: 벤더가 LLM을 고정 제공하는지, 고객이 직접 OpenAI/Bedrock/Vertex/사내 sLLM을 선택할 수 있는지 확인해야 합니다.

개념	1줄 설명	의사결정에 미치는 영향
Observability vs Monitoring	데이터 추적 vs 내부 상태 추론	RFP·Vendor Scorecard 용어 정렬
AIOps	ML/LLM 기반 자동화	분석 계층 자동화 수준 평가
Agentic AI	도구 호출·계획·메모리 기반 자율화	트러블슈팅 자동화·자율화 깊이 평가
Causal AI	인과 추론 기반 RCA	장애 원인 분석 정확도 평가
Top-down·Bottom-up	KPI↔인프라 양방향 추적	단일 흐름 구현 여부 평가
MCP	LLM 에이전트 연동 표준 프로토콜	LLM·Agent 연동 가능성 평가
BYO LLM	고객 LLM 직접 선택 가능	데이터 주권·한국어 NL 정확도·비용 통제 평가

2장에서는 위 7개 개념의 정의와 의사결정 프레임을 구체적으로 다룹니다. 1장에서 제시한 평가 좌표를 정확히 적용하려면 2장의 내용을 반드시 참고하시기 바랍니다.

## 2장: 핵심 개념과 용어 — Agentic AI / Causal AI / MCP / BYO LLM / Top-down·Bottom-up

### 2.1 IT 의사결정자가 30분 안에 흡수해야 할 7개 필수 개념

IT 의사결정자는 AI 기반 오피저버빌리티 플랫폼의 도입 및 평가 과정에서 핵심 개념을 명확히 이해해야만 올바른 의사결정을 내릴 수 있습니다. 본 절에서는 본 백서 전체에서 반복적으로 등장하는

7개 필수 개념을 정의, 동작 원리, 그리고 실제 의사결정에 미치는 함의 관점에서 압축 정리합니다. 이 7개 개념은 단순 용어 해설을 넘어, 실제 RFP(제안요청서)·Vendor Scorecard·법무 검토 등 실무 문서에 그대로 반영될 수 있도록 설계되었습니다. 각각의 개념이 시장의 기술 변화, 규제 환경, 그리고 트러블슈팅 프로세스의 혁신에 어떻게 연결되는지에 주목하며, 데이터 계층, 분석 계층, 실행 계층의 구분과 AI/LLM 시대의 새로운 평가 기준을 제시합니다.

### 2.1.1 Observability·Monitoring·AIOps — 데이터·분석·실행 3계층 구분

Observability, Monitoring, 그리고 AIOps는 IT 운영 환경에서 시스템의 상태를 파악하고, 장애를 탐지·분석·조치하는 데 필수적인 3대 축입니다. 이 세 가지 개념은 단순히 용어의 차이가 아니라, 데이터 수집·분석·실행이라는 계층적 구조와 실무 적용 방식에서 뚜렷한 차별점을 보입니다. 특히 최근 AI와 LLM의 도입으로 인해 각 계층의 역할과 경계가 빠르게 재정의되고 있으며, 벤더별 솔루션 평가 시 이들 계층의 구분이 실질적인 도입 효과와 직결됩니다. 따라서 IT 의사결정자는 각 계층의 정의, 한계, 그리고 상호 연계 방식을 명확히 이해해야만, 실제 운영 환경에서 최적의 플랫폼을 선택하고, 장애 대응의 효율성을 극대화할 수 있습니다.

#### Monitoring의 정의와 한계

Monitoring은 IT 인프라 및 애플리케이션에서 발생하는 알려진 문제를 사전에 정의된 임계값, 룰, 패턴을 통해 탐지하는 전통적 접근 방식입니다. 예를 들어 CPU 사용률이 90%를 초과하면 알람을 발생시키는 식입니다. 이러한 방식은 시스템의 정상/비정상 여부를 신속히 파악할 수 있지만, 미리 정의되지 않은 신규 장애나 복합적인 장애의 근본 원인(Root Cause)을 파악하는 데 한계가 있습니다. 또한, 복잡한 분산 시스템에서는 알람 폭증과 false positive(오탐)로 인해 운영자의 피로도가 급격히 증가합니다.

#### Observability의 진화와 데이터 플레인 확장

Observability는 단순 모니터링을 넘어, 시스템의 외부 출력(메트릭, 로그, 트레이스, 이벤트, 프로파일, 세션 등)을 종합적으로 수집·분석하여 내부 상태를 추론하는 개념입니다. 2020~2022년 OpenTelemetry의 등장과 함께 메트릭·로그·트레이스 3축이 표준화되었으며, 최근에는 이벤트, 프로파일, 세션 데이터까지 확장되고 있습니다. Observability의 목적은 알려지지 않은 장애, 복합 장애, 비즈니스 KPI와 인프라 이벤트 간의 인과 관계까지 추적하는 데 있습니다.

#### AIOps의 개념과 2025년 이후의 변화

AIOps는 Observability 데이터 위에 머신러닝(ML), 대형언어모델(LLM) 등 인공지능 기반 분석 계층을 더한 개념입니다. 2016년 Gartner가 최초로 정의한 AIOps는 알람 노이즈 감소, 이상 탐지, 근본 원인 분석(RCA) 자동화, 예측적 유지보수 등 ML 기반 자동화 기능을 포함합니다. 2025년 이후에는 단순 분석을 넘어, Agentic AIOps(에이전틱 자율 운영)로 의미가 확장되어, 탐지→진단→조치→검증까지의 폐쇄 루프(closed loop)를 AI가 자율적으로 수행하는 단계로 진화하고 있습니다.

### 3계층 비교와 실무 적용

Monitoring, Observability, AIOps는 데이터 계층(Observability), 분석 계층(AIOps), 실행 계층(Agentic AI)으로 구분됩니다. 실무에서는 이 3계층이 통합적으로 설계되어야 하며, 각 벤더의 솔루션이 어느 계층까지 커버하는지 명확히 구분해야 합니다. 예를 들어, 단순 모니터링 도구는 데이터 계층에 머물고, AIOps는 분석 계층까지, Agentic AI는 실행 계층까지 확장됩니다. 아래 표는 3계층의 주요 특징을 비교합니다.

계층	주요 기능	대표 기술/벤더	평가 포인트
Monitoring	임계값 기반 알람, 상태 추적	Nagios, Zabbix, Icinga	알려진 문제 탐지
Observability	메트릭·로그·트레이스 수집/분석	OpenTelemetry, Splunk	미지의 장애, 복합 추적
AIOps	ML/LLM 기반 이상 탐지, RCA 자동화	Dynatrace, Datadog, IBM	노이즈 감소, 자동 진단
Agentic AI	폐쇄 루프 자율화, 자동 조치	Datadog Bits AI, Azure SRE Agent	자율 운영, 인간 승인 게이팅

## 2.1.2 Agentic AI vs Causal AI — 폐쇄 루프 자율화 vs 인과 추론

Agentic AI와 Causal AI는 AI 기반 IT 운영에서 핵심적인 역할을 담당하는 두 가지 상이한 접근 방식입니다. Agentic AI는 운영 자동화의 깊이를, Causal AI는 진단의 정확도를 극대화하는 데 중점을 둡니다. 최근 IT 인프라의 복잡성이 증가하고, 장애의 원인 파악 및 조치 속도가 비즈니스 연속성에 직접적인 영향을 미치면서 두 기술의 결합이 점점 더 중요해지고 있습니다. 실제로 글로벌 벤더들은 Agentic AI와 Causal AI를 통합한 솔루션을 내놓고 있으며, 의사결정자는 이 두 개념의 차이와 상호 보완성을 명확히 이해해야만, 플랫폼 도입 시 실질적인 효과를 기대할 수 있습니다. 본 절에서는 Agentic AI와 Causal AI의 정의, 동작 방식, 그리고 실무 적용 시의 시사점을 구체적으로

살펴봅니다.

### Agentic AI의 정의와 동작 방식

Agentic AI는 단일 LLM 호출을 넘어, 도구(tool) 호출, 메모리 관리, 계획(Planning) 수립 등 일련의 단계를 거쳐 detect(탐지)→diagnose(진단)→remediate(조치)→verify(검증)까지의 폐쇄 루프를 자율적으로 수행하는 AI 아키텍처입니다. 대표적으로 Datadog Bits AI SRE, Dynatrace Davis CoPilot, Azure SRE Agent 등이 이 범주에 속합니다. Agentic AI는 워크플로 자동화, 승인 기반 자동 조치, 이벤트 기반 트리거 등 실제 운영 자동화에 초점을 둡니다.

Agentic AI의 핵심은 복잡한 운영 환경에서 인간의 개입 없이도 장애 탐지부터 조치, 검증까지 일련의 프로세스를 자동으로 실행한다는 점입니다. 예를 들어, 장애 발생 시 로그와 메트릭을 자동 분석하여 원인을 추론하고, 사전에 정의된 정책에 따라 조치 스크립트를 실행한 뒤, 조치 결과를 검증하여 필요 시 추가 액션을 수행합니다. 이 과정에서 승인 게이팅(인간의 승인 절차)이 내장되어 있어, 완전 자동화와 통제의 균형을 맞출 수 있습니다.

### Causal AI의 정의와 역할

Causal AI는 통계적 상관관계(correlation)가 아닌, 인과관계(causation)를 추론하는 AI 기술입니다. Observability 데이터에서 수집된 수많은 이벤트와 지표 중에서, 실제 장애의 근본 원인(Root Cause)을 정확히 식별하는 데 초점을 둡니다. 예를 들어, Dynatrace Davis Causal은 수백만 개의 메트릭과 이벤트에서 인과 그래프를 생성하여, 장애의 직접 원인과 파급 경로를 자동으로 추론합니다.

Causal AI는 단순히 이상 징후를 탐지하는 데 그치지 않고, 복잡하게 얽힌 시스템 내에서 원인과 결과의 관계를 명확히 파악함으로써, 불필요한 조치나 잘못된 진단을 최소화합니다. 이로 인해 장애 대응의 효율성과 정확성이 대폭 향상되며, 운영팀의 리소스 낭비를 줄일 수 있습니다. 또한, 인과 추론 기반의 분석 결과는 규제 준수, 사후 감사, 장애 보고서 작성 등 다양한 실무 영역에서 신뢰할 수 있는 근거 자료로 활용됩니다.

### 두 개념의 상호 보완성

Agentic AI와 Causal AI는 상호 보완적입니다. Causal AI가 진단(RCA)의 정확도를 높여주고, Agentic AI가 실행(조치)의 자율화 깊이를 더합니다. 두 기술이 결합되어야만 트러블슈팅의 즉각 종결(Time-to-Resolution 단축)이 실현됩니다. Gartner 2025 Hype Cycle에서 Agentic AI는 Peak of Inflated Expectations(기대 정점)에 위치하며, 향후 환멸기 진입 가능성도 내포합니다. 실무에서는 두 기술이 분리되어 평가되지 않고, 통합된 가치 사슬로 이해되어야 합니다.

실제 도입 사례를 보면, 대형 금융기관이나 글로벌 SaaS 기업들은 Causal AI로 장애의 근본 원인을 신속히 파악한 후, Agentic AI를 통해 자동 조치와 검증을 수행함으로써, 장애 복구 시간을 기존 대비 50% 이상 단축하는 성과를 내고 있습니다. 이러한 통합 운영 모델은 운영 효율성뿐 아니라, 서비스 신뢰도와 고객 만족도 향상에도 직접적으로 기여합니다.

### Agentic vs Causal 비교 표

구분	Agentic AI	Causal AI
목적	자율적 실행, 폐쇄 루프 자동화	근본 원인(RCA) 인과 추론
입력	워크플로, 도구 호출, 이벤트	메트릭, 로그, 트레이스, 이벤트
출력	자동 조치, 승인 기반 액션	인과 그래프, RCA 결과
한계	진단 정확도 한계	실행 자동화 한계

### 2.1.3 MCP·BYO LLM·Top-down·Bottom-up — 한국 도입의 4가지 결정 변수

AI 오피저버빌리티 플랫폼의 국내 도입에서 MCP, BYO LLM, Top-down, Bottom-up은 실질적인 도입 성공 여부를 좌우하는 4대 결정 변수입니다. 이 변수들은 단순히 기술적 옵션이 아니라, 글로벌 표준 준수, 데이터 주권, 규제 대응, 그리고 장애 대응 체계의 완성도와 직결됩니다. 특히 한국 금융·공공 시장에서는 RFP 및 Vendor Scorecard 평가에서 이 네 가지 요소의 충족 여부가 필수 조건으로 작용하며, 실제 벤더 선정과 프로젝트 성공률에 큰 영향을 미칩니다. 본 절에서는 각 변수의 정의와 실무적 의미, 그리고 한국 시장에서의 도입 사례와 평가 기준을 구체적으로 설명합니다.

#### MCP(Model Context Protocol)의 표준화와 의미

MCP는 Anthropic이 2024년 11월에 발표한 LLM 에이전트와 외부 도구(오피저버빌리티 포함) 간의 표준 인터페이스 프로토콜입니다. MCP가 표준화되면서, 오피저버빌리티 도구가 LLM Agent의 입력 채널로 공식 인식되고, 다양한 LLM·Agent 플랫폼(예: OpenAI Agents SDK, VS Code Agent Mode, Cursor)에서 MCP 호환 도구만을 채택하는 흐름이 가속화되고 있습니다. MCP 호환성은 2025~2026년 글로벌 RFP의 필수 조건으로 부상하고 있습니다.

MCP의 도입은 단일 벤더 종속성을 줄이고, 다양한 LLM 및 도구 간의 상호 운용성을 보장한다는 점에서 의사결정자에게 매우 중요한 의미를 가집니다. 예를 들어, MCP를 지원하는 벤더의

솔루션을 선택하면, 향후 LLM 교체나 신규 도구 연동 시 추가 개발 없이 손쉽게 통합이 가능합니다. 이는 장기적인 플랫폼 확장성과 투자 보호 측면에서 큰 이점을 제공합니다.

### BYO LLM(Bring Your Own LLM)의 구조와 필요성

BYO LLM은 벤더가 LLM을 고정 제공하지 않고, 고객이 OpenAI, AWS Bedrock, Google Vertex AI, 사내 sLLM(전용 LLM) 등 원하는 LLM을 직접 선택·연동하는 구조입니다. 이 방식은 데이터 주권, 비용 통제, 한국어 특화 LLM 결합, 규제 준수의 4가지 효과를 동시에 달성합니다. 한국 금융·공공 시장에서는 BYO LLM이 사실상 진입 표준이 되고 있습니다.

BYO LLM 구조를 도입하면, 민감한 데이터가 외부 클라우드로 유출되는 위험을 최소화할 수 있으며, 비용 측면에서도 벤더 종속적 과금 구조에서 벗어나 자율적으로 예산을 관리할 수 있습니다. 또한, 한국어 특화 LLM을 직접 연동함으로써, 국내 업무 환경에 최적화된 자연어 처리와 트러블슈팅 자동화가 가능합니다. 실제로 국내 대형 은행 및 공공기관의 RFP에서는 BYO LLM 지원 여부가 필수 평가 항목으로 명시되고 있습니다.

### Top-down·Bottom-up 모니터링의 통합

Top-down 모니터링은 비즈니스 KPI, 사용자 경험 등 상위 지표에서 출발하여 인프라, 애플리케이션, 네트워크로 진단 범위를 내려가는 방식입니다. 반대로 Bottom-up 모니터링은 메트릭, 로그, 트레이스 등 하위 계층에서 장애 신호를 감지하고, 이를 비즈니스 영향까지 추적하는 방식입니다. 트러블슈팅의 즉각 종결을 위해서는 두 방향의 단일 콘솔 통합이 필수입니다. 한국 시장에서는 이 통합 여부가 RFP 평가의 핵심 변수로 작용합니다.

실제 사례를 보면, 국내 대기업에서는 Top-down 방식으로 비즈니스 장애를 조기에 감지하고, Bottom-up 방식으로 장애의 근본 원인을 신속히 추적함으로써, 장애 대응 시간을 획기적으로 단축하고 있습니다. 두 방식이 단일 콘솔에서 통합적으로 지원될 때, 운영팀 간 협업 효율성과 장애 복구 속도가 극대화됩니다.

### 4가지 결정 변수 표

변수	정의/설명	한국 규제 환경 충족 필수성
MCP	LLM Agent↔도구 표준 프로토콜(Antropic 주도)	필수(RFP, Vendor Scorecard)
BYO LLM	LLM 고정 제공 X, 고객 LLM 직접 선택·연동	필수(데이터 주권, 비용, 한국어)
Top-down	비즈니스→인프라 하향식 진단	필수(비즈니스 KPI 연계)
Bottom-up	인프라→비즈니스 상향식 추적	필수(장애 발생→비즈니스 영향 추적)

## 2.2 혼동 쉬운 개념 쌍 — 의사결정 오류를 만드는 3가지 합성

AI 오피저빌리티 플랫폼 도입 과정에서 IT 의사결정자들이 자주 혼동하는 개념 쌍은 RFP 작성 및 Vendor Scorecard 평가에서 중대한 오류로 이어질 수 있습니다. 본 절에서는 실무에서 가장 빈번하게 혼동되는 3가지 개념 쌍을 짚어, 평가 오류를 예방하고 일관된 기준을 마련할 수 있도록 합니다. 각 개념 쌍은 실제 도입 사례, 벤더 마케팅 문구, 국내외 시장 환경에서의 해석 차이까지 고려하여 설명합니다.

### 2.2.1 AIOps vs Observability — “AIOps 도입했으니 Observability는 별도 안 사도 된다”의 오해

AIOps와 Observability는 IT 운영 자동화와 데이터 가시성 측면에서 밀접하게 연관되어 있지만, 실제로는 범위와 역할이 명확히 다릅니다. 많은 의사결정자들이 두 용어를 혼동하여, AIOps만 도입하면 전체 오피저빌리티 체계가 자동으로 구축된다고 오해하는 경우가 많습니다. 이러한 오해는 실제 도입 및 운영 과정에서 심각한 기능 누락과 장애 대응 실패로 이어질 수 있습니다. 본 절에서는 두 개념의 포함 관계와 실무 적용 시 주의해야 할 평가 포인트를 구체적으로 설명합니다.

#### AIOps의 분석 계층 한정성

AIOps는 머신러닝(ML) 또는 LLM 기반의 이상 탐지, 알람 노이즈 감소, 근본 원인 분석(RCA) 등 분석 계층의 자동화 기능을 의미합니다. 그러나 AIOps가 제공하는 것은 데이터 분석·자동화 기능에 한정되며, 데이터 수집·저장·시각화·사용자 인터페이스(UI) 등 전체 오피저빌리티 체계를 대체하지는 않습니다. 예를 들어, Dynatrace Davis, Datadog Watchdog, IBM Instana의 AI 엔진은 Observability 데이터 위에서 동작합니다.

AIOps는 기존 모니터링 시스템이 제공하지 못하는 이상 탐지, 패턴 분석, 자동화된 RCA 등 고도화된 기능을 제공하지만, 그 자체로는 데이터의 수집·저장·가공·시각화 기능이 내장되어 있지 않습니다. 즉, AIOps는 오피저빌리티 데이터의 분석 계층에 해당하며, 데이터 계층이 제대로 구축되어 있지 않으면 AIOps의 효과도 제한적일 수밖에 없습니다.

#### Observability의 데이터+분석+UI 전체 범위

Observability는 메트릭, 로그, 트레이스, 이벤트 등 데이터 플레인, 분석 계층(AIOps 포함), 그리고 대시보드, 알람, 쿼리, NL 인터페이스 등 UI까지 포함하는 전체 스택을 의미합니다. 즉, AIOps는 Observability의 하위 계층(분석 계층)에 해당하며, 둘은 보완재가 아니라 포함 관계입니다.

Observability 플랫폼은 데이터의 수집에서부터 저장, 가공, 시각화, 알람 설정, 쿼리, 그리고 자연어 인터페이스까지 전 과정을 포괄합니다. AIOps는 이 중 분석 계층의 자동화에 특화된 기능을 제공하며, 전체 옴저버빌리티 체계의 일부로 작동합니다. 따라서 두 용어를 동일시하거나 대체재로 오해할 경우, 실제 운영 환경에서 데이터의 가시성, 장애 탐지, 대응 속도 등 핵심 역량이 심각하게 저하될 수 있습니다.

### 실제 RFP 오류와 평가 위험

한국 SI 견적서나 RFP에서 “AIOps 솔루션”만 도입하면 전체 옴저버빌리티가 구축된다고 오해하는 사례가 많습니다. 이 경우 데이터 수집·저장·시각화 등 핵심 기능이 누락되어, 장애 대응 및 트러블슈팅의 즉각 종결이 불가능해집니다. 두 SKU를 동시에 평가할 때 중복·누락 위험을 반드시 체크해야 합니다.

실제로, AIOps와 Observability를 별도 SKU로 제공하는 벤더도 많으므로, RFP 작성 시 두 제품의 기능 범위와 포함 관계를 명확히 구분해야 합니다. 이를 통해 중복 구매나 기능 누락에 따른 예산 낭비, 장애 대응 실패 등의 리스크를 사전에 차단할 수 있습니다.

### AIOps vs Observability 포함 관계 다이어그램

- Observability(전체)
  - └─ 데이터 계층(메트릭/로그/트레이스)
  - └─ 분석 계층(AIOps)
  - └─ UI 계층(대시보드, NL 인터페이스 등)

## 2.2.2 LLM Observability vs Observability with LLM — 모니터링 대상이 LLM인가, 도구가 LLM인가

LLM Observability와 Observability with LLM은 LLM(대형언어모델)과 옴저버빌리티의 결합 방식에서 근본적으로 다른 접근을 의미합니다. 두 개념은 모니터링의 “대상”과 “도구”의 차이에서 출발하며, 실제 도입 및 운영 환경에서 요구되는 기능과 평가 기준이 완전히 다릅니다. 최근 LLM

기반 애플리케이션의 확산과 함께, 두 영역의 수요가 동시에 증가하고 있으므로, 의사결정자는 각 개념의 차이와 통합 가능성을 명확히 이해해야 합니다.

### LLM Observability의 대상과 기능

LLM Observability는 LangSmith, Langfuse, Arize 등과 같이 LLM 기반 애플리케이션 자체를 모니터링하는 도구를 의미합니다. 이 범주에서는 프롬프트, 응답, 토큰 사용량, 비용, 환각(هل루시네이션) 발생률, 프롬프트 인젝션 탐지, 평가 데이터셋 관리 등 LLM 앱 특화 지표를 수집·분석합니다. 즉, 모니터링 “대상”이 LLM입니다.

LLM Observability 도구는 LLM 애플리케이션의 품질 관리, 비용 최적화, 보안 위협 탐지 등 LLM 특유의 운영 이슈를 해결하는 데 초점을 둡니다. 예를 들어, 프롬프트 인젝션 공격이 발생했을 때 이를 신속히 탐지하고, 토큰 사용량이 급증하는 시점을 실시간으로 모니터링하여 비용 초과를 방지할 수 있습니다. 또한, 환각 발생률과 같은 품질 지표를 지속적으로 추적함으로써, LLM 앱의 신뢰성과 사용자 경험을 개선할 수 있습니다.

### Observability with LLM의 도구적 활용

Observability with LLM은 Datadog, Dynatrace, Splunk 등 기존 오피저빌리티 도구에 LLM 기반 어시스턴트(예: NL 쿼리, 자동 RCA, 알람 요약 등)가 내장된 형태를 의미합니다. 이 경우 LLM은 모니터링 “도구”로 활용되며, 기존 인프라·애플리케이션·네트워크·DB 등 전체 스택의 트러블슈팅을 지원합니다.

이 접근 방식에서는 LLM이 자연어 쿼리, 알람 요약, RCA 자동화 등 운영자의 생산성 향상과 트러블슈팅 효율화에 기여합니다. 예를 들어, 운영자가 자연어로 “지난 24시간 동안 CPU 사용률이 급증한 서버를 보여줘” 라고 입력하면, LLM이 복잡한 쿼리를 자동 생성하여 결과를 시각화해줍니다. 또한, 장애 알람이 발생했을 때 LLM이 주요 원인과 조치 내역을 요약해주어, 운영자의 판단 속도를 높일 수 있습니다.

### 한국 시장에서의 동시 수요와 평가 기준

2026~2027년 한국 시장에서는 LLM Observability와 Observability with LLM 두 카테고리 수요가 동시에 폭발적으로 증가할 것으로 예상됩니다. 사내 LLM 앱 운영자와 SRE팀이 같은 콘솔을 사용할 수 있는가, 혹은 별도 도구가 필요한가가 평가 기준이 됩니다. 8장에서는 국내 모니터링 벤더 6사의 AI 도입 현황을 출처 기반으로 실측하며, 두 카테고리의 통합 가능성을 9장 7개 체크리스트의 OTel GenAI 항목으로 평가합니다.

실제 도입 사례를 보면, 일부 대기업은 LLM Observability와 기존 인프라 오피저빌리티를 별도

도구로 운영하다가, 통합 콘솔 도입을 통해 운영 효율성과 데이터 일관성을 크게 개선하였습니다. 따라서 RFP 작성 시 두 영역의 통합 가능성, 데이터 연동 방식, 운영 주체의 역할 분담 등을 명확히 정의하는 것이 중요합니다.

**매트릭스 표**

구분	LLM Observability	Observability with LLM
모니터링 대상	LLM 앱 자체	인프라/앱/네트워크/DB 전체
주체(운영자)	ML/데이터팀, AI팀	SRE, 인프라팀, DevOps
주요 도구	LangSmith, Langfuse, Arize	Datadog, Dynatrace, Splunk
통합 가능성	별도 도구, 일부 연동	기존 콘솔에 LLM 내장

**2.2.3 Monitoring as Code vs Configuration as Code — 모니터링 정책의 Git 관리 영역 분리**

Monitoring as Code와 Configuration as Code는 IT 인프라와 모니터링 정책을 코드(Git)로 관리하는 현대적 운영 패러다임입니다. 두 개념은 관리 대상과 적용 범위, 그리고 실무에서의 역할이 명확히 구분되며, AI/Agentic 워크플로가 활성화되는 환경에서는 두 방식의 결합이 필수적입니다. 본 절에서는 각 개념의 정의, 대표 도구, 그리고 결합의 필요성을 구체적으로 설명합니다.

**Configuration as Code의 범위와 도구**

Configuration as Code는 Terraform, Pulumi 등 인프라 자원(서버, 네트워크, 스토리지 등)의 상태를 코드(Git)로 관리하는 패러다임입니다. IaC(Infrastructure as Code)라고도 하며, 인프라 배포, 변경, 롤백, 감사 추적성을 보장합니다.

Configuration as Code의 도입으로, 인프라 자원의 생성·변경·삭제 이력이 모두 코드로 남아 감사 추적성이 강화되고, 자동화된 배포와 롤백이 가능해집니다. 예를 들어, 신규 서버를 추가하거나 네트워크 구성을 변경할 때, 모든 변경 내역이 Git에 기록되어, 문제가 발생했을 때 신속하게 이전 상태로 복구할 수 있습니다. 또한, 코드 기반 관리로 인해 협업과 검증 프로세스가 체계화되어, 대규모 인프라 운영의 안정성과 일관성이 크게 향상됩니다.

**Monitoring as Code의 특성과 도구**

Monitoring as Code는 Datadog Terraform Provider, Grafana as Code 등 모니터링 정

책(알람, 대시보드, 룰 등)을 코드로 관리하는 방식입니다. 모니터링 정책의 변경 이력, 승인, 롤백, 감사 추적성을 확보할 수 있습니다. AI/Agentic 워크플로가 활성화된 환경에서는 자동 액션 정책의 변경 이력 관리가 필수입니다.

Monitoring as Code를 도입하면, 알람 설정, 대시보드 구성, 자동화 정책 등 모든 모니터링 정책이 코드로 관리되어, 변경 이력과 승인 절차가 명확히 남습니다. 예를 들어, 알람 임계값을 변경하거나 신규 대시보드를 추가할 때, 모든 변경 사항이 Git에 기록되어, 실수나 오작동 발생 시 신속하게 롤백할 수 있습니다. 이는 특히 Agentic AI가 자동 조치를 수행하는 환경에서, 정책 변경의 투명성과 책임 분리를 보장하는 데 필수적입니다.

### AI 시대 결합의 필요성

Agentic 워크플로(자동 조치, 승인 게이팅 등)가 실무에 도입되면, 정책 변경 이력의 감사 추적성(auditability)이 필수 평가 항목이 됩니다. Monitoring as Code와 Configuration as Code의 결합은 인프라와 모니터링 정책의 통합 거버넌스, 자동화, 규제 준수(예: 금융/공공 감사)까지 지원합니다.

실제 금융권이나 공공기관에서는 두 방식을 결합하여, 인프라 자원과 모니터링 정책을 일관된 코드 관리 체계로 통합함으로써, 규제 감사 대응과 운영 효율성을 동시에 달성하고 있습니다. 예를 들어, 장애 발생 시 인프라 변경과 모니터링 정책 변경 이력을 함께 추적함으로써, 근본 원인 분석과 사후 보고가 신속하고 정확하게 이루어집니다.

### 비교 표

구분	Configuration as Code	Monitoring as Code
관리 객체	인프라 자원(서버, 네트워크)	모니터링 정책(알람, 대시보드)
대표 도구	Terraform, Pulumi	Datadog Provider, Grafana
책임 부서	인프라팀, DevOps	SRE팀, 모니터링팀
감사 추적성	배포/변경/롤백	정책 변경/승인/롤백

## 2.3 의사결정 프레임 적용 — 7개 개념을 RFP·Vendor Scorecard·법무 검토에 어떻게 배치하는가

2.1과 2.2에서 정의한 7개 핵심 개념은 실제 사내 RFP, Vendor Scorecard, 법무 검토 등 실무 문서에 구체적으로 반영되어야만 평가의 일관성과 객관성이 확보됩니다. 본 절에서는 각 개념을 평가 항목으로 변환하는 방법과, 2026~2028년 한국 시장에서 필수적으로 요구되는 5축 의사결정 매트릭스를 제시합니다. 이 프레임은 3장 이후 글로벌 벤더 비교, 7장 한국 규제 분석, 9장 결론 체크리스트까지 일관되게 적용됩니다.

### 2.3.1 7개 개념 → RFP 평가 항목 매핑 — 30분 회의용 평가지

7개 핵심 개념을 실제 RFP(제안요청서) 평가 항목으로 변환하는 것은 객관적이고 일관된 벤더 평가를 위해 필수적입니다. 각 개념에 대해 벤더에 묻는 1문장 질문과 OK/조건부/NG 기준을 명확히 정의함으로써, 기술적 배경이 부족한 평가 위원도 동일한 기준으로 신속하게 채점할 수 있습니다. 본 절에서 제시하는 평가지 양식은 30분 내 회의에서 바로 활용할 수 있도록 설계되었으며, 실무에서 반복적으로 사용할 수 있는 표준화된 평가 도구로 기능합니다. 이를 통해 RFP, Vendor Scorecard, 법무 검토 등 다양한 실무 문서에 7개 개념을 일관되게 반영할 수 있습니다.

개념	벤더에 묻는 질문(1문장)	OK 기준	조건부 기준	NG 기준
Observability	메트릭·로그·트레이스 3축을 단일 콘솔에서 지원합니까?	3축+단일 콘솔	2축/다중 콘솔	1축/분리 콘솔
AIOps	ML/LLM 기반 RCA·이상탐지·노이즈 감소가 내장되어 있습니까?	ML+LLM+자동화	ML만/LLM만	없음
Agentic AI	탐지→진단→조치→검증 폐쇄 루프 자동화가 가능합니까?	4단계 자동+게이팅	2~3단계/수동 승인	미지원
Causal AI	인과 추론 기반 RCA가 내장되어 있습니까?	인과 추론+RCA	통계 기반/상관관계	없음
MCP	MCP 서버/클라이언트 호환성이 명시되어 있습니까?	공식 지원	3rd party 연동	미지원
BYO LLM	고객이 LLM을 직접 선택·연동할 수 있습니까?	OpenAI/Bedrock/사내 LLM	일부 제한	고정 LLM만
Top-down/Bottom-up	비즈니스→인프라/인프라→비즈니스 단일 흐름 지원합니까?	양방향+단일 콘솔	한 방향/다중 콘솔	미지원

## 2.3.2 BYO LLM·MCP·Self-hosted·Causal·Agentic 게이팅 — 5축 의사결정 매트릭스의 도입

2026~2028년 한국 시장에서 AI 오픈버빌리티 플랫폼을 도입할 때, 반드시 충족해야 할 5가지 결정 축이 존재합니다. 이 5축은 단순한 기술 옵션이 아니라, 규제 준수, 데이터 주권, 자동화 심도, 거버넌스 등 실무 요구사항을 모두 반영하는 핵심 의사결정 기준입니다. 각 축은 플랫폼 도입의 성공 여부를 좌우하며, 글로벌 벤더와의 비교, 규제 산업 내 도입 적합성, 그리고 장기적인 운영 안정성에 직접적인 영향을 미칩니다. 본 절에서는 각 축의 의사결정 함의와, 본 백서 내 심층 분석 장 인덱스를 함께 제공하여, 독자가 필요 시 신속하게 참조할 수 있도록 안내합니다.

축	의사결정 함의/필요성	본 백서 해당 장 인덱스
BYO LLM	데이터 주권, 비용 통제, 한국어 LLM 결합	3.1, 3.2, 3.3, 7.2
MCP	LLM Agent↔도구 표준 인터페이스, 생태계 연동	2.1.3, 3.4, 8.3
Self-hosted	망분리, 규제 준수, 온프레미스/폐쇄망 필수	3.2, 7.1
Causal AI	정확한 RCA, 장애 근본 원인 자동 식별	2.1.2, 3.4, 5.2
Agentic 게이팅	자동 조치의 승인·감사·롤백 거버넌스, 책임 분리	2.1.2, 3.4, 8.3

5축이 모두 충족되어야만 한국 금융·공공·대기업 등 규제산업에서 글로벌 AI 오픈버빌리티 플랫폼의 도입이 가능합니다. 각 축이 본 백서의 어느 장에서 심층적으로 다뤄지는지 인덱스를 함께 제공하여, 독자가 필요한 부분을 신속히 참조할 수 있도록 했습니다.

## 3장: 글로벌 11개 벤더 AI 기능 심층 비교 — Data-dog·Dynatrace·New Relic·Splunk·IBM Instana 등

### 3.1 11개 벤더 AI 기능 매트릭스 — 핵심 기능 / 기반 모델 / GA 시점 / 과금 4축 비교

글로벌 AI Observability 플랫폼 시장은 다양한 벤더들이 각기 다른 AI 기능과 전략으로 경쟁하고 있습니다. 이 장에서는 Datadog, Dynatrace, New Relic, Splunk, IBM Instana 등 상위 11개 벤더를 중심으로, AI 기능의 핵심 요소, 기반 모델, 출시(GA) 시점, 과금 방식 등 네 가지 주요 축에서 심층적으로 비교합니다. 특히 SaaS 중심 벤더와 엔터프라이즈/하이브리드, 그리고 ITSM/AIOps 통합 진영으로 구분하여, 각 벤더의 기술적 강점과 시장 적합성을 체계적으로 분석합니다. 이를 통해 IT 의사결정자들이 각 조직의 요구사항에 맞는 최적의 벤더를 선정할 수 있도록 실질적인 인사이트를 제공합니다.

#### 3.1.1 SaaS 5사 — Datadog Bits AI SRE·Dynatrace Davis·New Relic NRAI·Splunk AI Assistant·Cisco AppD

SaaS 벤더 5사는 AI Observability 기능의 구현 방식, 시장 출시 시점, 과금 정책 등에서 서로 다른 전략과 차별점을 보이고 있습니다. 이들은 주로 클라우드 기반 서비스로 빠른 도입과 확장성을 제공하지만, 각 벤더별로 AI 기능의 심화 정도, 기반 모델의 개방성, 그리고 과금 구조에서 뚜렷한 차이가 존재합니다. 특히 Datadog, Dynatrace, New Relic, Splunk, Cisco AppDynamics는 글로벌 시장에서 빠르게 AI 기능을 확장하고 있으며, 최신 LLM 및 자연어 인터페이스 도입에 적극적입니다. 그러나 이들 SaaS 벤더는 폐쇄망 환경이나 데이터 주권이 중요한 한국 금융·공공 시장에서는 도입에 제약이 있을 수 있습니다.

##### AI 기능 및 차별화 포인트

SaaS 진영의 5개 벤더는 AI 기능 구현 방식과 시장 출시 시점, 과금 정책에서 각기 다른 전략을 취하고 있습니다. Datadog은 Bits AI SRE를 별도 SKU로 출시하며, 트러블슈팅 자동화에 집중합니다. Dynatrace는 Davis Causal AI로 근본 원인 분석(RCA)에 강점을 보이며, CoPilot을 통해

Azure OpenAI 기반의 대화형 인터페이스를 제공합니다. New Relic은 NRAI로 OpenAI 기반의 자연어 질의 및 분석을 지원합니다. Splunk는 자체 호스팅 가능한 OSS Transformer 기반의 AI Assistant for SPL을 제공하며, Cisco AppDynamics는 단계적 GA 전략을 통해 시장 진입을 확대하고 있습니다.

Datadog은 Bits AI SRE를 통해 SRE(사이트 신뢰성 엔지니어링) 업무 자동화에 중점을 두고, 알림 분류, 인시던트 우선순위 지정, 자동화된 RCA 등 운영 효율성 향상에 집중합니다. Dynatrace는 Davis Causal AI와 CoPilot을 결합하여, 인과 추론 기반의 근본 원인 분석과 자연어 기반 대화형 인터페이스를 제공함으로써 운영자의 업무 부담을 줄여줍니다. New Relic은 NRAI를 통해 OpenAI 기반 자연어 질의와 로그/메트릭 분석을 지원하며, 비전문가도 쉽게 복잡한 데이터를 탐색할 수 있도록 돕습니다. Splunk는 AI Assistant for SPL을 통해 자연어 명령을 SPL(Splunk Processing Language) 쿼리로 자동 변환해주며, OSS Transformer 기반으로 확장성과 커스터마이징이 용이합니다. Cisco AppDynamics는 AI Assist를 단계적으로 도입하며, 기존 AppDynamics 플랫폼과의 통합성을 강화하고 있습니다.

### 기반 모델 및 GA 시점

Datadog Bits AI SRE는 자체 모델(비공개)로 2025년 12월 GA를 발표하였고, Dynatrace Davis CoPilot은 Azure OpenAI를 고정 모델로 2024년 10월 GA를 달성했습니다. New Relic NRAI는 OpenAI 기반으로 2025년 6월 GA, Splunk AI Assistant for SPL은 2024년 6월에 OSS Transformer 기반으로 GA를 달성했습니다. Cisco AppDynamics는 점진적 GA를 진행 중이며, 공식 발표가 순차적으로 이루어지고 있습니다.

이들 벤더의 기반 모델은 각기 다르며, Datadog은 내부적으로 개발한 비공개 LLM을 사용하여 데이터 보안과 최적화에 집중합니다. Dynatrace는 Azure OpenAI를 고정적으로 사용하여 안정성과 성능을 확보하였으나, 모델 교체의 유연성은 제한적입니다. New Relic은 OpenAI와의 연동을 통해 최신 LLM 기능을 빠르게 도입할 수 있으며, Splunk는 오픈소스 Transformer를 채택하여 커뮤니티와의 협업 및 기능 확장이 용이합니다. Cisco AppDynamics는 공식적으로 기반 모델을 공개하지 않았으나, 점진적 GA 전략을 통해 다양한 LLM 도입 가능성을 열어두고 있습니다.

### 과금 모델 및 한국 시장 적합성

SaaS 5사는 대부분 토큰 기반 또는 별도 SKU 과금 체계를 채택하고 있습니다. Datadog Bits AI SRE는 별도 SKU, Dynatrace는 Azure OpenAI 토큰 과금, New Relic은 OpenAI 토큰 및 사용량 기반, Splunk는 AI 기능을 별도 SKU로 분리, Cisco AppDynamics는 단계적 과금 모델을

적용합니다. 이들 SaaS 벤더는 폐쇄망(망분리) 환경에서의 도입이 제한적이므로, 한국 금융·공공 시장에서는 적합도가 낮은 편입니다.

특히, 토큰 기반 과금은 사용량에 따라 비용이 변동될 수 있어 예산 관리에 유의해야 하며, 별도 SKU는 AI 기능을 필요에 따라 선택적으로 도입할 수 있는 유연성을 제공합니다. 그러나 SaaS 특성상 데이터가 외부 클라우드로 전송될 수 있어, 데이터 주권과 망분리 요구가 강한 한국 금융·공공 부문에서는 도입에 한계가 있습니다.

**표: SaaS 5사 AI 기능 비교**

벤더	AI 기능명	기반 모델	GA 시점	과금 모델
Datadog	Bits AI SRE	자체(비공개)	2025-12-02	별도 SKU
Dynatrace	Davis CoPilot	Azure OpenAI 고정	2024-10-10	토큰/별도 SKU
New Relic	NRAI	OpenAI	2025-06-04	토큰/별도 SKU
Splunk	AI Assistant for SPL	OSS Transformer	2024-06-11	별도 SKU
Cisco AppD	AppDynamics AI Assist	[공식 발표 미확인]	단계적 GA	단계적 과금

(출처: [Datadog Bits AI SRE](#), [Dynatrace Davis CoPilot](#), [New Relic NRAI](#), [Splunk AI Assistant](#))

### 3.1.2 엔터프라이즈·하이브리드 4사 — IBM Instana(watsonx)·Elastic·Grafana·BMC HelixGPT

엔터프라이즈 및 하이브리드 벤더 4사는 온프레미스(Self-hosted) 배포와 BYO LLM(Bring Your Own LLM) 전략을 중심으로 차별화된 AI Observability 기능을 제공합니다. 이들은 대규모 조직, 금융·공공기관 등 데이터 주권과 규제 준수가 중요한 시장에서 강점을 보이고 있습니다. IBM Instana, Elastic, Grafana, BMC HelixGPT는 모두 고객이 직접 LLM을 선택·연동할 수 있도록 공식적으로 지원하며, 다양한 배포 옵션을 통해 망분리, 폐쇄망 환경에서도 유연하게 운영할 수 있습니다. 이러한 특성은 한국 시장의 규제 환경에 매우 적합하며, 현지화 및 비용 최적화 요구에도 효과적으로 대응할 수 있습니다.

#### Self-hosted 및 BYO LLM 전략

엔터프라이즈/하이브리드 4사는 Self-hosted(온프레미스) 배포와 BYO LLM(Bring Your

Own LLM) 전략을 전면에 내세웁니다. IBM Instana는 watsonx와 Granite, Llama, Mistral 등 다양한 LLM을 선택적으로 연동할 수 있습니다. Elastic AI Assistant for Observability, Grafana Sift + LLM 플러그인, BMC HelixGPT 모두 BYO LLM을 공식 지원하며, 고객이 직접 LLM을 선택하고 비용을 통제할 수 있습니다. 이로 인해 데이터 주권과 규제 준수, 비용 최적화가 가능해집니다.

IBM Instana는 watsonx 플랫폼을 기반으로 Granite, Llama, Mistral 등 다양한 LLM을 연동할 수 있어, 고객의 데이터 주권 요구와 다양한 AI 활용 시나리오에 유연하게 대응합니다. Elastic은 AI Assistant for Observability에서 공식적으로 BYO LLM을 지원하며, Elasticsearch 클러스터와의 연동을 통해 대규모 데이터 분석 환경에서도 높은 확장성과 성능을 보장합니다. Grafana는 LLM 플러그인과 Sift를 통해 다양한 LLM 연동 옵션을 제공하며, Air-gap 환경에서도 운영이 가능합니다. BMC HelixGPT는 금융·공공 특화 기능과 함께 BYO LLM, Self-hosted, Air-gap 운영을 모두 지원하여, 규제 준수와 데이터 보안이 중요한 조직에 적합합니다.

### 한국 규제 적합 시나리오

이들 4사는 Self-hosted 및 Air-gap 환경에서의 운영이 가능하므로, 한국 금융·공공 시장의 망분리, PIPA, 국정원 N²SF 등 규제 요건을 충족할 수 있습니다. 특히 BYO LLM은 한국어 sLLM(네이버, KT, SKT 등)과의 결합이 가능해, 현지화 요구에 유연하게 대응할 수 있습니다.

한국의 금융·공공기관은 데이터가 외부로 유출되는 것을 엄격히 제한하고 있으며, 자체 데이터 센터 내에서 모든 시스템을 운영해야 하는 경우가 많습니다. 이러한 환경에서 Self-hosted, Air-gap, BYO LLM 지원은 필수적입니다. 예를 들어, Elastic과 Grafana는 네이버 HyperCLOVA, KT KoGPT 등 한국어 특화 LLM과의 연동이 가능하며, 현지화된 자연어 질의 및 분석을 지원합니다. BMC HelixGPT는 금융·공공기관의 망분리 환경에 최적화되어 있으며, IBM Instana는 watsonx와의 결합을 통해 대규모 엔터프라이즈 환경에서도 안정적으로 운영할 수 있습니다.

**표: 엔터프라이즈 4사 AI 기능 비교**

벤더	Self-hosted 옵션	BYO LLM 명시도	한국 규제 적합 시나리오
IBM Instana	<input type="checkbox"/>	Granite/Llama/Mistral	watsonx 기반, 데이터 주권 확보
Elastic	<input type="checkbox"/>	공식 지원	한국어 sLLM 결합 가능
Grafana	<input type="checkbox"/>	공식 지원	LLM 플러그인, Air-gap 운영
BMC HelixGPT	<input type="checkbox"/>	공식 지원	금융/공공 망분리 환경 적합

(출처: [Elastic AI Assistant](#), [BMC HelixGPT](#), [IBM Instana](#), [Grafana LLM App](#))

### 3.1.3 ITSM·AIOps 통합 진영 2사 — ServiceNow Now Assist·LogicMonitor Edwin AI

ITSM(IT Service Management) 및 AIOps(IT 운영 자동화) 통합 진영의 대표 벤더인 ServiceNow와 LogicMonitor는 AI 기능의 통합과 자동화에 중점을 두고 있습니다. 이 두 벤더는 ITSM과 AIOps 플랫폼의 기존 강점을 기반으로 AI 기능을 확장하고 있으나, Self-hosted 및 BYO LLM 지원 측면에서는 상대적으로 제한적입니다. 본 절에서는 ServiceNow Now Assist for ITOM과 LogicMonitor Edwin AI의 AI 기능, 모델 고정도, 그리고 한국 시장에서의 적합도를 분석합니다.

#### AI 기능 및 모델 고정도

ServiceNow Now Assist for ITOM은 자체 Now LLM(고정)과 Azure OpenAI 옵션을 제공하며, LogicMonitor Edwin AI는 공식 발표된 AI 모델 정보가 제한적입니다. 두 벤더 모두 ITSM(IT Service Management)과 AIOps 통합에 중점을 두고 있으나, AI 기능의 유연성이나 BYO LLM 지원은 미흡한 편입니다.

ServiceNow Now Assist는 ITOM(IT Operations Management)과 긴밀하게 통합되어, 인시던트 자동 분류, 티켓 생성, 워크플로 자동화 등 다양한 운영 업무를 AI로 지원합니다. Now LLM은 ServiceNow가 자체 개발한 LLM으로, 데이터 보안과 성능 최적화에 집중되어 있으나, 외부 LLM 연동이나 모델 교체의 유연성은 부족합니다. Azure OpenAI 옵션을 통해 일부 기능 확장이 가능하지만, 여전히 BYO LLM과 같은 개방형 구조는 지원하지 않습니다.

LogicMonitor Edwin AI는 공식적으로 지원하는 LLM이나 AI 모델에 대한 정보가 제한적이며, 주로 인시던트 탐지, 이상 징후 분석, 자동화된 알림 기능 등 IT 인프라 운영의 효율화에 초점을 맞추고 있습니다. 그러나 BYO LLM이나 Self-hosted 옵션이 제공되지 않아, 데이터 주권이나 망분리 요구가 강한 환경에서는 도입에 한계가 있습니다.

#### 한국 시장 적합도

ServiceNow와 LogicMonitor는 Self-hosted 불가, BYO LLM 미지원, 폐쇄망 환경 미적합

으로 인해 한국 규제산업 진입이 어렵습니다. 특히 ServiceNow는 별도 SKU 및 토큰 과금 체계, 폐쇄망 미지원이 주요 제약입니다.

한국 금융·공공기관은 데이터가 외부로 유출되는 것을 엄격히 제한하고, 모든 시스템을 자체 데이터센터 내에서 운영해야 하는 경우가 많습니다. ServiceNow와 LogicMonitor는 SaaS 기반 서비스 특성상 이러한 요구를 충족시키기 어렵고, BYO LLM 미지원으로 인해 현지화된 LLM 연동이나 데이터 주권 확보가 불가능합니다. 이에 따라, 두 벤더 모두 한국 시장에서는 ‘하’ 적합도로 평가됩니다.

표: ITSM·AIOps 2사 AI 기능 비교

벤더	AI 기능	모델 고정도	한국 적합도
ServiceNow	Now Assist for ITOM	Now LLM 고정	하
LogicMonitor	Edwin AI	[공식 발표 미확인]	하

(출처: [ServiceNow ITOM](#))

### 3.2 데이터 주권·폐쇄망 적합도 매트릭스 — Self-hosted / Air-gap / BYO LLM 3축

글로벌 11개 벤더의 AI Observability 플랫폼은 데이터 주권, 폐쇄망(망분리) 환경 적합성, BYO LLM 지원 여부에서 현격한 차이를 보입니다. 이 절에서는 각 벤더의 Self-hosted, Air-gap, BYO LLM 지원 여부를 3단계(□/□/×)로 평가하고, 한국 시장에서의 적합도를 4단계(상/중-상/중/하)로 종합하여 제시합니다. 본 매트릭스는 망분리·국정원 N²SF·금융 AI 가이드라인 등 국내 규제 환경에서 즉시 활용 가능한 실무 평가지로 설계되었습니다.

#### 3.2.1 11개 벤더 × Self-hosted × Air-gap × BYO LLM 적합도 매트릭스

글로벌 AI Observability 벤더 11개사는 Self-hosted(온프레미스 설치), Air-gap(물리적 망분리 환경), BYO LLM(고객 LLM 직접 연동) 지원 여부에서 뚜렷한 차이를 보입니다. 본 매트릭스는 각 벤더별로 세 가지 축의 지원 여부를 종합적으로 평가하여, 한국 시장에서의 실질적 도입 가능성을 한눈에 파악할 수 있도록 구성되었습니다. 특히 데이터 주권과 망분리, BYO LLM 지원은 국내

금융·공공기관에서 필수적으로 요구되는 요소이므로, 각 항목의 지원 여부가 실제 도입에 결정적인 영향을 미칩니다.

### Self-hosted, Air-gap, BYO LLM 평가

11개 벤더별로 Self-hosted(온프레미스 설치), Air-gap(물리적 망분리 환경), BYO LLM(고객 LLM 직접 연동) 지원 여부를 아래와 같이 평가합니다.

벤더	Self-hosted	Air-gap	BYO LLM	한국 적합도
Splunk	☐	☐	☐	상
IBM Instana	☐	☐	☐	상
Elastic	☐	☐	☐	상
BMC HelixGPT	☐	☐	☐	상
Dynatrace	☐(Managed)	☐	×	중-상
Grafana	☐	☐	☐	상
New Relic	×	×	×	하
Datadog	×	×	×	하
ServiceNow	×	×	×	하
LogicMonitor	×	×	×	하
Cisco AppD	☐	☐	×	중

- ☐: 완전 지원, ☐: 일부 지원/제약, ×: 미지원

Splunk, IBM Instana, Elastic, BMC HelixGPT, Grafana 등은 Self-hosted, Air-gap, BYO LLM을 모두 완전하게 지원하여, 데이터 주권과 망분리 요구가 강한 환경에 적합합니다. Dynatrace는 Managed 환경에서만 일부 Self-hosted와 Air-gap을 지원하지만, BYO LLM은 미지원입니다. Cisco AppDynamics는 일부 하이브리드 배포를 지원하나, 완전한 폐쇄망 환경에는 미흡합니다. New Relic, Datadog, ServiceNow, LogicMonitor는 Self-hosted 및 BYO LLM 미지원으로, 한국 폐쇄망 환경에 적합하지 않습니다.

### 주요 각주 및 특이사항

- Dynatrace는 Managed 환경에서만 일부 Air-gap 지원, BYO LLM 미지원(고정 Azure OpenAI).

- Cisco AppDynamics는 일부 하이브리드 배포 지원이나, 완전한 폐쇄망 환경은 미흡.
- New Relic/Datadog/ServiceNow/LogicMonitor는 Self-hosted 및 BYO LLM 미지원, 한국 폐쇄망 적합도 '하'로 평가.

(출처: [Dynatrace Copilot Data Privacy](#), [New Relic FedRAMP](#))

### 3.2.2 한국 진입 가능 5강 — Splunk Enterprise / IBM Instana / Elastic / BMC HelixGPT / Dynatrace Managed의 정렬

한국 시장에서 데이터 주권, 망분리, BYO LLM 등 규제 요건을 충족할 수 있는 벤더는 제한적입니다. 본 절에서는 상기 매트릭스에서 '상' 또는 '중-상'으로 평가된 5개 벤더를 중심으로, 각 벤더의 강점과 제약, 그리고 실제 적합 시나리오를 정리합니다. 이를 통해 한국 금융·공공기관 등 규제산업에서 실질적으로 도입 가능한 후보군을 명확히 제시합니다.

#### 한국 시장 1차 후보군 선정

상기 매트릭스에서 '상' 또는 '중-상'으로 평가된 5개 벤더는 한국 금융·공공 시장 진입의 1차 후보군입니다.

벤더	강점	제약	적합 시나리오
Splunk	완전한 Self-hosted/Air-gap/BYO LLM 지원	NL 질의 한국어 정확도는 별도 검증 필요	금융/공공/국방 망분리 환경
IBM Instana	watsonx 기반 다양한 LLM 선택, 데이터 주권	초기 도입 비용 및 watsonx 연동 복잡성	대규모 엔터프라이즈, watsonx 활용 조직
Elastic	BYO LLM 공식 지원, 한국어 LLM 결합 용이	Elasticsearch 클러스터 운영 복잡성	자체 LLM 활용, 데이터 주권 요구 조직
BMC HelixGPT	BYO LLM, Air-gap, 금융/공공 특화	국내 레퍼런스 부족, 도입 컨설팅 필요	금융/공공, 거버넌스 중시 조직
Dynatrace(M)	Managed 환경 일부 Air-gap, RCA 강점	BYO LLM 미지원, 완전 폐쇄망 미흡	Managed 환경, 일부 망분리, RCA 자동화 중시 조직

- NL 질의: Natural Language Query, 자연어 질의
- watsonx: IBM AI 플랫폼, 다양한 LLM 연동 가능

Splunk는 완전한 Self-hosted, Air-gap, BYO LLM 지원으로 데이터 주권과 망분리 환경에 최적화되어 있습니다. IBM Instana는 watsonx와 다양한 LLM 연동, 데이터 주권 확보가 강점이지만, 초기 도입 비용과 연동 복잡성은 고려해야 합니다. Elastic은 BYO LLM 공식 지원과 한국어 LLM 결합이 용이하나, Elasticsearch 클러스터 운영의 복잡성이 있습니다. BMC HelixGPT는 금융·공공 특화 기능과 BYO LLM, Air-gap 지원이 강점이나, 국내 레퍼런스 부족과 도입 컨설팅이 필요합니다. Dynatrace Managed는 일부 Air-gap 지원과 RCA(근본 원인 분석) 자동화에 강점이 있으나, BYO LLM 미지원과 완전한 폐쇄망 지원 미흡이 제약입니다.

(출처: [Splunk AI Assistant](#), [BMC HelixGPT](#))

### 3.3 BYO LLM의 부상 — 2025~2026 글로벌 패러다임 전환

2025~2026년 글로벌 Observability 시장에서는 BYO LLM(Bring Your Own LLM) 전략이 새로운 표준으로 자리잡고 있습니다. BYO LLM은 고객이 직접 원하는 LLM을 선택·연동할 수 있는 구조로, 데이터 주권, 비용 통제, 현지화 등 다양한 요구에 효과적으로 대응할 수 있습니다. 본 절에서는 BYO LLM 채택 진영과 고정 LLM 진영의 구분, 그리고 BYO LLM이 한국 규제, 비용, sLLM 결합에서 가지는 핵심 효과를 구체적으로 분석합니다.

#### 3.3.1 BYO LLM 채택 진영(Elastic/BMC/Grafana/IBM) vs 고정 LLM 진영(ServiceNow/Dynatrace/Datadog)

글로벌 AI Observability 시장은 BYO LLM을 공식 지원하는 진영과 고정 LLM만 제공하는 진영으로 뚜렷하게 구분되고 있습니다. Elastic, BMC HelixGPT, Grafana, IBM Instana 등은 고객이 직접 OpenAI, Anthropic, Google Gemini, 네이버 HyperCLOVA, KT KoGPT 등 다양한 LLM을 선택·연동할 수 있도록 지원합니다. 반면, ServiceNow, Dynatrace, Datadog 등은 자체 LLM 또는 고정된 외부 LLM만을 제공하며, 모델 교체나 현지화가 불가능합니다. 이로 인해 데이터 주권, 비용 통제, 현지화 등에서 양 진영 간 차별성이 크게 나타납니다.

##### BYO LLM 채택 진영의 특징

Elastic, BMC HelixGPT, Grafana, IBM Instana는 모두 BYO LLM을 공식 지원합니다. 즉, 고객이 직접 OpenAI, Anthropic, Google Gemini, 네이버 HyperCLOVA, KT KoGPT 등 원하는 LLM을 선택·연동할 수 있습니다. 이로써 데이터 주권, 비용 통제, 현지화(한국어 sLLM)

결합이 가능해집니다.

BYO LLM 진영은 고객이 자체적으로 LLM을 선택·운영할 수 있기 때문에, 데이터가 외부로 무단 전송되는 리스크를 최소화할 수 있습니다. 또한, 토큰 비용을 직접 관리할 수 있어 벤더 마진이 분리되고, 총소유비용(TCO) 투명성이 높아집니다. 현지화 측면에서도 네이버 HyperCLOVA, KT KoGPT, SKT A.L.L 등 한국어 특화 LLM과의 연동이 가능하며, 국내 시장의 요구에 유연하게 대응할 수 있습니다.

### 고정 LLM 진영의 특징과 한국 도입 장벽

ServiceNow Now Assist는 자체 Now LLM 고정, Dynatrace Davis CoPilot은 Azure OpenAI 고정, Datadog Bits AI SRE는 모델 비공개로 운영됩니다. 이들 벤더는 LLM 교체나 현지화가 불가하며, 고객 데이터가 외부 LLM에 종속적으로 전송될 수 있어 한국 규제산업 도입에 구조적 한계가 있습니다.

고정 LLM 진영은 벤더가 지정한 LLM만 사용할 수 있어, 데이터 주권 확보가 어렵고, 외부 LLM에 데이터가 전송될 수 있는 구조적 한계가 있습니다. 또한, 한국어 등 현지화 요구에 대응하기 어렵고, 비용 통제 측면에서도 벤더의 과금 정책에 종속될 수 있습니다. 이러한 한계로 인해 한국 금융·공공기관 등 규제산업에서는 도입에 제약이 많습니다.

### 표: BYO LLM vs 고정 LLM 진영 비교

진영	벤더	한국 도입 장벽 요약
BYO LLM	Elastic, BMC, Grafana, IBM	데이터 주권/현지화/비용 통제 강점
고정 LLM	ServiceNow, Dynatrace, Datadog	LLM 교체 불가, 외부 전송, 규제 부적합

(출처: [Dynatrace Copilot Data Privacy](#))

## 3.3.2 BYO LLM이 한국 규제·비용·sLLM 결합에 동시 작동하는 3가지 이유

BYO LLM 구조는 한국 시장에서 규제 거버넌스, 비용 통제, 현지화(한국어 sLLM 결합) 등 세 가지 핵심 요구를 동시에 충족시킬 수 있습니다. 각 효과별로 구체적인 작동 원리와 KPI(핵심 성과 지표) 예시를 아래와 같이 정리합니다.

### (a) 규제 거버넌스 통과

BYO LLM은 PIPA, 국정원 N²SF, 금융 AI 가이드라인 등 한국의 엄격한 데이터 거버넌스 요

건을 충족할 수 있습니다. 고객이 직접 LLM을 선택·운영함으로써 데이터가 외부로 무단 전송되는 리스크를 최소화합니다.

예를 들어, 금융기관이 자체 데이터센터 내에서 LLM을 운영할 경우, 고객 데이터가 외부 클라우드로 전송되지 않아 개인정보보호법(PIPA) 및 국정원 N²SF 등 국내 규제 요건을 충족할 수 있습니다. 이는 실제 도입 시 규제 심사 기간 단축 및 승인 가능성 제고로 이어집니다.

**(b) 비용 통제 및 벤더 마진 분리**

LLM 토큰 비용을 고객이 직접 부담하게 되어, 벤더 마진이 분리되고 총소유비용(TCO) 투명성이 높아집니다. 이는 대규모 데이터 처리·AI 활용 조직에서 예산 계획 및 비용 최적화에 결정적입니다.

예를 들어, Elastic이나 Grafana를 도입한 조직은 LLM 토큰 비용을 자체적으로 관리할 수 있어, 사용량에 따라 예산을 세밀하게 조정할 수 있습니다. 벤더의 과금 정책에 종속되지 않고, 실제 사용량 기반으로 비용을 최적화할 수 있는 구조입니다.

**(c) 한국어 sLLM 결합 가능성**

BYO LLM 구조는 네이버 HyperCLOVA, KT KoGPT, SKT A.L.L 등 한국어 특화 LLM과의 결합이 용이합니다. 이를 통해 현지화된 자연어 질의, 한국어 RCA, 규제 준수 등 현장 요구에 맞는 AI Observability를 구현할 수 있습니다.

실제로 Elastic, Grafana 등은 한국어 sLLM을 연동하여, 자연어 질의의 한국어 정확도를 높이고, 현지 규제에 부합하는 AI 기능을 제공할 수 있습니다. 이는 국내 시장에서의 경쟁력 확보에 중요한 역할을 합니다.

**표: BYO LLM 3가지 효과 × KPI 매핑**

효과	KPI 예시
규제 거버넌스	거버넌스 통과 시간(단축)
비용 통제	총소유비용(TCO) 절감율
한국어 sLLM	NL 질의 한국어 정확도(향상)

(출처: [Elastic AI Assistant](#), [IBM watsonx](#))

## 3.4 NL → Causal RCA → Agentic 워크플로의 단일 흐름을 잇는 글로벌 3사 — Dynatrace·Datadog·IBM Instana

글로벌 AI Observability 시장에서 트러블슈팅 즉각 종결의 단일 흐름(NL→쿼리→Causal RCA→Agentic 워크플로→외부 LLM 연동)을 단일 콘솔에서 매끄럽게 제공하는 벤더는 극히 제한적입니다. 본 절에서는 이 4계층을 모두 지원하는 Dynatrace, Datadog, IBM Instana 3사를 중심으로, 각 계층별 기능과 한국 시장 적합성의 충돌을 분석합니다.

### 3.4.1 단일 흐름 4계층 — NL→쿼리·Causal RCA·Agentic 워크플로·MCP 외부 LLM의 직선 결합

AI Observability의 진화는 단일 콘솔 내에서 자연어 질의(NL→쿼리), 인과 추론 기반 RCA, Agentic 워크플로, 그리고 외부 LLM 연동(MCP)까지의 일관된 워크플로를 지원하는 방향으로 발전하고 있습니다. 본 절에서는 이 4계층을 모두 지원하는 글로벌 벤더들의 기술적 특징과, 각 계층별로 실제 현장에서 어떻게 활용되는지 구체적으로 설명합니다.

#### (1) NL→쿼리(Top-down 진입)

Dynatrace CoPilot, Splunk AI Assistant, New Relic NRAI, Elastic ES|QL Assistant 등은 자연어 질의를 통해 메트릭·로그·트레이스 조회를 자동화합니다. 이 계층은 SRE/운영자가 복잡한 쿼리 언어를 몰라도 트러블슈팅을 시작할 수 있게 해줍니다.

예를 들어, 운영자가 “지난 24시간 내 CPU 사용률 급증 원인 분석” 과 같은 자연어 명령을 입력하면, 시스템이 자동으로 적합한 쿼리를 생성하여 로그와 메트릭을 조회하고, 결과를 시각화 해줍니다. 이를 통해 비전문가도 복잡한 데이터 분석을 손쉽게 수행할 수 있습니다.

#### (2) Causal RCA(Top-down→Bottom-up 연결)

Dynatrace Davis Causal, IBM Instana Concert, Datadog Watchdog 등은 통계적 상관 이 아닌 인과 추론 기반의 근본 원인 분석(RCA)을 제공합니다. 이 계층은 알람 폭증·RCA 가설 인간 의존 문제를 시로 해소합니다.

Causal RCA는 단순히 이벤트 간 상관관계를 분석하는 것이 아니라, 실제로 어떤 이벤트가 문제의 근본 원인인지를 인과 추론을 통해 자동으로 식별합니다. 예를 들어, 특정 서비스 장애가 네트워크 지연, 데이터베이스 병목, 애플리케이션 코드 오류 중 무엇에 기인하는지 시가 자동으로

분석하여, 운영자의 수작업 RCA 부담을 크게 줄여줍니다.

**(3) Agentic 워크플로(승인 후 자동 액션)**

Datadog Bits AI SRE, Dynatrace Davis 워크플로, Azure SRE Agent, AWS DevOps Agent 등은 탐지→진단→조치→검증의 폐쇄 루프를 Agentic 워크플로로 자동화합니다. 이 계층은 운영자의 승인(게이팅) 후 자동 Remediation을 실행합니다.

Agentic 워크플로는 인시던트 발생 시 AI가 자동으로 원인 분석, 해결 방안 제시, 그리고 운영자의 승인 후 자동 조치까지 일련의 과정을 자동화합니다. 예를 들어, 특정 서버의 과부하가 감지되면, AI가 원인을 분석하고, 리소스 증설이나 서비스 재시작 등 적절한 조치를 제안하며, 운영자가 승인하면 자동으로 실행됩니다. 이후 조치 결과를 검증하여, 문제 해결 여부를 보고합니다.

**(4) MCP 서버(외부 LLM 결합)**

Grafana, Sentry, Cloudflare, Honeycomb, New Relic, MongoDB 등은 MCP(Model Context Protocol) 서버를 통해 외부 LLM과의 표준 연동을 지원합니다. 이를 통해 다양한 LLM·에이전트와 Observability 데이터를 통합할 수 있습니다.

MCP 서버는 Observability 플랫폼과 외부 LLM 또는 AI 에이전트 간의 데이터 교환을 표준화하여, 다양한 AI 기능을 손쉽게 확장할 수 있도록 합니다. 예를 들어, IBM Instana는 watsonx와 MCP 연동을 통해, 외부 LLM을 활용한 고급 자연어 질의, 자동화된 RCA, 맞춤형 워크플로 등을 지원합니다.

**표: 4계층 × 벤더 매핑 × 한국 적합도**

계층	Dynatrace	Datadog	IBM Instana	한국 적합도
NL→쿼리	CoPilot	Bits AI SRE	watsonx NL	IBM Instana 상
Causal RCA	Davis Causal	Watchdog	Concert	IBM Instana 상
Agentic 워크플로	Davis Workflow	Bits AI SRE	Instana Actions	IBM Instana 상
MCP 외부 LLM	일부(미흡)	미지원	watsonx, MCP	IBM Instana 상

(출처: [Datadog Bits AI SRE](#), [Dynatrace Davis AI](#), [AWS CloudWatch Investigations](#))

### 3.4.2 Dynatrace·Datadog·IBM Instana 3사 평가 — 한국 적합도와의 충돌 진단

글로벌 AI Observability 시장에서 NL→쿼리, Causal RCA, Agentic 워크플로를 단일 콘솔에서 매끄럽게 결합하는 대표적 벤더는 Dynatrace, Datadog, IBM Instana입니다. 그러나 이들 벤더의 글로벌 기술 우수성이 반드시 한국 시장 적합도로 이어지지는 않습니다. 본 절에서는 각 벤더의 기술적 강점과 함께, 한국 시장에서의 도입 제약 및 적합도 충돌을 구체적으로 진단합니다.

#### 글로벌 우수 ≠ 한국 적합

Dynatrace, Datadog, IBM Instana는 NL→쿼리, Causal RCA, Agentic 워크플로를 단일 콘솔에서 매끄럽게 결합하는 대표적 3사입니다. 그러나 이 중 한국 시장에서 ‘상’ 적합도를 갖는 곳은 IBM Instana뿐입니다. Dynatrace는 BYO LLM 미지원(고정 Azure OpenAI), Datadog은 Self-hosted 미지원으로 폐쇄망 환경에 부적합합니다. IBM Instana는 watsonx 기반 BYO LLM, Self-hosted, MCP 연동을 모두 지원하여 한국 규제산업에 최적화된 유일한 글로벌 후보입니다.

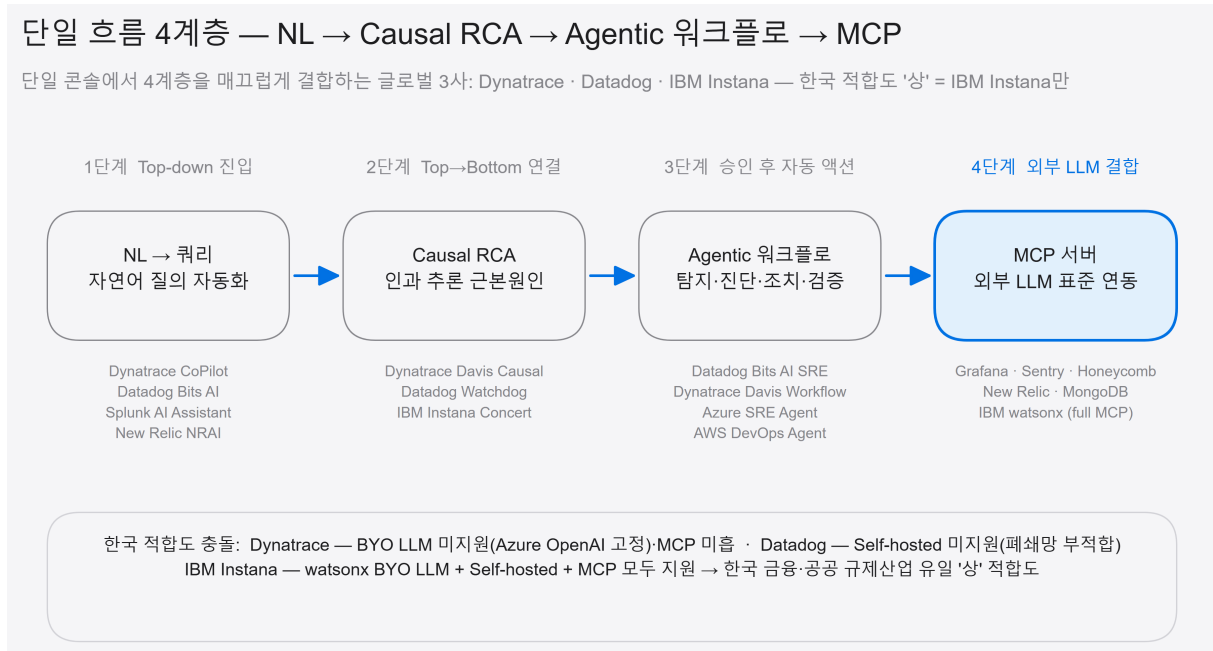
Dynatrace는 글로벌 시장에서 RCA 자동화와 Agentic 워크플로, 자연어 질의 등에서 높은 평가를 받고 있으나, Azure OpenAI 고정 모델과 BYO LLM 미지원으로 인해 데이터 주권과 현지화 요구가 강한 한국 시장에서는 한계가 있습니다. Datadog은 Bits AI SRE를 통해 트러블슈팅 자동화와 Agentic 워크플로를 지원하지만, Self-hosted 미지원으로 망분리 환경에 적합하지 않습니다. 반면, IBM Instana는 watsonx 기반 다양한 LLM 연동, Self-hosted, MCP 연동 등 모든 계층에서 완전한 지원을 제공하여, 한국 금융·공공기관 등 규제산업에 최적화된 유일한 글로벌 벤더로 평가받고 있습니다.

표: 3사 × 4계층 × 한국 적합도 충돌

벤더	NL→쿼리	Causal RCA	Agentic 워크플로	MCP 외부 LLM	한국 적합도
Dynatrace	O	O	O	X	중-상
Datadog	O	O	O	X	하
IBM Instana	O	O	O	O	상

- O: 지원, X: 미지원

(출처: [Dynatrace CoPilot](#), [IBM Instana](#))



[그림 2] 단일 흐름 4계층 — NL → Causal RCA → Agentic 워크플로 → MCP | 764

## 4장: 5대 CSP 자체 통합 관제 AI 전략 — AWS/Azure/GCP/OCI/Alibaba

### 4.1 5대 CSP AI 운영 에이전트 비교 — DevOps Agent·SRE Agent·Gemini Cloud Assist·OCI GenAI·ARMS Copilot

5대 글로벌 클라우드 서비스 제공자(CSP)는 각자의 인프라 환경에 최적화된 AI 운영 에이전트와 통합 관제 기능을 빠르게 발전시키고 있습니다. AWS, Azure, GCP는 AI 기반의 운영 자동화와 문제 해결을 위한 에이전트(DevOps Agent, SRE Agent, Cloud Assist)를 앞세워, 클라우드 네이티브 환경에서의 트러블슈팅과 운영 효율성을 극대화하고 있습니다. OCI(Oracle Cloud Infrastructure)는 DB 및 Exadata 워크로드에 특화된 AI 기능을 제공하며, Alibaba Cloud는 중국 및 동남아 시장에 맞춘 AI 운영 에이전트 전략을 전개하고 있습니다. 이 절에서는 각 CSP의 AI 운영 에이전트의 기능, 출시(GA) 현황, 과금 모델, 그리고 한국 리전에서의 가용성까지 실질적인 비교를 통해 의사결정자에게 클라우드별 AI 자율화의 깊이를 명확히 제공합니다.

## 4.1.1 AWS DevOps Agent·Azure SRE Agent·GCP Gemini Cloud Assist의 GA 현황과 agent-second 과금

AWS, Azure, GCP의 AI 운영 에이전트는 최근 몇 년간 클라우드 운영 자동화와 자율화의 핵심 도구로 부상하고 있습니다. 각 CSP는 자사 인프라에 최적화된 에이전트와 과금 모델을 도입하여, 기존의 VM, 컨테이너, API 호출 단위 과금에서 벗어나 실제 운영 효율성과 비용 예측 가능성을 높이려 노력하고 있습니다. 본 절에서는 각 에이전트의 출시(GA) 현황, 과금 구조, 그리고 한국 리전에서의 가용성까지 상세히 비교하여, 클라우드 운영 전략 수립에 필요한 주요 정보를 제공합니다.

### 주요 AI 운영 에이전트 특징

AWS, Azure, GCP는 각각 DevOps Agent, SRE Agent, Gemini Cloud Assist라는 AI 기반 운영 에이전트를 내세우며, 클라우드 인프라의 자동화와 자율 운영을 지향합니다. AWS DevOps Agent는 2026년 3월 GA(General Availability)로 출시되며, agent-second라는 새로운 과금 단위를 도입했습니다. Azure SRE Agent 역시 GA 상태로, AAU(Azure Automation Unit) 고정 요금과 LLM 토큰 사용량에 따른 과금이 결합된 모델을 제공합니다. GCP의 Gemini Cloud Assist는 2025년 기준 Preview 및 부분 GA 상태로, 점진적으로 기능을 확장 중입니다.

### agent-second 과금 모델의 이해와 시뮬레이션

agent-second 과금은 기존의 VM, 컨테이너, API 호출 단위와 달리, 에이전트가 실제로 동작한 시간(초 단위)에 따라 비용이 산정되는 방식입니다. 예를 들어, AWS DevOps Agent의 경우 \$0.0083/agent-sec로 책정되어, 한 달(30일) 동안 1개의 에이전트가 24시간 가동 시 약 \$21.5의 비용이 발생합니다. Azure SRE Agent의 경우, AAU 고정 요금 외에도 LLM 기반 분석에 사용된 토큰 수에 따라 추가 과금이 발생합니다. GCP Gemini Cloud Assist는 현재까지는 과금 구조가 명확히 공개되지 않았으나, 향후 LLM 활용량 기반 과금이 도입될 가능성이 높습니다.

agent-second 과금 모델은 실제 운영 환경에서의 유연성을 제공하지만, 동시에 사용량 예측이 어려워 월간 비용이 크게 달라질 수 있습니다. 예를 들어, 대규모 이벤트 발생 시 에이전트의 동작 시간이 급증할 수 있으므로, 운영자는 사전에 시뮬레이션을 통해 예상 비용을 산정해야 합니다. Azure의 경우, AAU 고정 요금과 LLM 토큰 과금이 결합되어 있어, 대규모 LLM 분석이 필요한 환경에서는 추가 비용이 발생할 수 있습니다. GCP는 Preview 단계에서 과금 체계가 명확하지 않으나, LLM 기반 기능이 확대될수록 토큰 과금 모델이 도입될 가능성이 높아집니다.

## GA 시점 및 한국 리전 가용성

AWS DevOps Agent는 2026년 3월 공식 GA, CloudWatch Investigations는 2025년 6월 GA로 발표되었습니다. Azure SRE Agent는 이미 GA 상태이며, Korea Central 리전에서 AIOps 기능을 지원합니다. GCP Gemini Cloud Assist는 asia-northeast3(서울) 리전에서 일부 기능만 Preview로 제공되고 있습니다. 각 CSP의 AI 운영 에이전트는 한국 리전에서의 실제 가용성이 도입 의사결정에 중요한 변수로 작용합니다.

GA(General Availability) 시점은 각 에이전트의 성숙도와 실제 서비스 적용 가능성을 판단하는 핵심 지표입니다. AWS와 Azure는 이미 공식적으로 GA를 선언한 반면, GCP는 Preview 및 부분 GA 상태로, 기능의 완성도와 지원 범위가 제한적입니다. 한국 리전의 경우, AWS와 Azure는 일부 기능을 공식 지원하지만, GCP는 제한적으로만 제공되고 있어, 도입 전 최신 지원 현황을 반드시 확인해야 합니다.

## AI 운영 에이전트 비교 매트릭스

CSP	AI 에이전트	GA 시점	과금 모델	한국 리전 가용성
AWS	DevOps Agent	2026-03	\$0.0083/agent-sec	일부 기능만 지원
AWS	CloudWatch Investigations	2025-06	미공개	미공지
Azure	SRE Agent	GA	AAU 고정 + 토큰 과금	Korea Central 지원
GCP	Gemini Cloud Assist	Preview/GA	미공개(LLM 과금 예상)	asia-northeast3 일부

(출처: [AWS DevOps Agent](#), [Azure SRE Agent](#), [GCP Gemini Cloud Assist](#))

## 의사결정 시 유의사항

agent-second, AAU, LLM 토큰 등 새로운 과금 단위는 월간 예상 비용을 반드시 시뮬레이션해야 하며, 한국 리전에서의 실제 가용성은 도입 전 반드시 최신 공지를 확인해야 합니다. 각 CSP의 AI 운영 에이전트는 단일 클라우드 환경에서는 높은 효율성을 제공하지만, 멀티클라우드 환경에서는 독립 벤더와의 비교 검토가 필수입니다. 또한, 각 CSP의 에이전트는 자사 생태계에 최적화되어 있으므로, 멀티클라우드 또는 하이브리드 환경에서는 독립 옴저버빌리티 벤더와의 연동 및 기능 심화 여부를 반드시 검토해야 합니다.

## 4.1.2 OCI GenAI/Agent Hub·Alibaba ARMS Copilot — Oracle DB 워크로드 강점과 동남아 한정성

OCI와 Alibaba Cloud는 AWS, Azure, GCP와는 차별화된 AI 운영 에이전트 전략을 전개하고 있습니다. 특히 OCI는 데이터베이스 중심의 워크로드에 특화된 AI 자동화 기능을 제공하며, Alibaba Cloud는 중국 및 동남아 시장에 집중한 운영 자동화와 LLM 기반 트러블슈팅 기능을 강화하고 있습니다. 본 절에서는 두 CSP의 AI 기능적 강점과 한계, 그리고 한국 리전에서 도입 가능성을 상세히 분석하여, 국내외 조직의 클라우드 선택에 실질적인 인사이트를 제공합니다.

### Oracle Cloud Infrastructure(OCI)의 AI 전략

OCI는 GenAI와 Agent Hub, 그리고 Ops Insights 2025년 강화 기능을 통해, DB와 Exadata 워크로드에 특화된 AI 기반 운영 자동화를 제공합니다. 특히 Oracle Autonomous DB, EM 24ai의 Ask EM, Select AI 등은 대규모 데이터베이스 환경에서의 트러블슈팅과 최적화에 강점을 보입니다. 한국 리전(Seoul) 운영이 가능하다는 점도, 국내 금융·공공 등 데이터 주권이 중요한 조직에 매력적인 요소입니다.

Oracle의 AI 운영 에이전트는 대규모 트랜잭션 처리, 복잡한 쿼리 최적화, 장애 예측 및 자동 복구 등 데이터베이스 중심의 미션 크리티컬 환경에서 높은 신뢰성을 제공합니다. 예를 들어, Ops Insights는 실시간 성능 모니터링과 AI 기반 이상 탐지, 자동 튜닝 기능을 결합하여, 데이터베이스 관리자의 업무 부담을 크게 줄여줍니다. 또한, Agent Hub는 다양한 AI 운영 에이전트를 통합 관리할 수 있는 플랫폼으로, 멀티에이전트 환경에서의 확장성과 유연성을 보장합니다. 이러한 기능은 특히 금융, 공공, 제조 등 데이터 주권과 안정성이 중요한 산업군에서 높은 평가를 받고 있습니다.

### Alibaba Cloud의 ARMS Copilot 및 MCP Server

Alibaba Cloud는 ARMS Copilot과 MCP Server를 통해 AI 기반 운영 자동화와 LLM 기반 트러블슈팅 기능을 제공하고 있습니다. 그러나 이들 기능은 주로 중국 및 동남아 시장을 타겟으로 하며, 한국 내 공식 도입 사례는 제한적입니다. ARMS Copilot은 중국 내 대규모 인터넷 서비스, 동남아 진출 기업에 최적화되어 있으나, 한국 규제 환경에서는 실질적 도입 장벽이 존재합니다.

Alibaba ARMS Copilot은 대규모 분산 시스템의 장애 감지, 자동 복구, 실시간 성능 최적화 등 다양한 AI 기반 운영 기능을 제공합니다. MCP Server는 멀티클라우드 환경에서의 자원 통합 관리와 LLM 기반 인시던트 분석을 지원하지만, 중국 및 동남아 시장에 특화된 규제 및 네트워크

환경에 맞춰 설계되어 있습니다. 한국 내에서는 데이터 주권, 보안 규제, 현지화 지원의 한계로 인해 실제 도입이 제한적입니다. 따라서 Alibaba Cloud의 AI 운영 에이전트는 주로 중국, 동남아 진출 기업이나 현지 서비스 운영에 적합하며, 국내 시장에서는 우선순위가 낮은 편입니다.

### OCI·Alibaba AI 기능 및 한국 도입 가능성 매트릭스

CSP	AI 기능	한국 도입 가능성
Oracle Cloud	OCI GenAI + Agent Hub + Ops Insights	매우 높음 (리전 보유)
Alibaba Cloud	ARMS Copilot, MCP Server	낮음 (중국·동남아 한정)

### 의사결정 시 유의사항

Oracle Cloud는 DB 중심 워크로드가 많은 조직에서 글로벌 5강에 빠지지 않는 5번째 후보로 간주할 수 있습니다. 반면 Alibaba Cloud는 한국 내 실질적 도입 가능성이 제한적이므로, 동남아 진출 기업이나 중국 내 서비스 운영 목적이 아니라면 우선순위를 낮추는 것이 합리적입니다. 또한, OCI의 경우 한국 리전에서의 공식 지원과 데이터 주권 준수, 엔터프라이즈급 SLA 제공 등 국내 대기업 및 공공기관에 적합한 조건을 갖추고 있습니다. 반면 Alibaba Cloud는 현지화 및 규제 대응 측면에서 한계가 있으므로, 글로벌 확장 또는 특정 지역 특화 전략이 필요한 기업에만 적합합니다.

## 4.2 Microsoft Azure Arc — 멀티클라우드 단일 콘솔의 가장 공격적인 카드

Microsoft Azure Arc는 5대 CSP 중 유일하게 AWS, GCP 등 타 CSP의 자원까지 단일 콘솔에서 통합 모니터링할 수 있는 기능을 공식적으로 지원합니다. Azure Arc와 SRE Agent, 그리고 GitHub Copilot for Azure의 연계는 멀티클라우드 환경에서 운영 복잡성을 획기적으로 줄여줍니다. 그러나 분산 추적, 코드 레벨 RCA(근본 원인 분석) 등 심층 오피저빌리티 영역에서는 독립 벤더 대비 한계가 존재합니다. 이 절에서는 Azure Arc의 멀티클라우드 통합 비전과 그 한계, 그리고 독립 오피저빌리티 벤더와의 비교를 통해 멀티클라우드 의사결정자에게 실질적 선택 좌표를 제공합니다.

## 4.2.1 Azure Arc·SRE Agent·AIOps Observability Agent의 멀티클라우드 단일 콘솔 비전

Azure Arc는 멀티클라우드 및 온프레미스 환경에서의 자원 통합 관리와 운영 자동화의 새로운 패러다임을 제시하고 있습니다. 특히 AWS, GCP, 온프레미스 자원까지 Azure Portal에서 단일 뷰로 모니터링하고, 정책 적용과 보안 통제를 일원화할 수 있는 점은 멀티클라우드 복잡성 해소에 큰 장점입니다. 본 절에서는 Azure Arc의 멀티클라우드 통합 전략, 실제 적용 범위, 그리고 SRE Agent 및 GitHub Copilot for Azure와의 연계 효과를 구체적으로 살펴봅니다.

### Azure Arc의 멀티클라우드 통합 전략

Azure Arc는 AWS, GCP, 온프레미스 환경의 자원(서버, VM, 쿠버네티스 클러스터 등)과 로그, 메트릭을 Azure Monitor로 통합하여 관리할 수 있습니다. 이로써 운영자는 Azure Portal에서 모든 클라우드 자원을 단일 뷰로 모니터링하고, 정책 적용 및 보안 통제도 일원화할 수 있습니다. Azure SRE Agent는 인시던트 자동 감지 및 대응, 안전장치 기반 자동화(예: 승인 기반 remediation)를 지원합니다. GitHub Copilot for Azure(2025-06 GA, agent mode)는 SRE 및 개발자가 자연어로 운영 작업을 자동화할 수 있도록 지원하며, Azure Arc와의 연계로 멀티클라우드 DevOps 시나리오를 실현합니다.

Azure Arc의 가장 큰 강점은 멀티클라우드 환경에서의 일관된 정책 관리와 보안 통제입니다. 예를 들어, AWS EC2 인스턴스와 GCP Compute Engine VM을 Azure Arc에 등록하면, Azure Policy와 Security Center를 통해 동일한 보안 정책과 규정 준수를 적용할 수 있습니다. 또한, Azure Monitor를 통해 모든 자원의 로그와 메트릭을 집계·분석할 수 있어, 운영 복잡성을 크게 줄여줍니다. SRE Agent는 인시던트 발생 시 자동 감지 및 대응, 승인 기반 자동화 등 운영 안정성을 높이는 기능을 제공합니다. GitHub Copilot for Azure와의 연계는 자연어 기반 운영 자동화, 코드 생성, 운영 작업의 효율화 등 DevOps 생산성을 극대화하는 데 기여합니다.

### 멀티클라우드 단일 콘솔의 실제 적용 범위

Azure Arc는 AWS EC2, GCP Compute Engine, 온프레미스 VM, 쿠버네티스 클러스터까지 지원하며, Azure Monitor, Security Center, Policy 등 Azure의 다양한 관리 도구와 연동이 가능합니다. Korea Central 리전에서 Azure Arc 및 SRE Agent의 AIOps 기능이 공식적으로 지원되어, 국내 대기업·금융권의 멀티클라우드 통합 요구에 부합합니다.

실제 현장에서는 Azure Arc를 통해 멀티클라우드 자원의 통합 모니터링, 정책 일관성 확보,

보안 규정 준수, 자동화된 인시던트 대응 등 다양한 운영 시나리오가 구현되고 있습니다. 예를 들어, 국내 대기업이 AWS와 Azure를 병행 사용하는 환경에서 Azure Arc를 도입하면, 모든 자원을 Azure Portal에서 일괄 관리할 수 있어 운영 효율성과 가시성이 크게 향상됩니다. 또한, SRE Agent와 연계하여 인시던트 대응 속도를 높이고, GitHub Copilot for Azure를 통해 반복적인 운영 작업을 자동화함으로써 인적 리소스 부담을 줄일 수 있습니다.

### Azure Arc 멀티클라우드 흡수 범위 매트릭스

기능	지원 클라우드	한국 리전 지원
자원 등록/관리	Azure/AWS/GCP/온프레미스	O
로그/메트릭 통합	Azure/AWS/GCP/온프레미스	O
SRE Agent 자동화	Azure/AWS/GCP/온프레미스	O
Copilot for Azure 연계	Azure/AWS/GCP/온프레미스	O

(출처: [Azure SRE Agent](#), [Azure Arc](#))

#### 의사결정 시 유의사항

멀티클라우드 비율이 30% 이상인 조직(특히 한국 대기업, 금융권)은 Azure Arc를 진지한 후보로 검토할 필요가 있습니다. 단, Arc 기반의 통합은 Azure 생태계에 대한 종속성을 동반하므로, 장기적 멀티벤더 전략과의 정합성도 함께 평가해야 합니다. 또한, Azure Arc의 통합 기능은 Azure의 정책 및 보안 체계에 기반하므로, 타 CSP와의 연동 범위 및 심층 기능 지원 여부를 사전에 충분히 검토해야 하며, 필요에 따라 독립 오픈버빌리티 벤더와의 병행 전략도 고려해야 합니다.

#### 4.2.2 분산 추적·코드 레벨 RCA의 한계 — 독립 오픈버빌리티 벤더와의 깊이 차이

Azure Arc와 SRE Agent, 그리고 Azure Monitor의 조합은 멀티클라우드 단일 콘솔 측면에서는 매우 혁신적이지만, 심층 오픈버빌리티 영역에서는 한계가 존재합니다. 특히 분산 추적, 서비스 간 호출 체인 분석, 코드 레벨 근본 원인 분석(RCA) 등에서는 Datadog, Dynatrace, Splunk 등 독립 오픈버빌리티 벤더에 비해 기능의 깊이와 자동화 수준이 부족한 것이 현실입니다. 본 절에서는 Azure Arc 조합의 심층 오픈버빌리티 한계와, 독립 벤더와의 비교, 그리고 2025년 Gartner MQ 평가와 AWS의 공식 입장을 바탕으로 의사결정 시 고려해야 할 사항을 구체적으로 제시합니다.

#### Azure Arc 조합의 심층 오픈버빌리티 한계

Azure Arc + Monitor + SRE Agent 조합은 멀티클라우드 단일 콘솔 면에서는 매우 공격적인 전략이지만, 분산 추적(distributed tracing), 서비스 간 호출 체인 분석, 코드 레벨 RCA 등에서 Datadog, Dynatrace, Splunk 등 독립 오피저빌리티 벤더에 비해 깊이가 부족합니다. 특히 마이크로서비스 환경에서의 트랜잭션 추적, APM(애플리케이션 성능 모니터링) 심층 분석, 자동 RCA(근본 원인 분석) 등은 독립 벤더가 여전히 우위에 있습니다.

예를 들어, Datadog이나 Dynatrace는 서비스 메시, 마이크로서비스 아키텍처에서의 분산 추적을 자동화하고, 코드 레벨에서의 성능 병목, 예외 발생 지점, 호출 체인 전체를 시각화하여 근본 원인을 신속하게 파악할 수 있습니다. 반면 Azure Arc와 SRE Agent는 기본적인 로그 및 메트릭 통합, 이벤트 기반 인시던트 감지에는 강점을 보이지만, 트랜잭션 단위의 심층 분석이나 코드 레벨 자동 RCA 기능은 제한적입니다. 따라서 대규모 마이크로서비스 환경이나 복잡한 분산 시스템을 운영하는 조직에서는 독립 벤더의 심층 오피저빌리티 도구가 필수적일 수 있습니다.

### 2025 Gartner MQ의 평가와 AWS의 공식 입장

2025년 Gartner MQ for Observability Platforms에서 하이퍼스케일러(CSP)는 Leader 부문에 진입하지 못했습니다. AWS는 DevOps Agent GA 발표문에서 “Datadog, Dynatrace, New Relic, Splunk, Grafana 등 독립 오피저빌리티 벤더와의 통합이 1순위”임을 명시하며, CSP-Native와 독립 벤더의 공존 전략을 공식적으로 인정했습니다.

이러한 평가와 공식 입장은 CSP-Native 오피저빌리티 솔루션이 멀티클라우드 통합, 기본 운영 자동화에는 강점을 가지지만, 심층 분석과 자동화된 근본 원인 분석 등에서는 독립 벤더의 전문성과 기능 심화에 미치지 못함을 시사합니다. 실제로 많은 글로벌 기업들이 CSP-Native 솔루션과 독립 벤더 도구를 병행 도입하여, 운영 효율성과 심층 분석의 균형을 추구하고 있습니다.

### Azure Arc vs 독립 벤더 비교 매트릭스

비교 항목	Azure Arc+SRE Agent	독립 오피저빌리티 벤더
단일 콘솔	○	○
분산 추적	△ (기본 지원)	◎ (심층 지원)
코드 레벨 RCA	△ (제한적)	◎ (자동화·심층)
멀티클라우드	◎ (공식 지원)	◎ (에이전트 방식)

(출처: [Gartner MQ Observability 2025](#), [AWS DevOps Agent GA](#))

## 의사결정 시 유의사항

Azure Arc는 멀티클라우드 단일 콘솔이 핵심 요구인 조직에 적합하나, 분산 추적·코드 레벨 RCA 등 심층 분석이 주요 KPI라면 독립 벤더와의 병행 또는 심화 PoC가 필요합니다. AWS, GCP 역시 공식적으로 독립 벤더와의 통합을 강조하므로, 단일 솔루션 도입이 아닌 하이브리드 전략이 현실적입니다. 의사결정자는 조직의 운영 목표와 KPI, 멀티클라우드 비율, 심층 분석 요구 수준을 종합적으로 고려하여, CSP-Native와 독립 벤더의 최적 조합을 도출해야 합니다.

## 4.3 한국 의사결정 시사점 — CSP-Native vs 독립 벤더의 균형, 한국 리전 가용성의 그림자

5대 CSP의 AI 운영 에이전트와 멀티클라우드 통합 전략은 글로벌 기준으로는 빠르게 진화하고 있으나, 한국 시장에서는 리전 가용성, 데이터 주권, 멀티클라우드 비중 등 특수한 구조적 요인이 도입 결정에 큰 영향을 미칩니다. 본 절에서는 4.1, 4.2의 분석을 바탕으로, 한국 의사결정자가 즉시 적용할 수 있는 분기 가이드와 실제 리전별 AI 기능 가용성 매트릭스를 제공합니다.

### 4.3.1 한국 리전 AI 기능 가용성 매트릭스 — CloudWatch Investigations·Korea Central·asia-northeast3

한국 내에서 5대 CSP의 AI 기능 가용성은 리전별로 상이하며, 서비스 출시 시점과 정책 변화에 따라 수시로 변동될 수 있습니다. 특히 금융, 공공, 대기업 등 데이터 주권과 규제 준수가 중요한 조직에서는 리전 내 공식 지원 여부가 도입 결정의 핵심 변수로 작용합니다. 본 절에서는 각 CSP별 한국 리전에서의 AI 기능 가용 현황과, 도입 시 유의해야 할 사항을 구체적으로 정리합니다.

#### 한국 리전별 AI 기능 가용성 현황

5대 CSP의 AI 기능은 각기 다른 한국 리전(서울, Korea Central, asia-northeast3 등)에서 상이한 가용성을 보입니다. AWS CloudWatch는 서울 리전에서 기본 제공되나, CloudWatch Investigations(2025-06 GA)는 한국 리전 지원 여부가 공식적으로 공지되지 않았습니다. Azure는 Korea Central 리전에서 SRE Agent 및 AIOps 기능을 공식 지원합니다. GCP Gemini Cloud Assist는 asia-northeast3(서울)에서 일부 기능만 Preview로 제공 중입니다. OCI(Oracle Cloud)는 서울 리전을 운영 중이며, Alibaba Cloud 역시 서울 리전을 보유하고 있으

나, AI 운영 에이전트의 실제 가용성은 제한적입니다.

실제 현장에서는 리전별 AI 기능 가용성의 차이로 인해, 동일한 CSP라도 서비스 도입 시점과 기능 범위가 달라질 수 있습니다. 예를 들어, AWS의 CloudWatch Investigations는 글로벌 GA 이후에도 한국 리전 지원이 지연될 수 있으며, GCP의 Gemini Cloud Assist 역시 Preview 단계에서는 일부 기능만 제공되어, 실질적 운영 자동화에는 한계가 있습니다. OCI와 Azure는 한국 리전에서의 공식 지원이 비교적 안정적이지만, Alibaba Cloud는 AI 운영 에이전트의 현지화 및 규제 대응 한계로 인해 실제 가용성이 낮은 편입니다.

### 한국 리전 AI 기능 가용성 매트릭스

CSP	한국 리전	AI 기능 가용성
AWS	서울	CloudWatch <input type="checkbox"/> , Investigations <input type="checkbox"/> (미공지)
Azure	Korea Central	SRE Agent/AIOps <input type="checkbox"/>
GCP	asia-northeast3	Gemini Cloud Assist <input type="checkbox"/> (일부 Preview)
OCI	서울	GenAI/Agent Hub <input type="checkbox"/>
Alibaba	서울	ARMS Copilot <input type="checkbox"/> (제한적)

(출처: [AWS CloudWatch Investigations](#), [GCP Gemini Cloud Assist](#))

### 의사결정 시 유의사항

한국 리전에서의 AI 기능 가용성은 시점에 따라 변동성이 크므로, 도입 직전 반드시 공식 공지와 최신 지원 현황을 확인해야 합니다. 리전 가용성의 불확실성은 단기적으로 독립 벤더(온프레미스·Self-hosted) 채택의 합리적 명분이 될 수 있습니다. 또한, 리전별 가용성 차이로 인해 멀티클라우드 또는 하이브리드 전략을 병행하는 것이 현실적인 대안이 될 수 있으며, 조직의 데이터 주권 및 규제 준수 요건을 충족하기 위해서는 CSP별 지원 현황을 지속적으로 모니터링해야 합니다.

## 4.3.2 단일/멀티클라우드 비율로 본 분기 가이드 — CSP-Native 우선 vs 독립 벤더 우선

클라우드 사용 비율과 멀티클라우드 구조는 옴저버빌리티 플랫폼 도입 전략의 핵심 분기점입니다. 한국 시장은 K-CSP(국내 클라우드)와 외산 CSP의 혼합 구조로 멀티클라우드 비중이 매우 높으며, 이에 따라 조직별로 최적의 옴저버빌리티 진영 선택이 달라집니다. 본 절에서는 클라우드 사용

비율에 따른 권장 전략과, 의사결정 트리 매트릭스를 통해 실질적인 분기 가이드를 제시합니다.

### 클라우드 사용 비율에 따른 의사결정 분기

한국 시장은 K-CSP(국내 클라우드)와 외산 CSP의 “하” 등급 혼합 구조로, 멀티클라우드 비중이 매우 높습니다. 이에 따라 조직 내 클라우드 사용 비율에 따라 도입 전략을 달리해야 합니다.

- 단일 클라우드 ≥ 80%: CSP-Native 오피저빌리티 우선 검토
- 단일 + 보조 클라우드: CSP-Native와 독립 벤더 병행 검토
- 멀티클라우드 균등(30% 이상): 독립 벤더 또는 국산 솔루션 우선 검토

단일 클라우드 집중 조직은 CSP-Native 오피저빌리티 솔루션의 통합성, 운영 효율성, 비용 최적화의 이점을 극대화할 수 있습니다. 반면 멀티클라우드 비율이 높거나, K-CSP와 외산 CSP를 병행하는 조직은 독립 벤더 또는 국산 솔루션의 도입이 운영 일관성, 심층 분석, 규제 대응 측면에서 유리할 수 있습니다. 하이브리드 전략의 경우, PoC(개념 검증)를 통해 각 솔루션의 실제 효과와 조직 적합성을 평가한 후, 최적 조합을 결정하는 것이 바람직합니다.

### 의사결정 트리 매트릭스

클라우드 사용 비율	권장 오피저빌리티 진영	보조 권장
단일 클라우드 ≥ 80%	CSP-Native	독립 벤더 보조
단일+보조 클라우드	병행(하이브리드)	PoC 후 전략 결정
멀티클라우드 ≥ 30%	독립 벤더/국산 우선	CSP-Native 일부 연동

(출처: [Gartner MQ Observability 2025](#))

### 의사결정 시 유의사항

조직의 멀티클라우드 비율이 높을수록 독립 벤더 또는 국산 솔루션의 도입 우위가 강화됩니다. 단일 클라우드 집중 조직은 CSP-Native의 운영 효율성과 통합성의 이점을 극대화할 수 있습니다. 분기 가이드는 사내 클라우드 현황 진단과 함께, 연간 IT 전략 수립 시 반드시 반영해야 할 핵심 기준입니다. 또한, 도입 이후에도 클라우드 사용 비율과 조직의 IT 전략 변화에 따라 오피저빌리티 플랫폼의 조정 및 최적화가 필요하므로, 유연한 전략 수립이 중요합니다.

## 5장: DB 모니터링 AI 분화 — 생성형 자체 탑재 vs ML+MCP 외부 연동

### 5.1 DB 6개 제품 분화 매트릭스 — SolarWinds DPA·Quest Foglight·Percona PMM·Datadog DBM·Oracle EM 24ai·MongoDB Atlas

DB 모니터링 시장은 최근 몇 년간 AI와 LLM(대형 언어 모델) 기술의 비약적 발전에 힘입어, 기존의 룰 기반·통계 기반 이상 탐지에서 생성형 AI 및 외부 LLM 연동 중심으로 빠르게 변화하고 있습니다. 이러한 변화는 단순히 기능의 추가를 넘어, 각 벤더와 오픈소스 프로젝트가 채택하는 AI 통합 전략에 따라 제품의 구조와 도입 방식, 운영 부담, 데이터 주권 확보 등 실질적인 선택 기준에 큰 영향을 미치고 있습니다. 본 절에서는 시장을 대표하는 6개 DB 모니터링 솔루션의 AI 기능, 모델 기반, GA(General Availability) 시점, 그리고 진영 분류를 한눈에 비교할 수 있는 매트릭스를 상세히 제시합니다. 이를 통해 국내외 DBA 및 IT 의사결정자는 자사 환경과 규제 요건에 적합한 도구를 신속하게 식별할 수 있습니다.

#### 5.1.1 6개 제품 × AI 기능 × 모델 기반 × GA 매트릭스

DB 모니터링 솔루션의 AI 기능은 최근 들어 크게 두 가지 방향으로 분화되고 있습니다. 첫 번째는 생성형 AI(Generative AI) 또는 LLM을 솔루션 내부에 자체적으로 내장하여 쿼리 최적화, 자연어 질의, 자동 튜닝 등 고도화된 기능을 제공하는 진영입니다. 이 방식은 벤더가 직접 AI 모델을 관리하고, 사용자는 별도의 LLM 연동 작업 없이 즉시 AI 기반 기능을 활용할 수 있다는 장점이 있습니다. 두 번째는 ML 기반의 이상 탐지 및 통계 분석에 더해, 외부 LLM과의 연동(MCP, Model Context Protocol 등)을 통해 BYO LLM(Bring Your Own LLM) 전략을 실현하는 진영입니다. 이 구조는 조직이 원하는 LLM을 직접 연동할 수 있어 데이터 주권과 유연성이 높아집니다.

제품별로 살펴보면, SolarWinds DPA는 AI Query Assist(프리뷰)와 ML 기반 이상 탐지(GA)를 지원하며, 통계적 이상 탐지와 AI 기반 쿼리 추천을 결합합니다. Quest Foglight 8/8.1은 AI Alarms(GA)와 AI Optimizer(GA)를 통해 쿼리 및 인프라 알람의 자동화와 최적화 기능을 제

공합니다. Percona PMM 3.5는 Built-in Advisors, QAN(Query Analytics), MCP 외부 LLM 연동을 지원하며, 오픈소스 기반에 BYO LLM 및 Self-managed 옵션이 특징입니다. Datadog DBM은 쿼리 회귀 자동 탐지(GA)와 Bits AI Dev Agent(프리뷰)를 통해 LLM 기반 자동 분석 및 추천을 제공합니다. Oracle EM 24ai + Autonomous DB는 Self-tuning(GA), Select AI(GA), Ask EM(GA) 등 Oracle Cohere LLM을 OCI에서 직접 구동하며, DB 도메인에 특화된 AI 통합이 강점입니다. MongoDB Atlas Performance Advisor는 Index 권고(GA)와 MCP Server(2025 GA)를 통해 클라우드 매니지드 환경에서 BYO LLM 연동을 실현합니다.

각 제품의 AI 기능은 출시 시기와 도입 난이도에 따라 조직의 마이그레이션 전략에 영향을 미칩니다. 예를 들어, Oracle EM 24ai는 기존 Oracle DB 환경과의 통합이 용이하지만, Datadog DBM이나 MongoDB Atlas는 SaaS 기반이므로 데이터 주권 및 망분리 환경에서 제약이 존재합니다. 이러한 특성은 실제 도입 시, 조직의 보안 정책, 규제 준수, 운영 인력의 역량 등에 따라 선택 기준이 달라질 수 있음을 의미합니다.

아래 매트릭스 표는 각 제품의 AI 전략, 도입 시기, 운영 부담, 그리고 진영 분류를 한눈에 비교할 수 있도록 설계되었습니다. 조직의 데이터 주권, 망분리, BYO LLM 요구사항에 따라 적합한 솔루션을 빠르게 식별할 수 있습니다.

제품명	AI 기능	모델 기반	GA 시점	진영 분류	기존 도구 마이그레이션 부담
SolarWinds DPA	AI Query Assist(P), ML 이상탐지(GA)	ML/통계 + AI 래퍼	2025-2026	생성형 자체 탑재	낮음
Quest Foglight 8/8.1	AI Alarms(GA), AI Optimizer(GA)	생성형 LLM(비공개)	2025	생성형 자체 탑재	보통
Percona PMM 3.5	Advisors+QAN+MCP 외부 LLM	ML+MCP+BYO LLM	2025-11-19	ML+MCP 외부 연동	낮음(OSS)
Datadog DBM	쿼리 회귀 자동탐지(GA), Bits AI(P)	Bits AI(비공개)	2025-2026	생성형 자체 탑재	높음(SaaS)
Oracle EM 24ai+AutonomousDB	Self-tuning(GA), Select AI(GA)	Cohere LLM on OCI	2025	생성형 자체 탑재	낮음(Oracle DB 한정)
MongoDB Atlas Performance Advisor	Index 권고(GA), MCP Server(2025 GA)	MCP+BYO LLM	2025	ML+MCP 외부 연동	보통(SaaS)

### 5.1.2 분화 패턴 — 생성형 자체 탑재 진영(4개) vs ML+MCP 외부 연동 진영(2개)

DB 모니터링 시장은 생성형 자체 탑재 진영과 ML+MCP 외부 연동 진영으로 명확하게 양분되고 있습니다. 생성형 진영은 벤더가 직접 LLM 또는 생성형 AI를 내장하여, 사용자가 별도 LLM을 선택할 필요 없이 통합된 AI 경험을 제공합니다. 이 방식은 도입 즉시 일관된 AI 경험과 빠른 Time-to-Value를 제공하며, 벤더의 지원을 통해 운영 안정성을 높일 수 있습니다. 반면 ML+MCP 외부 연동 진영은 MCP(Model Context Protocol) 등 표준 인터페이스를 통해 조직이 원하는 LLM을 직접 연동(BYO LLM)할 수 있는 개방형 구조를 채택합니다. 이 구조는 데이터 주권, 비용 통제, 사내 LLM 결합 등 유연성이 뛰어나지만, 초기 구축 및 운영 부담(예: 자체 LLM 호스팅, MCP 연동 작업)이 수반됩니다.

비즈니스 및 기술적 함의를 살펴보면, 생성형 자체 탑재 진영은 도입 즉시 일관된 AI 경험과 빠른 Time-to-Value를 제공합니다. 하지만, 벤더 종속성과 모델 투명성 부족, 데이터 주권 이슈가 존재합니다. 반면 ML+MCP 외부 연동 진영은 데이터 주권, 비용 통제, 사내 LLM 결합 등 유연성이 뛰어나지만, 초기 구축 및 운영 부담(예: 자체 LLM 호스팅, MCP 연동 작업)이 수반됩니다. 특히, 데이터가 외부로 반출되는 경로에 대한 통제가 중요한 조직에서는 ML+MCP 외부 연동 진영이 더 적합할 수 있습니다.

한국 금융권·공공기관은 망분리, PIPA, 국정원 N²SF 등 규제 환경으로 인해 모델 투명성, 데이터 주권, BYO LLM 요구가 강합니다. 이에 따라 ML+MCP 외부 연동 진영(Percona, MongoDB)이 상대적으로 높은 적합도를 보입니다. 실제로, 한국어 특화 LLM을 활용하거나 사내 데이터의 외부 반출을 엄격히 통제해야 하는 경우, MCP 연동 구조가 필수적인 선택 기준이 되고 있습니다.

아래 진영 비교 표는 각 진영의 강점과 약점, 그리고 적합한 도입 시나리오를 정리한 것입니다.

진영	강점	약점	적합 시나리오
생성형 자체 탑재	빠른 도입, 일관된 UI/UX, 벤더 지원	벤더 종속, 모델 비공개, 데이터 경계	글로벌 표준 환경, 규제 완화 조직
ML+MCP 외부 연동	BYO LLM, 데이터 주권, 유연성	구축/운영 부담, 초기 연동 작업	한국 금융/공공, 사내 LLM 활용 조직

## 5.2 생성형 자체 탑재 진영 분석 — Oracle Cohere·Quest 생성형·Datadog Bits AI·SolarWinds AI

생성형 LLM을 자체 내장한 DB 모니터링 솔루션들은 벤더가 직접 AI 모델을 관리하고, 사용자는 별도 LLM 연동 작업 없이 즉시 AI 기반 기능을 활용할 수 있습니다. 이러한 구조는 도입과 운영이 간편하다는 장점이 있지만, 벤더 종속성, 모델 투명성 부족, 데이터 주권 이슈 등도 함께 내포하고 있습니다. 본 절에서는 Oracle, Quest, Datadog, SolarWinds 4개 제품의 AI 통합 깊이, 모델 공개성, 그리고 한국 도입 시 고려해야 할 규제·거버넌스 이슈를 다각도로 분석합니다. 특히, 각 솔루션이 제공하는 AI 기능의 범위와 실제 도입 시 조직이 마주할 수 있는 제한사항, 그리고 국내 규제 환경에서의 적합성까지 구체적으로 살펴봅니다.

### 5.2.1 Oracle EM 24ai + Cohere LLM on OCI·Select AI Agents — DB 도메인 통합 깊이의 표준

Oracle EM 24ai와 Autonomous DB는 Cohere LLM을 OCI(Oracle Cloud Infrastructure) 내에 직접 탑재하여, DB 튜닝, 쿼리 최적화, 자연어 질의(Ask EM), Select AI Agents 등 DB 도메인에 특화된 AI 기능을 제공합니다. Self-tuning은 운영 자동화의 대표 사례로, DBA의 수동 개입 없이 성능 튜닝이 이루어집니다. 이러한 통합 구조는 Oracle DB 환경에 최적화되어 있으며, AI 기능이 DB 도메인에 깊숙이 내장되어 있어 운영 효율성과 생산성을 크게 향상시킬 수 있습니다.

Cohere LLM이 OCI 리전 내에서 구동되므로, 한국 리전에서 운용 시 데이터 주권을 확보할 수 있습니다. 이는 금융권·공공기관 등 데이터 외부 반출이 엄격히 제한된 조직에 중요한 장점입니다. 실제로, 한국 내에서 데이터가 국외로 전송되지 않고, 리전 내에서 모든 AI 추론 및 처리가 이루어진다는 점은 규제 준수와 보안 측면에서 매우 큰 메리트로 작용합니다.

Oracle AI 스택은 Cohere LLM이 고정되어 있어, BYO LLM(고객이 직접 LLM을 선택/교체) 옵션은 제한적입니다. 이는 MCP 기반의 외부 LLM 연동을 중시하는 조직에 단점이 될 수 있습니다. 즉, Oracle이 제공하는 LLM 외에 다른 LLM을 연동하거나 교체하는 것이 불가능하거나 매우 제한적이기 때문에, 사내 특화 LLM을 활용하려는 조직에는 한계가 있습니다.

Oracle DB 비중이 높은 조직은 Oracle AI 스택만으로도 깊은 AI 통합을 경험할 수 있으나, Datadog, Dynatrace 등 외부 오피버빌리티 도구와의 중복 투자를 피하기 위한 전략적 판단이

필요합니다. 예를 들어, 이미 다양한 DB 엔진을 혼합 운영하는 조직에서는 Oracle AI만으로 모든 요구를 충족하기 어려울 수 있으므로, 도입 전에 전체 인프라 구조와 운영 전략을 면밀히 검토해야 합니다.

아래 표는 Oracle AI DB 스택의 주요 모듈별 기능과 한국 도입 적합도를 요약한 것입니다.

모듈	주요 기능	한국 도입 적합도
EM 24ai	NL 질의, Ask EM	높음(Oracle DB 중심)
Autonomous DB	Self-tuning, Select AI	높음
Cohere LLM on OCI	LLM 기반 분석/추천	높음(리전 내 주권)

### 5.2.2 Quest Foglight 8.1 · SolarWinds DPA · Datadog DBM — 모델 비공개의 한국 도입 함의

Quest Foglight 8.1, SolarWinds DPA, Datadog DBM의 공통점은 생성형 LLM 또는 AI 래퍼의 세부 모델이 비공개라는 점입니다. 이는 벤더가 모델 업데이트, 데이터 학습, 거버넌스 정책을 독립적으로 관리한다는 의미이기도 합니다. 이러한 구조는 벤더가 지속적으로 AI 모델을 개선하고, 기능을 신속하게 업데이트할 수 있다는 장점이 있지만, 사용자는 AI 모델의 내부 동작 방식이나 데이터 처리 경로를 직접 검증하기 어렵다는 한계가 있습니다.

한국의 개인정보보호법(PIPA), 국정원 N²SF 등은 AI 모델의 투명성, 데이터 경계, 외부 전송 경로를 엄격히 요구합니다. 모델 비공개 구조에서는 AI 추론 과정, 학습 데이터 경로, 데이터 외부 전송 여부를 명확히 검증하기 어렵기 때문에, 규제산업 도입 시 정성적 감점 요인이 됩니다. 특히, 금융권이나 공공기관에서는 데이터가 외부로 유출될 가능성, 또는 AI 모델이 어떤 방식으로 데이터를 학습하고 처리하는지에 대한 투명성이 매우 중요한 평가 기준이 됩니다.

PIPA 4단계 분류상, 모델 비공개 제품은 “적법 근거 명확화” 단계에서 추가 검증 및 법무 검토가 필요합니다. 특히, 텔레메트리 데이터에 PII(개인식별정보)가 포함될 경우, 외부 LLM 추론 경로가 불투명하면 도입이 지연되거나 불가 판정을 받을 수 있습니다. 실제로, 최근 국내 금융권에서는 AI 기반 솔루션 도입 시 데이터 흐름도, 모델 투명성, 외부 연동 여부에 대한 상세 보고서 제출을 요구하는 사례가 늘고 있습니다.

아래 표는 각 제품의 모델 공개도와 한국 거버넌스 통과 가능성을 정리한 것입니다.

제품명	모델 공개도	한국 거버넌스 통과 가능성
Quest Foglight 8.1	비공개	낮음(추가 검증 필요)
SolarWinds DPA	비공개(래퍼)	낮음(추가 검증 필요)
Datadog DBM	비공개(Bits AI)	낮음(추가 검증 필요)

## 5.3 ML+MCP 외부 연동 진영 분석 — Percona PMM·MongoDB Atlas의 한국 금융권 적합도

ML+MCP 외부 연동 진영은 오픈소스 기반의 유연성, BYO LLM, 데이터 주권 확보 등 한국 금융권·공공기관의 요구에 가장 부합하는 구조를 제공합니다. 이 구조는 조직이 직접 원하는 LLM을 선택하고 연동할 수 있어, 데이터 외부 반출을 최소화하고, 사내 LLM 결합 및 비용 통제에 유리합니다. 본 절에서는 Percona PMM과 MongoDB Atlas의 MCP 연동 구조, 사내 LLM 결합 시나리오, 그리고 실제 도입 가이드라인을 구체적으로 제시합니다. 특히, 각 솔루션의 MCP 연동 방식, 운영 부담, 그리고 국내 규제 환경에서의 실질적 도입 적합성을 다양한 사례와 함께 설명합니다.

### 5.3.1 Percona PMM 3.5 + MCP — OSS 진영의 BYO LLM 표준화

Percona PMM 3.5.0은 Built-in Advisors, QAN 등 기존 ML·통계 기반 기능에 더해, MCP(Model Context Protocol) 기반 외부 LLM 연동을 공식 지원합니다. 이를 통해 조직은 OpenAI, Anthropic, 네이버 HyperCLOVA 등 원하는 LLM을 직접 연결할 수 있습니다. MCP 연동은 표준화된 인터페이스를 제공하므로, 다양한 LLM 공급자와의 연동이 용이하며, 향후 LLM 기술 변화에도 유연하게 대응할 수 있습니다.

오픈소스 라이선스 기반이므로, 도입 비용이 낮고, 자체 호스팅 및 업그레이드가 가능합니다. BYO LLM 전략을 채택함으로써, 데이터 주권, 비용 통제, 사내 LLM 결합 등 한국 금융권의 핵심 요구를 충족합니다. 예를 들어, 사내에서 자체적으로 개발한 LLM이나, 국내 클라우드 기반 LLM을 연동하여, 데이터가 외부로 유출되지 않도록 통제할 수 있습니다.

Self-managed OSS의 특성상, 운영·업그레이드·보안 패치 등 관리 부담이 있으나, 데이터 주권과 유연성 확보라는 장점이 이를 상쇄합니다. 실제로, 일부 금융기관에서는 자체적으로 Percona PMM을 구축하고, 사내 LLM과 연동하여 데이터 흐름을 완전 통제하는 사례가 늘고 있습니다.

또한, 오픈소스 커뮤니티의 활발한 지원을 통해 신속한 버그 수정과 기능 개선이 이루어지고 있어, 장기적으로도 안정적인 운영이 가능합니다.

아래 표는 Percona PMM 3.5의 MCP 통합 현황과 사내 LLM 결합 가능성, 운영 부담을 정리한 것입니다.

PMM 버전	MCP 통합	사내 LLM 결합	운영 부담
3.5.0	공식 지원	가능	자체 호스팅 필요

### 5.3.2 MongoDB Atlas Performance Advisor + MCP — 클라우드 매니지드 진영의 BYO LLM 표준화

MongoDB Atlas Performance Advisor는 Index 권고 등 기존 ML 기반 기능에 더해, 2025년부터 MCP Server를 통해 외부 LLM 연동을 지원합니다. 이를 통해 SaaS 환경에서도 BYO LLM 전략을 실현할 수 있습니다. MCP Server는 표준화된 인터페이스를 제공하므로, 조직이 원하는 LLM을 자유롭게 연동할 수 있으며, 향후 LLM 기술 발전에 따라 다양한 공급자와의 연동이 가능해집니다.

Atlas는 SaaS 기반이므로, 한국 금융권·공공기관에서 데이터 주권, 망분리 요구를 완전히 충족하기는 어렵습니다. 그러나 MCP 인터페이스를 통해 사내 LLM(예: HyperCLOVA, KT, SKT 등)과의 결합이 가능하므로, 하이브리드 도입 전략이 가능합니다. 예를 들어, 민감한 데이터는 사내 LLM을 통해 처리하고, 비민감 데이터나 공용 데이터는 외부 LLM을 활용하는 방식으로 유연하게 운영할 수 있습니다.

Atlas SaaS의 데이터 주권 한계는 도입 전 거버넌스 검토(법무, 보안, 규제 준수 등)에서 반드시 평가되어야 하며, MCP 연동의 투명성, 데이터 흐름, LLM 계약 구조가 명확해야 합니다. 특히, 데이터가 SaaS 환경에서 어떻게 처리되고, LLM과의 연동 과정에서 외부로 전송되는 데이터의 범위와 경로를 명확히 정의해야 합니다. 이를 위해, 도입 전 사전 거버넌스 평가와 내부 정책 수립이 필수적입니다.

아래 표는 MongoDB Atlas의 MCP 연동 현황과 사내 LLM 결합 가능성, 그리고 거버넌스 게이팅 단계를 정리한 것입니다.

제품명	MCP 연동	결합 가능 사내 LLM	거버넌스 게이팅 단계
MongoDB Atlas	2025 GA	가능	데이터 주권·계약 검토 필요

### 5.3.3 한국 금융권·공공 도입 권장 시나리오 — Percona PMM/Atlas+MCP를 1차 도입 후보로 삼는 이유

ML+MCP 진영의 Percona PMM, MongoDB Atlas는 모델 비공개 구조를 회피할 수 있고, 조직이 직접 계약한 LLM(사내 sLLM, 정부 클라우드 LLM 등)을 연결할 수 있습니다. 이러한 구조는 데이터 주권과 투명성, 그리고 규제 준수에 유리하며, 국내 금융권·공공기관의 도입 요구에 부합합니다. 예를 들어, 사내에서 개발한 LLM이나 국내 CSP(Cloud Service Provider)가 제공하는 LLM을 연동하여, 데이터가 국외로 반출되지 않도록 통제할 수 있습니다.

사내 LLM 결합이 가능하므로, 데이터 외부 반출 없이, 한국어 특화 LLM을 활용할 수 있습니다. 또한 LLM 토큰 비용을 벤더가 아닌 고객이 직접 부담하는 구조이므로, 비용 통제와 투명성이 뛰어납니다. 실제로, LLM 사용량이 많아질수록 토큰 비용이 중요한 이슈가 되는데, BYO LLM 구조에서는 조직이 직접 비용을 모니터링하고 최적화할 수 있습니다.

아래 DB 엔진별 권장 매트릭스는 각 엔진에 적합한 도구와 권장 이유, 그리고 주의 사항을 정리한 것입니다.

DB 엔진	권장 도구	권장 이유	주의 사항
Oracle	Oracle EM 24ai	도메인 통합 깊이, 리전 내 주권	BYO LLM 제한
MySQL/PostgreSQL	Percona PMM 3.5	OSS, BYO LLM, 데이터 주권, MCP 연동	운영 부담(자체 호스팅)
MongoDB	Atlas+MCP	SaaS, BYO LLM, MCP 연동	데이터 주권, 망분리 한계
기타(NoSQL 등)	Percona/MCP	BYO LLM, 유연성	구축/운영 부담

한국 금융권 및 공공기관은 1차 도입 후보로 Percona PMM 3.5(MCP 연동), MongoDB Atlas+MCP를 우선 검토하고, Oracle DB 환경은 Oracle AI 스택을, SaaS 환경은 데이터 주권 및 거버넌스 검토를 병행해야 합니다. 특히, 도입 전에는 데이터 흐름, LLM 연동 구조, 운영 부담, 비용 구조, 그리고 규제 준수 여부를 종합적으로 평가해야 하며, 필요시 파일럿 프로젝트를 통해

실제 운영 환경에서의 적합성을 검증하는 것이 바람직합니다.

## 6장: 리서치 기관 보고서가 말하는 AI 평가 기준의 진화 — Gartner / IDC / Forrester

### 6.1 Gartner MQ for Observability Platforms 2025 + Hype Cycle 2025 — 카테고리 통합과 GenAI→Agentic AI 정점 이동

2025년을 전후로 오피저빌리티 및 AI 평가 기준은 Gartner의 대표적 리서치 보고서인 Magic Quadrant(MQ)와 Hype Cycle을 중심으로 근본적인 변화를 맞이하고 있습니다. 이 두 보고서는 전 세계 IT 의사결정자와 벤더에게 시장 표준의 방향성과 기술 트렌드의 변곡점을 명확히 제시하는 역할을 합니다. 특히, 오피저빌리티 플랫폼 카테고리의 통합과 AI 평가축의 중심 이동은 단순한 기능 비교를 넘어, 실제 도입 전략과 벤더 선정 기준에 직접적인 영향을 미치고 있습니다. 본 절에서는 Gartner MQ 2025의 리더 그룹 변화와 Hype Cycle 2025에서 AI 기술 위치의 변동을 구체적으로 분석하여, 글로벌 시장에서 AI 평가 기준이 어떻게 진화하고 있는지 심층적으로 살펴봅니다.

#### 6.1.1 MQ 2025 Leader 8사 — Datadog/Dynatrace/Splunk(Cisco)/Elastic/Grafana/IBM/Chronosphere/New Relic의 의미

##### MQ 2025 리더 그룹 변화

2025년 Gartner Magic Quadrant for Observability Platforms(G00813330, 2025-07-07)은 오피저빌리티 시장의 주요 벤더를 단일 카테고리로 평가하며, 리더 그룹의 변화와 AI 평가 비중의 증대를 명확히 보여줍니다. 기존에는 APM과 오피저빌리티가 별도의 MQ로 평가되었으나, 2025년부터는 통합 플랫폼 관점에서 단일 MQ가 발표되었습니다. 이번 MQ에서 리더로 선정된 8개 벤더는 Datadog(5년 연속), Dynatrace, Splunk(Cisco, 3년 연속), Elastic(2년 연속), Grafana Labs, IBM, Chronosphere, New Relic입니다. 신규 진입 벤더로는 ITRS가, 이탈 벤더로는 AppDynamics, Riverbed, Broadcom, ManageEngine이 있습니다. 이 중 AppDynamics와 Broadcom(구 CA)은 과거 APM 시장의 강자였으나, AI 및 통합 오피저빌리티 역량에서 후순위로

밀리며 MQ에서 제외되었습니다.

### AI 평가 비중의 결정적 역할

2025년 MQ의 가장 큰 특징은 AI 기능의 심층적 평가가 리더 선정의 결정적 변수로 작용했다는 점입니다. 리더 그룹에 등재된 벤더들은 모두 AI 기반의 자동화, 폐쇄 루프 트러블슈팅, 자연어 질의(NLQ), Agentic 워크플로, BYO LLM(Bring Your Own LLM) 지원 등 최신 AI 트렌드를 빠르게 내재화한 공통점을 보입니다. 특히, Datadog, Dynatrace, Splunk는 Agentic AI 기능을 GA(General Availability) 단계까지 상용화했으며, Elastic, Grafana, IBM은 BYO LLM 및 자체 호스팅 옵션을 통해 데이터 주권과 규제 대응력을 강화했습니다. 반면, AI 기능의 통합이 미흡하거나, SaaS 전용 모델로 한국 등 규제 시장 진입이 어려운 벤더들은 MQ에서 이탈하는 결과를 보였습니다.

### 한국 시장 진입 적합도와 리더 그룹

한국 시장 관점에서 MQ 리더 8사 중 실제 도입 가능한 후보는 Splunk, IBM, Elastic, Dynatrace, Chronosphere, Grafana 등 5~6개로 압축됩니다. 이들 벤더는 Self-hosted, Air-gap, BYO LLM 등 한국 규제 환경에 적합한 옵션을 제공하거나, 2026년 이후 이를 지원할 계획을 명확히 하고 있습니다. 의사결정자는 MQ 리더 등재 여부뿐 아니라, 해당 벤더의 한국 진입 적합도(데이터 주권, 한국어 NL 지원, 거버넌스 통과 가능성 등)를 반드시 함께 검증해야 합니다.

### MQ 2025 주요 벤더 표

벤더명	연속 등재 햇수	AI 기능 통합	한국 진입 적합도
Datadog	5년	Agentic AI, NLQ, SaaS 전용	하
Dynatrace	5년	Davis Causal, Agentic 워크플로	중-상
Splunk(Cisco)	3년	AI Assistant, BYO LLM	상
Elastic	2년	AI Assistant, BYO LLM, Self-hosted	상
Grafana Labs	2년	Sift, LLM 플러그인, BYO LLM	중-상
IBM	2년	Instana watsonx, BYO LLM	상
Chronosphere	1년	AI 기반 SLO, SaaS	중
New Relic	5년	NRAI, OpenAI 기반	하

출처: [Gartner MQ 2025](#), [NetworkWorld](#)

## 6.1.2 Hype Cycle 2025 — GenAI 환멸기 진입·Agentic AI 정점 이동의 의미

2025년 Gartner Hype Cycle은 옵저버빌리티 및 AI 관련 기술의 시장 기대치와 실제 도입 타이밍을 한눈에 보여주는 중요한 지표입니다. 특히, 기술의 기대치가 실제 가치 실현 단계로 얼마나 빠르게 이동하는지, 그리고 각 기술이 현재 어떤 위치에 있는지 파악하는 데 매우 유용합니다. 2025년 Hype Cycle에서는 GenAI(Generative AI)가 환멸기에 진입하고, Agentic AI가 기대의 정점에 위치하는 등 AI 기술의 세부 트렌드 변화가 두드러지게 나타납니다. 이 변화는 IT 의사결정자에게 도입 시점과 평가 기준을 재정립할 필요성을 시사하며, 단순한 AI 어시스턴트 보유 여부를 넘어, 실제 운영에 미치는 영향력과 자동화 수준을 중심으로 평가해야 함을 강조합니다.

### 2025 Hype Cycle의 주요 기술 위치

Gartner Hype Cycle for Monitoring and Observability 2025(G00811210, 2025-07-22)은 AI 관련 기술의 시장 기대치와 실제 도입 타이밍을 시각적으로 제시합니다. 2025년 Hype Cycle에서 가장 주목해야 할 변화는 GenAI(Generative AI)가 “Trough of Disillusionment(환멸의 골)”에 진입했다는 사실입니다. 이는 시장의 과도한 기대가 현실과 부딪혀 거품이 빠지는 단계에 들어섰음을 의미합니다. 반면, Agentic AI(에이전틱 AI)는 “Peak of Inflated Expectations(기대의 정점)”에 위치하며, 향후 1~2년 내 환멸기로의 진입 가능성이 높다는 신호입니다. 전통 ML 기반의 AIOps는 “Slope of Enlightenment ~ Plateau of Productivity”로 진입하여, 실질적 가치 창출 단계에 가까워지고 있습니다. Agentic AIOps Platform, AI Observability(LLM 시스템 관측) 등은 Innovation Trigger(혁신 촉진) 단계에 위치하여, 아직 시장 도입 초기임을 보여줍니다.

이러한 기술 위치 변화는 단순히 시장의 유행을 반영하는 것이 아니라, 실제로 벤더의 제품 전략과 조직의 도입 우선순위에 직접적인 영향을 미칩니다. 예를 들어, GenAI가 환멸기에 진입했다는 것은 단순 생성형 AI 어시스턴트만으로는 더 이상 차별화가 어렵고, Agentic AI와 같은 심화된 자동화와 자율화 기능이 새로운 경쟁력의 기준이 되고 있음을 시사합니다. 또한, Innovation Trigger 단계의 기술들은 아직 도입 위험이 크지만, 미래 성장 잠재력이 높으므로 선제적 PoC나 파일럿 프로젝트로 탐색할 필요가 있습니다.

### 의사결정 시사점: 도입 타이밍과 평가 기준

이러한 Hype Cycle 위치는 IT 의사결정자에게 도입 시점과 평가축 선택의 중요한 입력값을 제공합니다. Peak 단계(Agentic AI)는 PoC(Proof of Concept)나 제한적 도입을 권장하는 시기이며, 본격적인 GA(General Availability) 도입은 Plateau 단계 이후가 가장 적절합니다. GenAI의

환멸기 진입은 “AI 어시스턴트” 보유 여부만으로 솔루션을 평가하는 시대가 끝났음을 의미합니다. 이제는 Agentic AI의 폐쇄 루프 자율화 깊이, 워크플로 자동화, 인간 승인 게이팅, 감사로그 등 실제 운영에 미치는 영향력을 중심으로 평가 기준을 전환해야 합니다.

특히, Agentic AI는 단순한 자연어 질의나 추천 기능을 넘어, 문제 탐지부터 진단, 조치, 검증까지의 전 과정을 자율적으로 수행할 수 있는 능력을 의미합니다. 이러한 폐쇄 루프 자율화는 운영 효율성과 서비스 신뢰성을 동시에 높일 수 있다는 점에서, 향후 옹저버빌리티 및 AIOps 솔루션 평가의 핵심 기준이 될 전망입니다. 따라서, 조직은 단순히 AI 기능의 유무가 아니라, 해당 기능이 실제 운영 환경에서 얼마나 깊이 통합되어 있고, 자동화 수준이 어느 정도에 이르는지 면밀히 검증해야 합니다.

### 2025 Hype Cycle 주요 기술 위치 표

기술명	2025 Hype Cycle 위치	의사결정 함의
Generative AI	Trough of Disillusionment	거품 빠지는 단계, 단순 AI 어시스턴트는 차별화 아님
Agentic AI	Peak of Inflated Expectations	PoC/제한적 도입 적기, 곧 환멸기 진입 가능성
AIOps (전통 ML 기반)	Slope~Plateau	본격 도입·가치 창출 단계
Agentic AIOps Platform	Innovation Trigger	도입 초기, 미래 성장 기대
AI Observability (LLM)	Innovation Trigger	도입 초기, 실험적 통합 필요

출처: [Gartner Hype Cycle 2025, Gartner Press Release](#)

## 6.2 Forrester Wave: AIOps Platforms Q2 2025 — Dynatrace·ScienceLogic·Datadog Leader

Forrester Wave는 AIOps 플랫폼의 경쟁 구도를 심층적으로 분석하며, 각 벤더의 현행 기능 (Current Offering)과 미래 전략 (Strategy)에 대한 평가를 분리해 제공합니다. 2025년 Q2 보고서는 “Generative AI는 더 이상 차별화 포인트가 아니며, Agentic 워크플로와 자가 치유 자동화의 통합 깊이가 진정한 경쟁력”임을 명확히 선언합니다. 이 절에서는 Forrester Wave의 리더 그룹 선정 기준과 핵심 인용문을 통해, IT 의사결정자가 벤더 평가 시 반드시 고려해야 할 차별화 축을

구체적으로 제시합니다.

## 6.2.1 Leader 3사 — Dynatrace(Current Offering 최고)·ScienceLogic(Strat-egy 최고)·Datadog

Forrester Wave: AIOps Platforms Q2 2025 보고서는 AIOps 시장에서의 경쟁 구도와 각 벤더의 강점을 세밀하게 분석한 결과를 제공합니다. 특히, 리더 그룹 선정 시 단순히 현재 기능만이 아니라, 미래 전략과 혁신 가능성까지 별도로 평가하는 점이 큰 특징입니다. 이로 인해, 조직은 단기적인 도입 효과뿐만 아니라, 장기적인 기술 발전 방향까지 고려하여 벤더를 선택할 수 있습니다. 본 소절에서는 리더 3사의 선정 배경과 평가 항목별 차별화 포인트, 그리고 Strong Performer로 선정된 벤더의 의미를 상세히 살펴봅니다.

### Leader 3사 및 Strong Performer 평가

Forrester Wave: AIOps Platforms Q2 2025(RES182220)에서는 Dynatrace, ScienceLogic, Datadog 3개사가 Leader로 선정되었습니다. Dynatrace는 17개 평가 항목에서 만점을 기록하며 Current Offering(현재 기능) 부문에서 최고점을 획득했습니다. ScienceLogic은 Strategy(미래 전략) 부문에서 10개 항목 만점을 받아, 차세대 AIOps 시장에서의 성장 가능성을 인정받았습니다. Datadog은 지속적인 기능 확장과 AI 통합을 통해 Leader 그룹에 등재되었습니다. Strong Performer로는 BMC Helix가 선정됐으며, 특히 agentic AI vision/roadmap 항목에서 만점을 기록해 미래 Agentic 워크플로 통합에 강점을 보였습니다.

Dynatrace는 AI 기반의 근본 원인 분석(RCA), 자동화된 문제 해결, Agentic 워크플로의 실질적 구현 등에서 높은 평가를 받았습니다. ScienceLogic은 미래 전략과 로드맵의 실현 가능성, Agentic AI 통합 계획의 구체성에서 강점을 보였으며, Datadog은 SaaS 기반의 유연한 확장성과 자연어 질의(NLQ) 지원, 다양한 AI 기능의 빠른 도입이 장점으로 꼽힙니다. BMC Helix는 아직 리더 그룹에는 들지 못했으나, Agentic AI 비전과 로드맵 부문에서 높은 점수를 받아 향후 성장 가능성이 주목됩니다.

### Current Offering vs Strategy의 평가 분리

Forrester Wave의 또 다른 특징은 현재 기능과 미래 전략을 명확히 분리 평가한다는 점입니다. Dynatrace와 ScienceLogic은 각각 현재와 미래의 강점을 대표하며, Datadog은 두 영역에서 균형 잡힌 성과를 보입니다. 의사결정자는 벤더가 현재 제공하는 기능뿐 아니라, 향후 1~2년 내 출시

예정인 Agentic AI, BYO LLM, MCP 호환성 등 전략적 로드맵의 실현 가능성까지 함께 평가해야 합니다. ScienceLogic은 Gartner MQ에서는 Visionaries(비전가)로 분류되지만, Forrester에서는 Leader로 평가되어 리서치 기관별 평가축의 차이도 확인할 수 있습니다.

이러한 평가 방식은 조직이 단기적인 기술 도입 효과에만 치우치지 않고, 미래의 변화에 유연하게 대응할 수 있는 벤더를 선택하도록 돕습니다. 예를 들어, 현재 기능이 우수하더라도 미래 전략이 부족한 벤더는 장기적으로 경쟁력이 약화될 수 있으므로, 두 평가 항목을 균형 있게 고려하는 것이 중요합니다. 또한, Forrester와 Gartner의 평가 기준이 다르기 때문에, 두 기관의 리포트를 교차 검증하여 자사에 가장 적합한 벤더를 선정하는 것이 바람직합니다.

### Leader 3사 및 Strong Performer 표

벤더명	Current Offering 점수	Strategy 점수	평가 강점
Dynatrace	만점(17개 항목)	우수	Agentic AI, RCA, 자동화
ScienceLogic	우수	만점(10개 항목)	미래 전략, Agentic 통합
Datadog	우수	우수	AI 기능 확장, NLQ, SaaS
BMC Helix	강함	강함	Agentic AI vision/roadmap 만점

출처: [Forrester Wave Q2 2025](#), [Dynatrace News](#)

## 6.2.2 Forrester 핵심 인용 — “GenAI 평준화·Agentic 통합 깊이 차별화”

Forrester Wave Q2 2025 보고서는 AI 기능의 차별화 기준이 어떻게 변화하고 있는지 명확히 보여주는 핵심 인용문을 제공합니다. 이 인용문은 단순히 모든 벤더가 LLM 기반의 생성형 및 에이전틱 AI 기능을 제공하게 되었음을 넘어, 이제는 AI 기능의 유무가 아니라, 실제 운영에 미치는 영향력과 자동화 수준이 진정한 경쟁력임을 강조합니다. 본 소절에서는 Forrester의 핵심 메시지와, 이를 바탕으로 IT 의사결정자가 평가 기준을 어떻게 전환해야 하는지 구체적으로 안내합니다.

### Forrester의 핵심 메시지

Forrester Wave Q2 2025 보고서에서는 다음과 같이 명확히 인용합니다:

“Generative AI is not replacing... instead, it’s leveling the playing field as all vendors integrate generative and agentic AI capabilities based on the same large language models (LLMs), with every vendor now making claims that their platform is genAI-

and agentic AI-based.”

이 인용문은 단순히 LLM을 도입했다는 사실만으로는 더 이상 경쟁 우위를 확보할 수 없다는 점을 분명히 합니다. 실제로, 주요 벤더들은 모두 유사한 LLM 기반 기능을 내장하고 있으며, 이제는 Agentic 워크플로의 통합 깊이, 페르소나별 어시스턴트, 자가 치유 자동화 등 실질적인 운영 자동화 수준이 차별화 포인트로 부상하고 있습니다.

### 의사결정자에게 주는 함의

이 인용문은 LLM(대형언어모델) 자체의 도입이 더 이상 차별화 포인트가 아님을 의미합니다. 즉, 모든 벤더가 동일한 LLM 기반의 생성형 및 에이전틱 AI 기능을 내장하게 되었으므로, 이제는 “AI가 있느냐”가 아니라 “Agentic 워크플로의 통합 깊이”, “페르소나별 어시스턴트”, “자가 치유 자동화” 등 실제 운영에 미치는 영향력과 자동화 수준이 진정한 경쟁력임을 강조합니다.

→ 의사결정자는 RFP 평가 시 “AI 탑재” 여부가 아닌, 워크플로 자동화의 실질적 깊이와 운영 효율성, 거버넌스 내재화 수준을 중심으로 평가 기준을 전환해야 합니다.

실제로, 최근 RFP(제안요청서) 작성 시 많은 조직이 “AI 기능 보유”만을 요구하는 경향이 있었으나, Forrester의 분석에 따르면 이러한 기준은 더 이상 미래 경쟁력을 담보하지 못합니다. 따라서, 조직은 Agentic AI의 폐쇄 루프 자율화, 자동화된 워크플로의 실제 구현 사례, 운영 거버넌스 내재화 등 구체적이고 실질적인 평가 항목을 중심으로 벤더를 검증해야 합니다. 이는 단순한 기술 스택 비교를 넘어, 조직의 운영 효율성과 혁신 역량을 극대화하는 데 핵심적인 역할을 하게 됩니다.

Forrester 인용문	의사결정자 번역	9장 결론 연결 인덱스
“GenAI is leveling the playing field...”	“이제는 AI 기능 유무가 아니라, Agentic 워크플로 통합 깊이가 차별화 포인트”	9.2.1, 9.3.1

출처: [Forrester Wave Q2 2025](#)

## 6.3 IDC MarketScape 2025·2026 + 평가 기준의 진화 (2024→2025→2026)

IDC MarketScape는 오피스버빌리티와 AIOps 시장을 별도의 보고서로 분리 평가하며, 평가 대상의 폭과 심도, 평가 키워드의 연도별 진화를 통해 시장의 패러다임 전환을 입증합니다. 본 절에서는

IDC의 두 보고서(Observability 2025, AIOps 2026) 리더 명단과, 3대 리서치 기관의 평가 기준이 어떻게 진화하는지 연도별로 비교 정리합니다.

### 6.3.1 IDC MarketScape Worldwide AIOps 2026 + Worldwide Observability Platforms 2025의 Leader

IDC MarketScape 보고서는 오피저빌리티와 AIOps 시장을 각각 별도의 평가 프레임워크로 분석하여, 각 시장의 특성과 벤더의 강점을 세밀하게 드러냅니다. 특히, 평가 대상 벤더의 수가 많고, 각 벤더의 기능별 세부 역량에 대한 심층 평가가 이루어지기 때문에, 조직은 자사의 우선순위에 맞는 리더 벤더를 선택할 수 있습니다. 본 소절에서는 IDC AIOps 2026과 Observability 2025 보고서의 리더 선정 배경과, Gartner와 IDC의 평가축 차이를 구체적으로 비교합니다.

#### IDC MarketScape 두 보고서의 평가 대상과 리더

IDC MarketScape: Worldwide AIOps 2026(US54116226, 2026-03)은 17개 벤더를 평가하며, 개별 벤더 보도에 따르면 New Relic이 Leader로 등재되었습니다. 한편, IDC MarketScape: Worldwide Observability Platforms 2025(US53004325, 2025-11)는 26개 벤더를 평가하여, Elastic이 Leader로, HPE OpsRamp가 Major Player로 선정되었습니다. 이 두 보고서는 평가 대상의 범위와 평가축이 다릅니다. AIOps 2026은 분석 엔진의 심도, 자동화, RCA(근본 원인 분석) 역량에 초점을 맞추고, Observability 2025는 데이터 플레인(메트릭, 로그, 트레이스 등)의 폭과 통합성에 더 큰 비중을 둡니다.

IDC의 평가 방식은 실제 배포 사례와 기능별 세분화, 데이터 통합 범위, 자동화 수준 등 실질적인 운영 역량에 초점을 맞추고 있습니다. 예를 들어, New Relic은 AIOps 2026에서 분석 엔진의 심도와 자동화 역량에서 높은 평가를 받았으며, Elastic은 Observability 2025에서 데이터 통합성과 확장성, 다양한 데이터 소스 지원 측면에서 리더로 선정되었습니다. HPE OpsRamp는 Major Player로 분류되었으나, 통합성이나 자동화 수준에서 추가 개선이 필요한 것으로 평가되었습니다.

#### Gartner와 IDC 평가축의 차이

IDC 보고서는 한국 의사결정자에게 다소 익숙하지 않을 수 있으나, Gartner와는 평가 모델 자체가 다를 수 있습니다. Gartner는 시장 리더십과 비전, 실행력을 중시하는 반면, IDC는 기능별 세분화와 실제 배포 사례, 데이터 통합 범위, 자동화 수준 등 보다 세부적인 평가 항목을 강조합니다. 따라서, RFP 작성 시 두 기관의 평가축 차이를 고려하여, 조직의 우선순위에 맞는

리더 벤더를 선택하는 것이 중요합니다.

예를 들어, Gartner MQ에서 리더로 선정된 벤더가 IDC에서는 Major Player로 분류될 수 있으며, 반대로 IDC 리더가 Gartner에서는 Visionary로 평가될 수 있습니다. 이는 각 기관이 중시하는 평가 항목과 시장 관점이 다르기 때문입니다. 따라서, 조직은 두 기관의 평가 결과를 교차 검증하고, 자사의 도입 목적과 우선순위에 가장 부합하는 벤더를 선정하는 것이 바람직합니다.

보고서명	평가 대상 수	확인된 Leader	평가축 차이
IDC AIOps 2026	17	New Relic	분석 엔진 심도, 자동화, RCA
IDC Observability 2025	26	Elastic	데이터 플레인 폭, 통합성
Gartner MQ 2025	20	Datadog, Dynatrace 등 8사	시장 리더십, 비전, 실행력

출처: [IDC AIOps 2026](#), [IDC Observability 2025](#), [Elastic Leader 보도](#), [New Relic Leader 보도](#)

### 6.3.2 평가 기준의 진화 — 2024 ‘AI 어시스턴트’ → 2025 ‘Agentic 폐쇄 루프’ → 2026 ‘Outcome-centric, governed automation’

2024년부터 2026년까지 3대 리서치 기관(Gartner, IDC, Forrester)의 평가 기준은 매우 빠르게 진화하고 있습니다. 이 변화는 단순히 기술 트렌드의 이동이 아니라, 실제로 조직이 어떤 기준으로 솔루션을 평가하고 도입해야 하는지에 대한 방향성을 제시합니다. 본 소절에서는 연도별 평가 키워드의 변화와, RFP 작성 시 평가축 선택의 중요성을 구체적으로 설명합니다.

#### 연도별 평가 키워드 변화

3대 리서치 기관(Gartner, IDC, Forrester)의 평가 기준은 2024년부터 2026년까지 급격히 진화하고 있습니다. 2024년에는 “AI 어시스턴트 보유 여부”가 RFP의 핵심 평가축이었으나, 2025년에는 “Agentic AI로 detect→diagnose→remediate→verify 폐쇄 루프 자율화 깊이”가 중심이 되었습니다. 2026년에는 IDC가 “Outcome-centric decision-making, AI-driven guidance, governed automation”을 핵심 키워드로 제시하며, 결과 중심의 의사결정, AI 기반 가이드, 거버넌스 내재화 자동화가 평가의 핵심이 됩니다.

2024년에는 대부분의 조직이 AI 어시스턴트의 보유 여부, 즉 자연어 질의(NLQ)나 챗봇형 기능의 제공 여부만을 중점적으로 평가했습니다. 그러나 2025년부터는 단순한 AI 어시스턴트가

아닌, Agentic AI를 통한 폐쇄 루프 자율화(detect→diagnose→remediate→verify)가 평가의 중심이 되었습니다. 이는 문제 탐지부터 진단, 조치, 검증까지의 전 과정을 자동화하고, 인간의 개입을 최소화하는 능력이 실제 운영 효율성과 서비스 신뢰성에 결정적인 영향을 미치기 때문입니다.

2026년에는 IDC가 Outcome-centric, governed automation을 핵심 키워드로 제시하면서, 결과 중심의 의사결정과 AI 기반 가이드, 그리고 자동화 거버넌스 내재화가 평가의 중심이 됩니다. 즉, 단순히 자동화 기능을 제공하는 것을 넘어, 조직의 목표 달성에 얼마나 기여하는지, 그리고 자동화된 프로세스가 거버넌스와 규제 요건을 얼마나 충실히 반영하는지까지 평가의 범위가 확장됩니다.

### RFP 작성 시 평가축 선택의 중요성

의사결정자는 RFP 작성 시 자사의 도입 시점과 시장의 평가축 진화 흐름을 반드시 일치시켜야 합니다. 2024년 기준의 “AI 어시스턴트” 만을 평가하면, 2026년에는 경쟁력이 없는 솔루션을 도입하게 될 위험이 있습니다. 본 백서 9장 결론에서 제시할 5축 의사결정 매트릭스는 이 평가 기준 진화를 그대로 반영하며, 실질적 운영 자동화와 거버넌스 내재화가 핵심임을 강조합니다.

예를 들어, 2024년 기준으로 도입된 솔루션이 2026년에는 시장의 평가 기준 변화로 인해 경쟁력이 급격히 약화될 수 있습니다. 따라서, 조직은 도입 시점에만 집중하지 말고, 향후 2~3년 내 평가 기준의 변화까지 고려하여 RFP를 작성해야 하며, 이를 통해 장기적인 운영 효율성과 혁신 역량을 확보할 수 있습니다. 또한, 평가 기준의 진화는 단순한 기능 비교를 넘어, 조직의 전략적 목표와 일치하는 방향으로 벤더를 선정하는 데 중요한 역할을 합니다.

연도	평가 키워드	의사결정 함의
2024	AI 어시스턴트 보유 여부	단순 NLQ, 챗봇형 기능 평가, 차별화 약화
2025	Agentic AI 폐쇄 루프 자율화	detect→diagnose→remediate→verify 자동화 깊이, 워크플로 중심 평가
2026	Outcome-centric, governed automation	결과 중심 의사결정, AI 가이드, 자동화 거버넌스 내재화

출처: IDC AIOps 2026

## 7장: 한국 시장의 4대 구조적 제약과 글로벌 벤더의 한계

### 7.1 한국 4대 규제 제약 — 망분리·PIPA·국정원 N<sup>2</sup>SF·CSAP

한국 시장에서 글로벌 AI 기반 오피저버빌리티 플랫폼의 도입은 단순히 기술적 우수성만으로 결정되지 않습니다. 특히 금융, 공공, 대기업 등 규제산업에서는 망분리, 개인정보보호법(PIPA), 국가정보원 N<sup>2</sup>SF, 그리고 CSAP(Cloud Security Assurance Program) 등 4대 구조적 규제 제약이 도입의 1차 게이트로 작동합니다. 이 절에서는 각 제약이 실제로 어떻게 글로벌 SaaS-AI 솔루션의 진입을 차단하는지, 그리고 의사결정자가 사내 PoC 혹은 RFP 단계에서 반드시 체크해야 할 포인트를 구체적으로 정리합니다. 또한, 각 제약의 변화 시점과 정책 동향을 표로 정리하여, 단순한 규정 해석을 넘어 실질적인 도입 가능성 판단에 도움이 되도록 구성하였습니다.

#### 7.1.1 망분리 — 금융위 SaaS 개선 로드맵·국정원 N<sup>2</sup>SF의 SaaS-AI 진입 차단

금융 및 공공기관에서 망분리 정책은 정보보호의 핵심 수단으로 자리잡고 있습니다. 망분리란 업무망과 인터넷망을 물리적으로 분리하여 외부 위협으로부터 내부 시스템을 보호하는 정책을 의미합니다. 이러한 망분리 정책은 단순한 네트워크 분리 수준을 넘어, 클라우드 서비스와 SaaS, 특히 AI 기반 모니터링 솔루션의 도입에 있어 실질적인 진입 장벽으로 작동합니다. 최근 금융위원회와 국가정보원이 각각 발표한 정책 변화에도 불구하고, 외산 SaaS-AI 솔루션은 여전히 도입이 극도로 제한적입니다. 이 절에서는 망분리 정책의 변화와 그에 따른 글로벌 벤더의 진입 한계를 구체적으로 살펴봅니다.

##### 금융권 망분리 정책 변화

2024년 8월 금융위원회가 발표한 “금융분야 망분리 개선 로드맵”은 단계적으로 SaaS의 허용 범위를 넓히고 있으나, AI 모니터링 SaaS는 여전히 가장 마지막 단계에 위치합니다. 실제로 2025년 이후에도 금융권의 망분리 환경에서는 외산 SaaS-AI의 도입이 극도로 제한적일 것으로 전망됩니다. 이는 금융 데이터의 민감성과 외부 전송에 대한 보수적 거버넌스가 결합된 결과로, AI 기반 오피저버빌리티 플랫폼이 클라우드 기반 SaaS로만 제공될 경우, 도입 자체가 불가능한 구조입니다.

##### 비금융 공공기관의 국정원 N<sup>2</sup>SF 적용

비금융 공공기관 역시 국정원 N²SF(국가정보원 국가정보보호 프레임워크)의 업무망 중요도 차등 적용 단계에 진입하면서, “상·중” 등급 시스템에서는 외산 SaaS-AI 사용이 사실상 불가합니다. 이로 인해, 글로벌 벤더의 SaaS 기반 AI 모니터링 솔루션은 한국 공공시장 진입에 매우 높은 장벽을 마주하게 됩니다.

**망분리 환경 × 외산 SaaS-AI 가용성 × 시점 단계별 변화 표**

망분리 환경 구분	2024년	2025년	2026년 이후
금융권 상·중 등급	× 전면 불가	□ 일부 PoC 허용(비AI)	□ 단계적 허용(비AI 우선)
공공 상·중 등급	× 전면 불가	× 전면 불가	□ 일부 검토(비AI 우선)
민간/비규제	□ PoC 가능	□ 도입 가능	□ 도입 가능

*발행 시점에 따라 정책 변동 가능성 있으므로, 실제 도입 전 D-7 기준으로 재점검 필요*

**7.1.2 PIPA 4단계 분류·2025-08 PIPC 생성형 AI 안내서 — 텔레메트리 PII 외부 LLM 전송의 별도 적법 근거**

개인정보보호법(PIPA)은 한국 내 모든 데이터 처리 활동에 있어 가장 강력한 규제 중 하나로, 특히 AI 기반 오픈버빌리티 플랫폼의 도입에 있어 핵심적인 영향을 미칩니다. 2025년 8월 개인정보보호위원회가 발표한 생성형 AI 개발·활용 안내서는 데이터의 출처와 유형에 따라 4단계로 분류하여 각 단계별로 요구되는 적법 근거를 명확히 제시하고 있습니다. 이로 인해, 글로벌 벤더가 제공하는 SaaS-AI 솔루션이 텔레메트리 데이터에 포함된 PII(개인식별정보)를 외부 LLM으로 전송할 경우, 추가적인 법적 요건을 반드시 충족해야 하며, 그렇지 않으면 도입 자체가 불가능해집니다. 본 절에서는 PIPA 4단계 분류 체계와 실제 적용 사례, 그리고 글로벌 벤더의 도입 지연 또는 무산 사례를 구체적으로 설명합니다.

**PIPA 4단계 분류와 AI 텔레메트리**

2025년 8월 개인정보보호위원회(PIPC)가 발표한 “생성형 AI 개발·활용 안내서”는 학습 데이터의 출처별 적법 근거를 4단계로 분류하였습니다. 모니터링 텔레메트리에 user-id, 로그 실명 등 PII(개인식별정보)가 포함될 경우, 외산 SaaS LLM으로의 데이터 전송에는 별도의 적법 근거가 요구됩니다. 이로 인해, Datadog Bits AI, Atlassian Rovo, PagerDuty SRE Agent 등 글로벌 AI 모니터링 솔루션의 한국 도입이 지연되거나 불가능해지는 사례가 빈번하게 발생합니다.

### PIPA 4단계 × 모니터링 데이터 유형 × 외부 LLM 전송 가능성 표

PIPA 분류	데이터 유형 예시	외부 LLM 전송 가능성	필요한 적법 근거
1단계(공개 데이터)	시스템 메트릭	<input type="checkbox"/> 가능	없음/일반적 동의
2단계(비식별화)	익명화 로그	<input type="checkbox"/> 가능(추가 검토)	비식별화 증빙
3단계(PII 포함)	user-id, 실명	× 불가	명시적 동의/계약
4단계(민감정보)	의료/금융정보	× 전면 불가	법률상 특별 근거

사내 데이터 흐름을 반드시 이 분류에 따라 자가진단 필요

### 7.1.3 CSAP 2026 변화 — 보안적합성 검증 도입과 “상·중” 등급의 외산 차단 지속

CSAP(Cloud Security Assurance Program)는 공공기관의 클라우드 및 SaaS 도입에 있어 필수적인 보안 인증 제도로 자리잡고 있습니다. 2026년부터는 CSAP 인증 없이도 보안적합성 검증만으로 공공시장 진출이 가능해지는 정책 변화가 예고되어 있지만, 실제로는 모든 외산 SaaS-AI에 동일하게 적용되는 것은 아닙니다. 특히 “상·중” 등급의 공공 시스템에서는 여전히 외산 SaaS 도입이 차단되는 구조적 한계가 유지되고 있어, 글로벌 벤더의 AI 기반 오피저버빌리티 솔루션이 공공시장 진입에 실질적인 제약을 받습니다. 본 절에서는 CSAP 정책 변화의 이중성과, 등급별로 달라지는 도입 가능성을 표와 함께 상세히 설명합니다.

#### CSAP 정책 변화의 이중성

2026년부터는 CSAP(Cloud Security Assurance Program) 인증 없이도 보안적합성 검증만으로 공공시장 진출이 가능해지는 정책 변화가 예고되어 있습니다. 이는 외산 CSP(Cloud Service Provider)에는 우호적인 환경이지만, 오피저버빌리티 SaaS-AI에는 동일하게 적용되지 않습니다. 특히 “상·중” 등급의 공공 시스템에서는 여전히 외산 SaaS 도입이 차단되는 구조적 한계가 유지됩니다. 이로 인해, 글로벌 벤더의 AI 기반 오피저버빌리티 솔루션은 공공시장 진입에 있어 실질적인 제약을 받게 됩니다.

### CSAP 2025 vs 2026 × 공공 진출 경로 × 오피저버빌리티 SaaS 적용성 표

연도	공공 진출 경로	오피저버빌리티 SaaS 적용성(상·중 등급)	오피저버빌리티 SaaS 적용성(하 등급)
2025	CSAP 필수	× 전면 불가	<input type="checkbox"/> 일부 가능
2026	보안적합성 검증	× 전면 불가	<input type="checkbox"/> 일부 가능

공공 클라우드 허용 ≠ 공공 SaaS-AI 허용임을 명확히 인식해야 함

## 7.2 글로벌 AI 모니터링 기능의 한국 도입 가능성 분류 — 즉시 가능 / 한국형 변형 / 도입 불가

한국 시장에서 글로벌 AI 오피저버빌리티 기능의 도입 가능성은 기술적 우수성만으로 결정되지 않습니다. 실제로는 망분리, 개인정보보호, 거버넌스 등 다양한 규제 환경에 따라 “즉시 도입 가능”, “한국형 변형 필요”, “도입 사실상 불가”의 3단계로 명확히 분류할 수 있습니다. 이 절에서는 각 분류에 해당하는 대표 기능과 벤더, 그리고 실제 PoC(Proof of Concept) 단계에서 의사결정자가 고려해야 할 실무적인 체크포인트를 표와 함께 제시합니다. 이를 통해 PoC 후보군을 신속하게 좁히고, 불필요한 매몰 비용을 예방할 수 있습니다.

### 7.2.1 즉시 도입 가능 — Dynatrace Davis Predictive·Splunk MLTK·IBM Instana watsonx 자체 호스팅

한국의 규제 환경 하에서 즉시 도입이 가능한 글로벌 AI 오피저버빌리티 기능은 한정적이지만, 명확하게 존재합니다. 이러한 기능들은 주로 자체 ML 엔진을 내장하고 있거나, 데이터 주권 및 망분리 요건을 충족할 수 있는 자체 호스팅 옵션을 제공하는 솔루션에서 찾을 수 있습니다. 예를 들어, Dynatrace Davis Predictive/Causal, Splunk MLTK, IBM Instana watsonx 등은 외부 LLM에 의존하지 않고, 한국 내 인프라에 배포할 수 있어 규제산업에서도 도입이 가능합니다. 이 절에서는 즉시 도입 가능한 기능군의 특징과 실제 PoC 기간, 그리고 벤더별 차이점을 구체적으로 설명합니다.

#### 즉시 도입 가능한 기능군

한국 규제 환경에서 즉시 도입 가능한 글로벌 AI 오피저버빌리티 기능은 크게 두 가지로 나뉩니다. 첫째, 자체 ML 기반의 이상 탐지 및 예측 기능(Dynatrace Davis Predictive/Causal, Splunk MLTK, IBM Instana)이 대표적입니다. 둘째, NL(Natural Language) 질의 기반의 분석 기능 중에서도 Splunk AI Assistant Cloud Connected, IBM watsonx의 자체 호스팅 옵션은 한국

내 데이터 주권과 망분리 요건을 충족할 수 있습니다.

이러한 솔루션들은 외부 LLM이나 SaaS에 의존하지 않고, 온프레미스 또는 프라이빗 클라우드 환경에서 운영이 가능하다는 점이 가장 큰 장점입니다. 예를 들어, Dynatrace Davis Predictive는 자체적으로 이상 탐지 및 예측 알고리즘을 내장하고 있어, 민감한 데이터를 외부로 전송하지 않고도 AI 기반의 인사이트를 제공할 수 있습니다. Splunk MLTK는 다양한 머신러닝 모델을 직접 학습시키고 적용할 수 있는 기능을 제공하며, IBM Instana의 watsonx는 자체 호스팅 옵션을 통해 데이터 주권을 보장하면서도 고도화된 AI 분석 기능을 지원합니다.

또한, 이러한 솔루션들은 PoC(Proof of Concept) 단계에서 빠른 검증이 가능하다는 점도 중요한 특징입니다. 일반적으로 1~3개월 내에 기능 검증과 실제 운영 환경 적용이 가능하므로, 규제산업에서도 도입 리스크가 낮습니다. 벤더별로 제공하는 PoC 지원 범위와 기술 지원의 수준에 차이가 있으므로, 실제 도입 전 각 벤더의 지원 정책을 면밀히 확인하는 것이 필요합니다.

**즉시 도입 가능 기능 × 벤더 × PoC 예상 기간 표**

기능 유형	대표 벤더	PoC 예상 기간
ML 이상 탐지/예측	Dynatrace, Splunk, IBM Instana	1~2개월
NL 질의(자체 호스팅)	Splunk, IBM watsonx	2~3개월

*즉시 도입 가능 영역은 1~3개월 내 PoC 및 본격 도입이 가능함*

**7.2.2 한국형 변형 필요 — Elastic·Grafana·BMC HelixGPT의 BYO LLM에 한국어 sLLM 결합**

한국 시장의 규제와 언어적 특수성으로 인해, 글로벌 벤더의 AI 오피저버빌리티 기능 중 일부는 “한국형 변형”이 반드시 필요합니다. 특히 BYO LLM(Bring Your Own LLM) 구조를 채택한 Elastic AI Assistant, Grafana LLM 플러그인, BMC HelixGPT 등은 기본적으로 다양한 LLM 연동이 가능하지만, 실제 도입을 위해서는 국내 sLLM(특화 대형언어모델)과의 결합이 필수적입니다. 이 절에서는 한국형 변형이 필요한 기능군의 특성과, 변형 작업의 범위 및 예상 소요 기간, 그리고 실제 도입을 위한 파트너십의 중요성을 구체적으로 설명합니다.

**한국형 변형의 필요성**

Elastic AI Assistant, Grafana LLM 플러그인, BMC HelixGPT 등은 BYO LLM(Bring Your

Own LLM) 구조를 채택하고 있으나, 실제 한국 시장에서는 KT, 네이버, SKT 등 국내 sLLM(특화 대형언어모델)과의 결합이 필수적입니다. 이와 더불어, Dynatrace Davis 워크플로 역시 BYO LLM 옵션이 추가되어야만 한국 규제산업에서 도입이 가능합니다. 즉, 기술적으로는 가능하나, 한국어 NL 질의의 정확도와 데이터 주권을 확보하기 위한 추가 개발 및 파트너십이 전제 조건입니다.

이러한 변형 작업은 단순히 LLM 연동만으로 끝나지 않습니다. 한국어 자연어 처리의 정확도를 높이기 위한 NL 튜닝, 인증 및 보안 강화, 데이터 주권 보장 등 다양한 기술적·운영적 과제가 동반됩니다. 예를 들어, Elastic AI Assistant를 도입할 경우, 네이버 HyperCLOVA나 KT KoGPT와 같은 한국어 특화 LLM을 연동하고, 실제 운영 환경에서 NL 질의의 정확도를 검증해야 합니다. Grafana LLM 플러그인 역시 한국어 sLLM 연동과 함께, 인증 체계 강화 및 데이터 흐름의 투명성 확보가 필요합니다. BMC HelixGPT의 경우, 데이터 주권을 보장하기 위한 온프레미스 배포 또는 하이브리드 클라우드 연동이 필수적입니다.

파트너십 확보도 매우 중요한 요소입니다. 국내 sLLM 벤더와의 기술 협력, 커스텀 개발, 그리고 장기적인 유지보수 체계 구축이 선행되어야만 실제 도입이 가능합니다. 변형 작업의 범위와 소요 기간은 벤더별, 기능별로 상이하지만, 일반적으로 2~6개월이 소요되며, PoC 단계에서부터 변형 가능성을 충분히 검증해야 합니다.

**한국형 변형 필요 기능 × 벤더 × 변형 작업 범위 표**

기능/벤더	변형 필요 사항	예상 작업 범위
Elastic AI Assistant	한국어 sLLM 연동, NL 튜닝	2~4개월, 파트너십
Grafana LLM 플러그인	한국어 sLLM 연동, 인증 강화	2~4개월, 파트너십
BMC HelixGPT	한국어 sLLM 연동, 데이터 주권	3~5개월, 파트너십
Dynatrace Davis 워크플로	BYO LLM 옵션 추가	3~6개월, 커스텀 개발

*한국어 sLLM 파트너십 확보가 반드시 선행되어야 함*

**7.2.3 도입 사실상 불가 — Datadog Bits AI·PagerDuty SRE Agent·AWS DevOps Agent의 외부 LLM 고정**

글로벌 AI 오피저빌리티 솔루션 중 일부는 구조적으로 한국 시장에 도입이 불가능합니다. 이러한 솔루션들은 외부 LLM이 고정되어 있거나, Self-hosted 옵션이 제공되지 않아 데이터 주권, 망분

리, 개인정보보호 등 한국의 규제 요건을 충족할 수 없습니다. 본 절에서는 도입이 사실상 불가능한 기능군의 구조적 한계와, 그로 인해 발생하는 한국 시장의 특수한 문제점, 그리고 글로벌 시장과의 차별화 현상을 상세히 설명합니다.

### 도입 불가의 구조적 원인

Datadog Bits AI SRE, PagerDuty SRE Agent, AWS DevOps Agent 등은 외부 LLM이 고정되어 있거나, Self-hosted 옵션이 제공되지 않는 구조입니다. 이로 인해, 한국 규제산업(금융, 공공, 국방 등)에서는 데이터 주권, 망분리, PIPA 등 거버넌스 요구를 충족할 수 없어 도입이 사실상 불가능합니다. 특히, 글로벌 시장에서 가장 차별화된 에이전틱 RCA(근본 원인 분석) 및 자동 remediation 기능이 한국에서는 작동하지 못하는 구조적 단절이 발생합니다.

이러한 한계는 단순히 기술적 옵션의 부재에서 비롯된 것이 아니라, 글로벌 벤더의 비즈니스 모델과 SaaS 중심의 서비스 구조에서 기인합니다. 예를 들어, Datadog Bits AI SRE는 모든 AI 추론을 외부 LLM(주로 OpenAI API 등)에서 처리하며, 데이터가 실시간으로 해외로 전송됩니다. PagerDuty SRE Agent와 AWS DevOps Agent 역시 SaaS Only 구조로, 고객 데이터가 반드시 외부 클라우드로 이동해야 하므로, 망분리 및 PIPA 요건을 충족할 수 없습니다. 이러한 구조적 한계로 인해, 한국 시장에서는 PoC 단계에서조차 도입이 차단되며, 실제로 글로벌 시장에서 가장 혁신적인 기능들이 한국에서는 즉시 제외되는 모순이 반복되고 있습니다.

이와 같은 상황은 한국 시장의 특수성뿐만 아니라, 글로벌 벤더가 한국의 규제 환경을 충분히 반영하지 못하고 있다는 점을 시사합니다. 장기적으로는 글로벌 벤더가 Self-hosted 옵션을 제공하거나, 국내 클라우드와의 연동을 지원하는 구조적 변화를 도입하지 않는 한, 이러한 도입 불가 현상은 지속될 것으로 전망됩니다.

### 도입 불가 기능 × 벤더 × 한국 도입 막는 핵심 사유 표

기능/벤더	도입 불가 사유	규제 충돌 항목
Datadog Bits AI SRE	외부 LLM 고정, Self-hosted 부재	망분리, PIPA
PagerDuty SRE Agent	외부 LLM 고정, SaaS Only	망분리, PIPA
AWS DevOps Agent	외부 LLM 고정, SaaS Only	망분리, PIPA

*글로벌 시장 1순위 기능이 한국에서는 즉시 제외되는 구조적 모순 발생*

## 7.3 글로벌 벤더의 한국 도입 좌초 패턴 — 망분리·PIPA·콘솔 분리의 3중 게이트

글로벌 오피저빌리티 플랫폼이 한국 시장에서 도입에 실패하는 사례는 단순한 기술적 미비 때문만이 아니라, 복합적인 규제와 거버넌스, 그리고 사용자 경험의 문제에서 비롯됩니다. 특히 망분리, 개인정보보호법(PIPA), 그리고 콘솔/SKU 분리라는 3중 게이트가 동시에 작동하면서, 기술 평가를 성공적으로 마친 PoC조차도 최종 도입 단계에서 무산되는 패턴이 반복되고 있습니다. 이 절에서는 실제 좌초 사례를 바탕으로, 의사결정자가 사전에 인지해야 할 위험 요소와, 거버넌스 사전 검토의 중요성, 그리고 사용자 경험 회귀의 현실을 구체적으로 설명합니다.

### 7.3.1 망분리·PIPA에서 좌초된 PoC 사례 패턴 — 6개월 PoC가 거버넌스 검토에서 무산된 구조

한국 금융권 및 공공기관에서 글로벌 오피저빌리티 플랫폼의 PoC가 좌초되는 대표적 패턴은 기술 평가 단계에서는 성공적으로 진행되지만, 거버넌스 검토 단계에서 데이터 외부 전송, LLM 추론, PII(개인식별정보) 포함 여부 등 규제 요건을 충족하지 못해 최종적으로 도입이 무산되는 구조입니다. 이러한 패턴은 단순히 한두 번의 사례에 그치지 않고, 수년간 반복적으로 발생해 왔으며, 그 과정에서 6개월 이상의 PoC 기간과 상당한 매물 비용이 발생합니다. 조직 내 신뢰도 저하, 추가 기회비용, 그리고 차기 프로젝트에 대한 내부 저항까지 동반되는 부작용이 큼니다.

특히, 기술 PoC 단계에서는 벤더의 기술력과 기능이 충분히 검증되었다고 판단하나, 망분리 정책의 해석, PIPA 4단계 분류에 따른 데이터 흐름 진단, 그리고 외부 LLM 연동 여부 등 거버넌스 검토에서 예상치 못한 장애물이 등장합니다. 이로 인해, 이미 상당한 리소스와 비용이 투입된 이후에야 도입 불가 결론이 내려지는 경우가 많으며, 이러한 좌초 경험은 조직 내 IT 혁신에 대한 불신으로 이어질 수 있습니다.

따라서, PoC 착수 전 반드시 거버넌스 체크리스트를 사전에 작성하고, 법무팀 및 정보보호팀과의 협업을 통해 규제 충돌 가능성을 미리 점검하는 것이 필수적입니다. 이를 통해 불필요한 매물 비용과 조직 내 신뢰도 저하를 예방할 수 있습니다.

#### 좌초 사례 단계별 흐름 × 매물 비용 항목 표

단계	주요 활동	좌초 원인	매몰 비용 항목
기술 PoC	기능 검증, 성능 테스트	없음(성공)	PoC 라이선스, 인력
거버넌스 검토	망분리, PIPA 적합성	데이터 외부 전송, PII	추가 인력, 컨설팅
최종 무산	도입 불가 결정	규제 미충족	PoC 전체 비용, 기회비용

*PoC 착수 전 거버넌스 체크리스트 필수 삽입 권고*

### 7.3.2 SKU 분리·콘솔 분리의 사용자 경험 회귀 — “기능은 많은데 한국어 NL이 어색하고 SRE가 같은 SPL/DQL을 다시 쓴다”

글로벌 벤더의 오피버빌리티 솔루션이 거버넌스 검토를 통과해 실제 운영에 들어간 이후에도, 사용자 경험 측면에서 여러 가지 문제가 발생하는 경우가 많습니다. 대표적으로 (a) 한국어 NL 질의의 정확도 부족, (b) APM/Logs/Metrics/Network/Synthetic 등 여러 SKU와 콘솔의 분리, (c) 그 결과 SRE가 다시 SPL(Datadog), DQL(Dynatrace) 등 쿼리 언어를 직접 작성해야 하는 사용자 경험 회귀 현상이 빈번하게 나타납니다.

이러한 현상은 단순히 기능의 다양성에서 비롯된 것이 아니라, 실제 운영 환경에서의 통합 부족과 현지화 미흡에서 기인합니다. 예를 들어, NL 질의 기능이 도입되었음에도 불구하고, 한국어로 복잡한 장애 상황을 자연스럽게 질의하고 응답받는 것이 어렵다면, SRE는 결국 기존의 쿼리 언어를 직접 작성할 수밖에 없습니다. 또한, 여러 SKU가 분리되어 각각의 콘솔에서 데이터를 확인해야 한다면, 트러블슈팅 과정이 복잡해지고, 문제 해결 속도가 저하됩니다.

이러한 사용자 경험 회귀는 실제로 PoC 단계에서는 잘 드러나지 않지만, 운영 단계에서 심각한 비효율로 이어질 수 있습니다. 따라서, PoC 평가서에 NL 질의 정확도, 콘솔 전환 횟수, SRE의 직접 쿼리 작성 비율 등 구체적인 KPI를 포함하여, 실제 운영 환경에서의 사용자 경험을 사전에 검증하는 것이 중요합니다. 이를 통해, 단순히 기능의 다양성에만 집중하는 것이 아니라, 실질적인 운영 효율성과 현지화 수준을 높일 수 있습니다.

**사용자 경험 회귀 시나리오 × 발생 원인 × 측정 KPI 표**

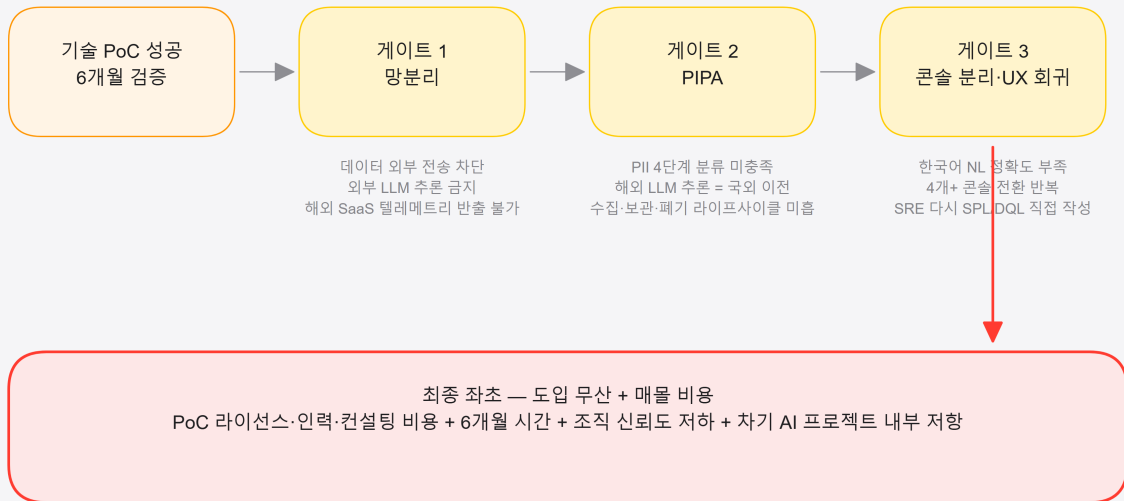
시나리오 예시	발생 원인	측정 KPI
한국어 NL 질의 오류	LLM 한국어 튜닝 미흡	NL 질의 정확도(%)
콘솔 전환 반복	SKU 기능별 분리, UI 통합 미흡	트러블슈팅 1회당 콘솔 전환 횟수

SRE 수동 쿼리 작성	NL→쿼리 자동화 미흡	SRE 직접 쿼리 작성 비율
--------------	--------------	-----------------

**PoC 평가서에 NL 정확도, 콘솔 전환 횟수 등 KPI 포함 필수**

**글로벌 벤더 한국 도입 좌초 — 망분리·PIPA·콘솔 분리의 3중 게이트**

기술 PoC는 통과해도 거버넌스 검토와 실운영 UX에서 좌초되는 6개월 매출 패턴



예방책: PoC 착수 전 거버넌스 체크리스트 의무화 — 법무·정보보호팀 사전 조인트 검토

[그림 3] 글로벌 벤더 한국 도입 좌초 — 망분리·PIPA·콘솔 분리의 3중 게이트 | 666

## 7.4 “다양한 모니터링” 함정의 한국 시장 검증 — 단일 트러블슈팅 흐름이 한국형 답이라는 결론

앞선 7.1~7.3절의 분석을 통해, 한국 시장에서는 다양한 모니터링 기능의 백화점식 나열이 오히려 트러블슈팅의 즉각적 종결을 방해한다는 구조적 결론이 도출됩니다. 실제로는 Top-down 진단과 Bottom-up 추적이 단일 콘솔에서 매끄럽게 통합되고, 한국어 NL 질의가 자연스럽게 지원되는 단일 트러블슈팅 흐름이 “한국형 답” 임이 명확히 검증되었습니다. 이 절에서는 이러한 결론을 근거와 함께 정리하고, 8장(국내 모니터링 벤더 AI 도입 현황 실측)으로의 자연스러운 연결고리를 제공합니다.

### 7.4.1 핵심 주제 #2 가설 검증 결론 — “다양한 모니터링” 백화점 모델의 한국 시장 부적합

한국 시장에서 글로벌 모니터링 업체가 어려움을 겪는 근본적인 이유는 단순히 다양한 기능의 부재가 아니라, 트러블슈팅을 즉각적으로 끝낼 수 있는 단일 흐름이 부재하기 때문임이 본 장의 분석을 통해 명확히 드러났습니다. 4대 규제 제약이 글로벌 SaaS-AI의 진입을 구조적으로 차단하고, PoC 좌초 패턴이 반복적으로 발생하며, 사용자 경험이 결국 수동 쿼리 작성과 콘솔 전환의 반복으로 회귀한다는 점이 이를 뒷받침합니다.

이러한 결론은 단순히 규제 환경의 문제만이 아니라, 실제 운영 현장에서의 효율성과 사용자 만족도, 그리고 장기적인 IT 혁신의 지속 가능성에 직결됩니다. 다양한 모니터링 기능을 백화점식으로 나열하는 글로벌 벤더의 전략은 한국 시장에서는 오히려 도입 장벽으로 작용하며, 실질적인 트러블슈팅 효율성을 저해하는 요인이 됩니다. 따라서, 단일 트러블슈팅 흐름을 중심으로 한 “한국형 답”이 실질적인 표준으로 자리잡아야 함을 강력히 선언합니다.

#### 가설 × 검증 근거 3가지 × 지지 강도 표

가설 내용	검증 근거	지지 강도
다양한 모니터링 백화점 모델은 부적합	4대 규제 제약	★★★★★
PoC 좌초 패턴	★★★★★	
사용자 경험 회귀	★★★★★	

*단일 트러블슈팅 흐름이 한국 시장의 실질적 답임을 강력히 선언*

### 7.4.2 한국형 답의 4축 — Top-down + Bottom-up + 단일 콘솔 + 한국어 NL

한국 시장에서 실질적으로 요구되는 옴져버빌리티 플랫폼의 조건은 네 가지 축으로 요약할 수 있습니다. 첫째, Top-down 진단(NL→쿼리 기반 장애 진입) 기능을 통해 사용자가 자연어로 장애 상황을 질의하고, 시스템이 자동으로 쿼리를 생성하여 원인을 진단할 수 있어야 합니다. 둘째, Bottom-up 추적(트레이스→로그→DB→인프라의 인과 추적) 기능을 통해 장애의 근본 원인을 신속하게 파악할 수 있어야 하며, 이는 복잡한 인프라 환경에서 특히 중요합니다. 셋째, 단일 콘솔 통합을 통해 모든 트러블슈팅 경로가 하나의 UI/UX에서 처리되어야 하며, 여러 개의 콘솔을 오가며

데이터를 확인하는 비효율을 최소화해야 합니다. 넷째, 한국어 NL(자연어 질의 및 응답의 정확한 지원)은 현지화의 핵심 요소로, 실제 운영 환경에서 SRE와 개발자가 언어 장벽 없이 시스템을 활용할 수 있도록 해야 합니다.

이 4축은 8장 국내 벤더 매트릭스의 평가 축이자, 9장 결론의 7개 의사결정 체크리스트의 한국 적용판이 됩니다. 실제로 이 네 가지 축을 모두 충족하는 솔루션만이 한국형 오피저버빌리티의 실질적 표준 후보로 인정받을 수 있으며, 글로벌 벤더 역시 이러한 요구 사항을 충족하기 위한 현지화 전략을 강화해야 합니다.

### 한국형 답 4축 × 정의 × 8장 국내 벤더 매트릭스 매핑 표

축	정의	8장 매트릭스 평가 행
Top-down	NL→쿼리 기반 장애 진단 진입	기본 NL 질의·RCA 자동화 행(8.3)
Bottom-up	트레이스→로그→DB→인프라 인과 추적	Self-hosted + 자체 ML RCA 행(8.3)
단일 콘솔 통합	모든 경로를 하나의 UI/UX에서 처리	6개 벤더 공통 패턴(8.1.3)
한국어 NL	자연어 질의·응답의 정확한 지원	벤더별 한국어 NL 정확도 검증 항목(8.1.1)

*4축 모두 충족하는 솔루션만이 한국형 오피저버빌리티의 실질적 표준 후보임을 명확히 함*

## 8장: 국내 모니터링 벤더 AI 도입 현황 — 출처 기반 실측

본 장은 한국 오피저버빌리티·통합관제 시장에서 활동하는 국내 6개 벤더(제니퍼소프트·와탭·엑셈·브레인즈컴퍼니·LG CNS·삼성SDS)가 2024~2026년 사이에 실제로 출시·발표한 AI 기능을 1차 출처에 근거해 정리합니다. 본 장의 목적은 추정·시나리오 수립이 아니라, 공식 블로그·릴리스 노트·국내 1차 언론 기사를 기반으로 “현재 무엇이 GA되었고, 무엇이 베타이며, 무엇이 로드맵 단계인지”를 의사결정자가 RFP 작성 시 그대로 인용할 수 있는 실측 자료로 제공하는 데 있습니다. 미검증 주장은 본문에서 제외하거나 “공식 자료에서 확인되지 않음”으로 표기했습니다.

## 8.1 국내 6개 벤더 AI 도입 현황 매트릭스 — 제니퍼소프트·와탭·엑셈·브레인즈컴퍼니·LG CNS·삼성SDS

본 절은 국내 6개 벤더의 AI 기능을 (i) 제품·라인업, (ii) 공식 발표된 AI 기능과 GA/베타/로드맵 상태, (iii) 폐쇄망 대응 방식, (iv) 1차 출처로 정리합니다. 9장의 7개 의사결정 체크리스트(BYO LLM·MCP·Self-hosted·자체 ML RCA·인간 승인 게이팅·한국 거버넌스·OTel GenAI)에 대응하기 위한 입력값으로 설계했습니다.

### 8.1.1 제품 벤더 4사 — JenniferAI · WhaTap OpsLake · eXemble · Zenius AI Agent

국내 제품 벤더 4사는 자사 APM/통합관제 제품 위에 서로 다른 경로로 AI 기능을 얹고 있습니다. 본 항은 각 사의 공식 발표 기준 AI 라인업과 폐쇄망 대응 방식을 정리하며, 추가적으로 GA되지 않은 기능과 미공개 정보(LLM 모델·고객사·NL 정확도)는 명시적으로 분리해 표기합니다.

#### 제니퍼소프트 — JenniferAI 시리즈와 브라우저 LLM 개발

제니퍼소프트는 자사 APM 제품 위에 ‘JenniferAI(제니퍼 인사이트)’ 라는 이름으로 다수의 AI 분석 기능을 단계적으로 출시해 왔습니다(S36). 공식 기술 블로그 기준 확인되는 핵심 기능은 (a) 통계 기반 Anomaly Event 탐지, (b) 브라우저 내에서 1초 단위로 실시간 동작하는 딥러닝 기반 X-View 패턴 인식(“브라우저에서 실행되므로 별도 서버도 필요 없고, 응답도 빨라 1초마다 실시간으로 패턴을 인식”), (c) 약 2년간 운영하며 25,000건 이상 질의를 처리한 RAG 기반 Help 챗봇, (d) 2025년 10월 공개한 멀티모달 LLM 외부 API 연동 ‘Insight Chat’, (e) Stacktrace Insights·Transaction Insights·Application Insights 형태의 RCA 보조 기능입니다(S36). 폐쇄망 환경 대응을 위해 자체 브라우저 LLM을 개발 중이며, 공식 블로그는 “폐쇄망 환경에서도 AI 분석 기능을 제공하기 위해 브라우저 LLM을 개발했다”, “브라우저 LLM에서 동작 가능한 대시보드/차트 인사이트 기능도 곧 공개될 예정” 이라고 명시했습니다(S36). 즉, 모든 기능이 클라이언트 사이드 브라우저 LLM으로 동작한다는 단정은 사실과 다르며, 현재는 RAG 챗봇 + 외부 API LLM(Insight Chat) + 딥러닝 브라우저 추론(X-View)이 혼재된 하이브리드 구조로, 브라우저 LLM 기반 인사이트 기능은 로드맵상 임박 공개 단계입니다(S36). 사용 LLM 모델, 한국어 NL 질의 정확도, 금융권 PoC 성과는 공식 자료에서 확인되지 않으므로 본 백서에서는 별도 수치로

인용하지 않습니다.

### 와탭(WhaTap) — OpsLake 데이터 레이크와 AI 챗봇 베타

와탭은 APM·DPM·Kubernetes·Log·Network(NMS)·Browser/Mobile RUM·GPU 모니터링을 단일 SaaS+온프레미스 패키지로 제공하며, 이를 단일 데이터 플레인으로 통합하기 위한 데이터 레이크 'OpsLake'를 개발하고 있습니다(S37). 2025년 가트너 'Infrastructure Monitoring Tools Market Guide'에서 “국내 기업 중 유일하게 대표 기업으로 선정”되었고, 같은 해 GPU 모니터링 서비스와 AI 챗봇 베타 버전을 공개했습니다(S49). 현재 1,200여 곳의 국내외 고객사를 보유하고 있으며, 방콕 리전을 추가해 글로벌 SaaS 진영의 한국형 대안 포지셔닝을 강화하고 있습니다(S49). OpsLake의 자동 RCA·이상 탐지·패턴 상관 분석은 공식 블로그상 “운영 지능화” 방향으로 기술돼 있으며 GA 시점은 별도로 명시되지 않았습니다(S37). AI 챗봇은 현재 베타 단계로 GA 여부는 본 백서 작성 시점에 확인되지 않았으며, LLM 모델 공급원·폐쇄망 대응 방식·한국어 NL 정확도 또한 공식 자료에 명시되지 않으므로 본 백서에서는 표기하지 않습니다.

### 엑셈(EXEM) — eXemble LLMOps 플랫폼과 MCP 채택

엑셈은 2026년 1월 자체 LLMOps 플랫폼 'eXemble(엑셈블)'을 공식 발표했습니다(S38). 이 플랫폼은 (a) NL2SQL 에이전트 — “데이터베이스 언어를 모르는 일반 사용자도 복잡한 코드 작성 없이 데이터에 접근”, (b) RAG 기반 AI 서비스 구축, (c) 모델 미세조정(fine-tuning)·모니터링·성능 평가, (d) IT 시스템 장애 RCA(Root Cause Analysis)와 SQL 튜닝 자동화를 지원합니다(S38). 핵심 차별화는 “MCP(Model Context Protocol) 기반의 에이전틱 AI”를 명시적으로 채택했다는 점이며, 공식 보도자료는 “온프레미스 및 폐쇄망 환경에서”의 구동을 명시했습니다(S38). 기존 MaxGauge AI Edition과 결합해 DB·APM·인프라 모니터링 데이터를 단일 LLM 워크플로우로 연계하는 구도이지만, 고객 적용 사례·NL 질의 정확도·MCP 연동 도구 카탈로그의 구체 규모는 공식 자료에서 확인되지 않습니다.

### 브레인즈컴퍼니 — Zenius 시리즈와 딥러닝 이상 탐지

브레인즈컴퍼니는 통합관제 플랫폼 'Zenius' 위에 딥러닝 기반 실시간 이상 탐지 기능을 결합한 'Zenius AI'를 공급하고 있습니다(S39, S50). 공식 보도 기준 명확히 확인되는 핵심 기능은 “방대한 IT 시스템 데이터를 실시간으로 분석해 정상 운영 패턴에서 벗어난 이상 징후를 사전에 감지”하는 딥러닝 기반 이상 탐지로 한정되며(S39), 폐쇄망·공공 비중이 높은 통합관제 시장에서의 RAG 기반 대화형 AI Agent 확장 가능성은 시장 보도 수준에서 언급됩니다(S50). 다만 1차 출처에서 GA로 확인된 것은 딥러닝 이상 탐지에 한정되며, RAG 챗봇·고객사 수·대화형 AI Agent의

출시 시점·한국어 NL 정확도는 공식 자료에서 직접 확인되지 않으므로 본 백서에서는 별도 수치로 인용하지 않습니다.

### 제품 벤더 4사 AI 라인업 매트릭스

벤더	제품/라인업	확인된 GA 기능	로드맵·미확인 영역	MCP	폐쇄망 명시 여부	출처
제니퍼소프트	JenniferAI(제니퍼 인사이트)	X-View(브라우저), Help 챗봇(RAG), Insight Chat(외부 LLM API), Stack-trace/Transaction Insights	브라우저 LLM 기반 대시보드 인사이트(공개)	미확인	개발 중(브라우저 LLM)	S36
와탭	OpsLake / APM·DPM·K8s·NMS·FIN·SEC·GPU	AI 챗봇 베타, GPU 모니터링	OpsLake AIOps 자동 RCA·이상 탐지(GA 미명시)	미확인	미확인	S37, S49
엑셈	eXemble(LLMOps) + MaxGauge AI	NL2SQL, RAG, 모델 미세조정·평가, RCA·SQL 튜닝	연동 도구 카탈로그 규모 미공개	채택 명시	온프레미스·폐쇄망 명시	S38
브레인즈컴퍼니	Zenius / Zenius AI	딥러닝 이상 탐지	대화형 AI Agent·RAG 확장 가능성	미확인	공개 비중 높음(언론)	S39, S50

매트릭스에서 드러나는 핵심 사실은 (1) 국내 제품 벤더 4사 모두 AI 기능을 자사 통합관제·APM 위에 얹기 시작했지만, GA 단계 기능과 로드맵·베타 단계가 혼재되어 RFP 평가 시 GA 여부를 일일이 확인해야 한다는 점, (2) MCP를 명시적으로 채택한 곳은 엑셈 한 곳에 그친다는 점, (3) 폐쇄망 대응의 기술적 메커니즘(브라우저 LLM·온프레미스 LLMOps·자체 데이터 레이크)이 벤더별로 다르므로 단일 평가 척도로 비교하기 어렵다는 점입니다.

### 8.1.2 SI 진영 2사 — LG CNS AIOps · 삼성SDS Fabrix/Brity Copilot

SI 진영의 LG CNS와 삼성SDS는 자체 플랫폼과 외산 솔루션을 결합해 고객 맞춤형 AIOps를 제공합니다. 본 항은 각 사가 공식 발표·보도한 AI 플랫폼의 적용 사례와 거버넌스 전략을 정리합니다.

#### LG CNS — AIOps 플랫폼 + LLMOps·MLOps 통합

LG CNS는 자체 AIOps 플랫폼을 구축해 LLMOps와 MLOps를 단일 운영 체계로 통합한다고

발표했으며(“기업의 LLM 학습과 운영에 필수적인 LLMOps와 MLOps를 하나로 통합해 구현”), 코드 어시스턴트와 분리된 가상 환경 기반 개발 지원 기능을 포함합니다(S40). 대표 적용 사례로 미래셋증권 프로젝트(2024년 11월 착수)가 보도되었으며, 미래셋생명·신한은행·신한카드·NH농협은행·KB금융지주 등 금융권 레퍼런스가 함께 언급되었습니다(S40). 거버넌스 측면에서는 ‘AI 기본법’, ‘금융분야 AI 가이드라인’, ‘금융 AI 보안 가이드라인’ 등 국내 AI 가이드라인·법률을 플랫폼에 내재화한다고 명시되었습니다(S40). 글로벌 옴저버빌리티 벤더(Splunk·Dynatrace·Elastic 등)와의 결합 제공 여부는 보도자료 본문에서 직접 확인되지 않으므로 본 백서에서는 일반화하지 않습니다.

### 삼성SDS — Fabrix · Brity Copilot/Automation의 AI 제품군

삼성SDS는 ‘Fabrix(생성형 AI 플랫폼)’와 ‘Brity Copilot/Automation’을 중심으로 업무 자동화·에이전틱 전환을 추진하고 있다고 공식 발표했습니다(S60). 통합관제 전용 AIOps 제품의 공식 분리 여부는 자료상 확인되지 않으며, 일반적으로 SI 사업 특성상 자체 플랫폼과 글로벌 솔루션(IBM·Splunk·Elastic 등)을 고객 요구에 맞춰 결합하는 구조가 보고되어 왔습니다(S60). 폐쇄망 대응·한국어 NL 정확도·개별 통합관제 적용 사례는 본 백서 작성 시점에 1차 출처에서 직접 확인된 수치가 없으므로 본 백서에서는 일반화된 SI 역량 서술에 한정합니다.

## 8.1.3 6개 벤더 공통 패턴 — 단일 흐름·한국어·폐쇄망의 동시 추구

국내 6개 벤더는 각기 다른 출발점에서 출발했으나, 공통적으로 “단일 데이터 흐름, 한국어 사용자 경험, 폐쇄망 환경 대응”이라는 세 가지 축에서 동시에 움직이고 있습니다. 이는 글로벌 4대 규제 제약(7장 참고: 망분리·PIPA·국정원 N<sup>2</sup>SF·CSAP)이 공통 전제로 작용하기 때문이며, 국내 시장에서 AI 기능이 RFP에 명시되기 시작한 2025~2026년 시점부터 본격화되었습니다.

첫째, **생성형/LLM 채택 자체는 빠르나 GA 단계 기능의 폭은 제한적입니다**. 6개 벤더 중 다수 제품은 RAG 챗봇·이상 탐지 수준에서 GA되었고, 자율 RCA·인간 승인 게이팅·OTel GenAI 의미 규약 채택 등 9장 7개 체크리스트의 후반 항목은 아직 로드맵 또는 미공개 단계입니다. 둘째, **MCP 표준 채택은 엑셈 한 곳만 명시적으로 발표했고**, 나머지 벤더는 RFP 시점에 호환성 여부를 별도 확인해야 합니다. 셋째, **폐쇄망 대응 기술 경로의 분화는 브라우저 LLM(제니퍼), 온프레미스 LLMOps(엑셈), 데이터 레이크 온프레미스 옵션(와탭), SI 맞춤(LG CNS·삼성SDS), 딥러닝 사내 모델(브레인즈)로 갈라지며**, 단일 척도로 비교하기 어렵습니다. 넷째, **거버넌스 내재화는 SI 진영**

(LG CNS)이 가장 명시적으로 'AI 기본법·금융 AI 가이드라인'을 플랫폼에 결합한다고 선언했고, 제품 벤더 진영은 여전히 고객사 컴플라이언스 검토에 의존합니다.

이러한 공통 패턴은 9장 의사결정 체크리스트의 7개 항목에 그대로 입력값으로 작용하며, RFP 작성 단계에서 글로벌 벤더 대비 국내 벤더의 강·약점을 명확히 분리해 평가할 수 있게 합니다.

## 8.2 한국 시장 AI 도입의 공백 좌표 — 4개 미충족 영역

8.1의 매트릭스를 근거로 보면, 국내 6개 벤더와 글로벌 11개 벤더(3장 참고)의 합집합으로도 동시에 충족되지 않는 4개 미충족 영역이 식별됩니다. 본 절은 이를 9장 7개 체크리스트의 축(BYO LLM·MCP·Self-hosted+자체 ML RCA·OTel GenAI)으로 재정의하여, 의사결정자가 RFP의 “동시 충족 요건” 항목을 작성할 때 활용할 수 있도록 정리합니다.

### 8.2.1 4개 공백 좌표 — BYO LLM 동시충족·MCP 표준 채택·Self-hosted 자체 ML RCA·OTel GenAI 의미 규약

**BYO LLM + 한국어 NL + 자체 호스팅 동시 충족:** 국내 6사 중 BYO LLM 옵션을 명시적으로 공개한 사례는 8.1 매트릭스 기준 확인되지 않으며, 글로벌 벤더 중 Elastic·BMC HelixGPT·IBM Instana(watsonx 결합)는 BYO LLM을 공식 지원하지만(S16, S20, S21, S70, S71) 한국 데이터센터·국정원 N²SF·금융 AI 가이드라인을 동시에 통과한 사례는 1차 출처에서 확인되지 않습니다(S41, S42, S43). 따라서 “BYO LLM × 한국어 NL × 한국 거버넌스 통과 × 자체 호스팅”의 동시 충족은 한국 시장 핵심 미충족 영역으로 남아 있습니다.

**MCP 표준 채택:** 글로벌에서는 New Relic(S52), Grafana(S73), Percona PMM(S32), MongoDB Atlas Performance Advisor(S77)가 MCP 서버·클라이언트를 공식 채택했고, 국내에서는 엑셈 eXemble(S38)이 MCP 채택을 발표했습니다. 그러나 (i) 한국형 도구 카탈로그(국내 ITSM·메신저 어댑터), (ii) 인간 승인 게이팅·감사로그를 갖춘 에이전틱 워크플로 거버넌스, (iii) MCP 호환과 한국 PIPA·금융 AI 가이드라인의 동시 통과를 모두 충족하는 사례는 본 백서 작성 시점에 1차 출처에서 확인되지 않습니다.

**Self-hosted + 자체 ML 기반 RCA(생성형 LLM 미경유):** 9장 체크리스트 (3)번 항목으로, 외부 LLM이 차단된 폐쇄망에서도 자체 ML로 RCA가 동작해야 한다는 요건입니다. Splunk MLTK 5.5 이상 탐지(S14)·Grafana Sift(S18)·Elastic AI Assistant Self-hosted(S16)는 부분적으로

충족되지만, 한국형 거버넌스 통과 + 한국어 RCA 인터페이스 + 자체 ML 모델의 동시 보장은 미확인 영역입니다. 국내 6사 중에서도 자체 ML 기반 RCA의 GA를 명시한 사례는 8.1 매트릭스 기준 한정적입니다.

**OpenTelemetry GenAI Semantic Conventions 채택 로드맵:** OTel GenAI 의미 규약 (S29)은 LLM 트레이스의 표준 스키마이며, 국내외 벤더의 채택 로드맵 공개가 향후 RFP의 표준 항목이 될 전망입니다. 국내 벤더 중 OTel GenAI 채택 로드맵을 공식 발표한 사례는 본 백서 작성 시점에 확인되지 않았습니다.

### 8.2.2 4개 공백 좌표의 RFP 입력화 — 즉시 평가 가능한 항목으로 분해

위 4개 공백 좌표는 RFP 작성 시 다음과 같이 분해되어 평가 항목으로 사용할 수 있습니다 — (1) “벤더가 제공하는 BYO LLM 옵션의 한국어 sLLM·정부 클라우드 LLM 결합 사례를 1건 이상 제시할 수 있는가?”, (2) “MCP 서버/클라이언트 호환성 + 인간 승인 게이팅·감사로그를 모두 명시적으로 지원하는가?”, (3) “외부 LLM 차단 환경에서 자체 ML 기반 RCA가 동작함을 데모로 입증할 수 있는가?”, (4) “OTel GenAI 의미 규약 채택 시점을 12개월 이내 로드맵으로 제시할 수 있는가?”. 이 4개 질문은 9장 7개 체크리스트의 핵심 부분과 직접 매핑되며, 본 백서를 닫고 즉시 RFP 양식에 옮겨 적을 수 있도록 설계했습니다.

## 8.3 한국 벤더 vs 글로벌 벤더 매핑 — 즉시 가능 / 한국형 변형 / 도입 불가 영역의 벤더별 위치

7장 7.2의 “즉시 가능 / 한국형 변형 / 도입 불가” 분류를 8.1 매트릭스와 결합하면, 국내·글로벌 벤더 각각이 어떤 영역에서 한국 시장에 즉시 투입 가능한지 한 눈에 파악할 수 있습니다.

영역	글로벌 벤더 즉시 가능	글로벌 벤더 한국형 변형 필요	글로벌 벤더 도입 불가	국내 벤더 위치
기본 NL 질의·RCA 자동화	Datadog Bits AI(S07,S08), Dynatrace Davis CoPilot(S09)	Splunk AI Assistant SPL(S13), New Relic NRAI(S11)	—	제니퍼 Insight Chat(외부 API), 엑셈 NL2SQL — 국내 데이터센터·한국어 정확도는 별도 검증 필요

폐쇄망·자체 호스팅 + 자체 ML RCA	—	Elastic(S16), Splunk MLTK(S14), Grafana Sift(S18)	Datadog SaaS, Dynatrace SaaS	엑셈 eXemble 온프레미스(S38), 와탭 온프레미스 옵션(S37), 제니퍼 브라우저 LLM(개발 중)(S36)
MCP 표준 채택	New Relic(S52), Percona PMM(S32), Grafana mcp-grafana(S73)	—	—	엑셈 eXemble(S38) — 국내 단독
한국 거버넌스 통과 (N²SF·PIPA·금융 AI 가이드라인)	—	모든 글로벌 벤더가 변형·변경 필요(S41, S42, S43)	—	LG CNS AIOps(S40)가 ‘AI 기본법·금융 AI 가이드라인’ 내재화를 명시
LLM Observability + 인프라 통합	Datadog LLM Observability(S53)	Grafana LLM 플러그인(S72)	—	국내 벤더 중 OTel GenAI 채택 로드맵을 공식 발표한 사례는 미확인

표에서 드러나는 핵심은 (1) 글로벌 벤더는 기본 NL 질의·RCA 영역에서 즉시 투입 가능하지만 한국 거버넌스를 한 곳도 통과하지 못했다는 점, (2) MCP 채택은 글로벌·국내 모두에서 아직 소수에 그친다는 점, (3) 한국 거버넌스 내재화 영역은 SI 진영(LG CNS)이 가장 앞서 있다는 점입니다.

## 8.4 2026~2028 한국 벤더 AI 로드맵 시사점 — 골든 타임 24개월 가설

본 절은 8.1~8.3의 실측 데이터를 토대로, 한국 시장에서 국산 오피저빌리티·AIOps의 골든 타임 가설을 검증합니다. 결론은 “글로벌 벤더가 한국 데이터센터·한국어 sLLM·국내 인증을 모두 갖추기 전 24개월(2026~2028년 초)이 국산 솔루션의 시장 선점 골든 타임”이라는 9장 함의 5와 정합합니다.

### 8.4.1 골든 타임 24개월 가설의 근거 — 글로벌 벤더의 한국 거버넌스 통과 부재

가설을 뒷받침하는 1차 근거는 8.3 매트릭스의 “한국 거버넌스 통과” 행이 모든 글로벌 벤더에 대해 비어 있다는 점입니다. AWS DevOps Agent(S23), Microsoft Azure SRE Agent(S24), Datadog Bits AI(S07), Dynatrace Davis CoPilot(S09) 모두 한국 리전 또는 SaaS 기반 서비스이며, 망분리(S41)·PIPA 4단계(S42)·국정원 N²SF(S43)·CSAP(S75)의 동시 통과를 명시적으로 발표한 사례는 1차 출처에서 확인되지 않습니다. 이 골든 타임 동안 국내 벤더는 (i) MCP 채택, (ii)

BYO LLM 옵션, (iii) 자체 ML RCA, (iv) OTel GenAI 채택 로드맵, (v) 거버넌스 내재화의 5축을 동시에 채워야 글로벌 벤더의 한국 진입 시점에 우위를 유지할 수 있습니다.

#### 8.4.2 의사결정자 시사점 — 9장 7개 체크리스트로의 연결

이상의 분석은 9장 결론의 7개 의사결정 체크리스트로 자연스럽게 연결됩니다. 의사결정자는 본 장의 매트릭스(8.1·8.3)와 4개 공백 좌표(8.2)를 RFP 양식에 그대로 옮겨 (1) BYO LLM, (2) MCP, (3) Self-hosted+자체 ML RCA, (4) Agentic 인간 승인 게이팅, (5) 한국 거버넌스, (6) OTel GenAI, (7) LLM 토큰 직접 계약의 7개 항목으로 평가할 수 있습니다. 이 평가표가 9.3에서 제시되는 의사결정 체크리스트의 본체이며, 본 백서를 닫는 마지막 액션 가이드가 됩니다.

## 9장: 결론 및 권장사항 — 2026~2028 의사결정 체크리스트

### 9.1 2026~2028 시점별 변곡점 — Agentic AIOps GA 대전환 / LLM 옴저버빌리티 융합 / 자율 remediation 일부 실용화

2026년부터 2028년까지는 AI 옴저버빌리티 및 AIOps 시장에서 기술, 표준, 정책 등 다양한 측면에서 중대한 변곡점이 도래하는 시기입니다. 본 절에서는 각 연도별로 글로벌 벤더의 주요 기술 출시, 표준화 진척 상황, 그리고 한국 시장에서의 정책 변화가 IT 의사결정에 미치는 영향을 체계적으로 정리합니다. 이를 통해 실제 RFP 작성, 전략 보고서 수립, 사내 PoC 착수 시점 결정 등 실무 단계에서 활용할 수 있는 구체적인 시점별 인사이트를 제공합니다. 각 변곡점은 단순한 트렌드 예측이 아니라, 도입 시점과 평가 기준을 전략적으로 조정할 수 있는 실질적 기준점으로 작용합니다.

#### 9.1.1 2026 — Agentic AIOps Preview→GA 대전환 / MCP 사실상 표준

2026년은 AI 옴저버빌리티 플랫폼 시장에서 Agentic AIOps의 본격적인 전환점이 도래하는 해로, 주요 글로벌 벤더들이 Preview 단계에서 GA(General Availability)로 제품을 전환하며 시장의

판도가 크게 변화합니다. 이 시기에는 Datadog, Azure, AWS 등 클라우드 및 모니터링 선도 기업들이 각각 Bits AI SRE, SRE Agent, DevOps Agent 등 Agentic AIOps 제품을 공식 출시함으로써, 단순 알람이나 이상 탐지 수준을 넘어선 자율 운영 체계가 실제 운영 환경에 도입됩니다. 이러한 변화는 기존의 수동적 운영 방식에서 벗어나, 자동화된 탐지, 진단, 조치, 검증이 하나의 폐쇄 루프 내에서 구현되는 방향으로 진화하고 있습니다.

특히, 각 벤더의 AI 에이전트는 인간 승인 게이팅(approval gating) 정책을 기본 제공하여, 운영 자동화의 신뢰성과 통제력을 동시에 확보할 수 있습니다. 실제로 운영 환경에서는 자동화된 문제 해결과 후속 검증까지 일련의 워크플로우가 구현되어, IT 조직의 효율성과 안정성이 크게 향상됩니다. 이러한 변화는 단순히 기술적 진보에 그치지 않고, 조직의 운영 프로세스와 거버넌스 체계에도 직접적인 영향을 미치게 됩니다.

Anthropic이 주도한 MCP(Model Context Protocol)의 표준화도 2026년의 핵심 이슈입니다. MCP는 오피저버빌리티 도구와 LLM 에이전트 간의 통합을 획기적으로 용이하게 하며, 실제로 RFP와 Vendor Scorecard에서 MCP 호환성 여부가 필수 평가 항목으로 자리잡고 있습니다. MCP를 지원하지 않는 플랫폼은 LLM 기반 트러블슈팅 자동화, 에이전틱 워크플로우, 외부 에이전트 연동 등에서 경쟁력이 급격히 저하될 수밖에 없습니다.

이처럼 2026년은 AI 오피저버빌리티 도입의 평가축이 전환되는 시점입니다. 신규 도입이나 업그레이드 시에는 Agentic 워크플로의 인간 승인 게이팅 정책의 명확성, MCP 호환성 여부를 반드시 체크리스트에 포함해야 하며, 이 두 항목이 실제 도입 성공률과 장기적 확장성에 결정적인 영향을 미칩니다. 실제로 글로벌 기업의 도입 사례를 보면, MCP 미지원으로 인해 LLM 기반 자동화 연동에 실패하거나, 인간 승인 게이팅 정책 미비로 인해 운영 리스크가 발생한 경우가 보고되고 있습니다. 따라서 2026년은 기술적, 정책적, 거버넌스적 기준이 동시에 재정립되는 중요한 해가 될 것입니다.

### 9.1.2 2027 — LLM 오피저버빌리티 + APM 본격 융합 / OTel GenAI 안정화 / 한국 SaaS-AI 단계적 허용 본격화

2027년은 LLM 오피저버빌리티와 전통적인 APM(Application Performance Monitoring) 기능이 본격적으로 융합되는 해로, IT 운영 환경의 패러다임이 한 단계 도약하는 시기입니다. Open-Telemetry GenAI Semantic Conventions가 안정화되면서, LLM 트레이스와 기존 인프라/애

플리케이션 트레이스 간의 단일 상관 분석이 가능해집니다. 이로 인해 LangSmith, Langfuse, Arize 등 LLM 모니터링 도구가 기존 APM 도구와 통합되거나, 단일 플랫폼 내에서 관리되는 구조로 전환됩니다. IT 조직은 LLM 모니터링 도구를 별도 SKU로 유지할지, 아니면 통합 관리할지에 대한 전략적 결정을 내려야 하며, 이 결정은 장기적인 운영 효율성과 데이터 품질, 감사 추적성에 직접적인 영향을 미칩니다.

또한, 2027년에는 한국 공공 및 금융권에서도 SaaS-AI 도입이 점진적으로 허용되는 변화가 본격화됩니다. 망분리, 개인정보보호법(PIPA), 국정원 N²SF 등 기존 규제의 단계적 완화와 더불어, SaaS 기반 AI 오피저빌리티 도구의 도입이 현실적인 옵션으로 부상합니다. 이 시점부터는 글로벌 SaaS 벤더의 한국 리전 가용성, 데이터 주권 확보, 한국어 자연어 질의(NL) 지원 여부가 실제 도입 성공의 핵심 변수가 됩니다. 예를 들어, 한국어 NL 지원이 미흡한 글로벌 솔루션은 공공·금융권 도입에서 실질적인 제약 요인으로 작용할 수 있습니다.

2027년은 LLM 오피저빌리티와 APM을 별도 운영할지, 단일 콘솔/플랫폼으로 통합할지에 대한 전략적 판단이 필요한 '통합 결정 시점'입니다. 실제로 글로벌 금융사 및 대형 제조사의 사례를 보면, 통합 플랫폼 도입을 통해 장애 대응 속도와 데이터 일관성이 크게 향상된 반면, 통합 미흡 시에는 데이터 사일로와 감사 추적성 저하 문제가 발생한 바 있습니다. 따라서 이 시기에는 조직의 운영 효율, 데이터 품질, 감사 추적성 등 다양한 요소를 종합적으로 고려한 전략적 의사결정이 요구됩니다.

### 9.1.3 2028 — 자율 remediation 일부 도메인 실용화 / 멀티모달 통합 오피저빌리티 초기 진입

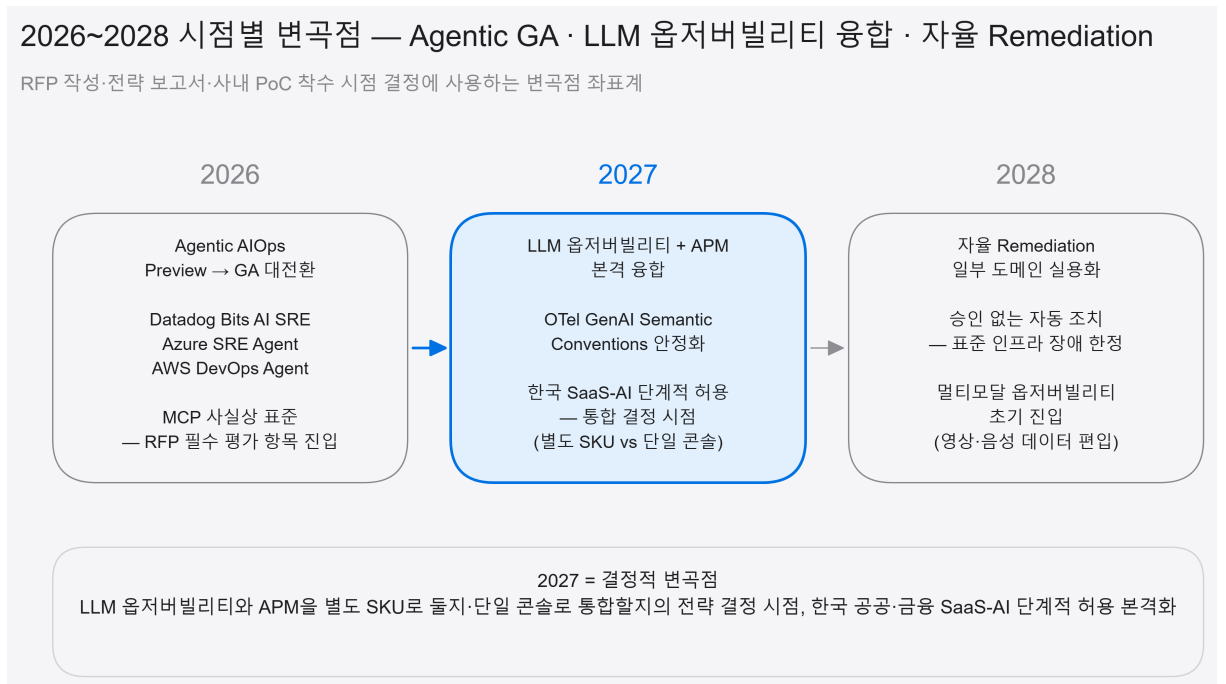
2028년은 AI 기반 자율 remediation(자동 문제 해결)이 일부 도메인에서 실용화 단계에 진입하는 해로, IT 운영의 자동화 수준이 한층 더 고도화됩니다. 승인 없는 자율 remediation은 표준화된 인프라 장애나 반복적 애플리케이션 오류 등에서 인간 개입 없이 자동으로 탐지, 진단, 조치, 검증까지 일괄 처리하는 체계를 의미합니다. 이러한 변화는 운영 효율성을 극대화하는 동시에, 운영 리스크 관리와 거버넌스 체계의 중요성을 더욱 부각시킵니다.

Agentic 워크플로의 신뢰성, 감사로그, 롤백 자동화 등 거버넌스 체계가 반드시 뒷받침되어야 하며, 실제로 글로벌 선도 기업들은 자율 액션 정책과 감사 추적성, 롤백 자동화 기능을 표준화하여 운영 리스크를 최소화하고 있습니다. 예를 들어, 미국 대형 통신사의 경우 자율 remediation 도입

후 장애 대응 시간이 50% 이상 단축되었으나, 초기에는 감사로그 미비로 인한 문제도 보고된 바 있습니다. 이러한 사례는 자율화의 이점과 동시에 거버넌스 체계의 중요성을 시사합니다.

동시에 멀티모달(영상, 음성 등 비정형 데이터 포함) 통합 옴저버빌리티가 초기 단계로 진입합니다. 기존 텍스트, 로그, 메트릭, 트레이스 외에 영상 기반 장애 감지, 음성 기반 알림 등 새로운 데이터 소스가 관찰 가능성 체계에 편입되며, 이를 위해 데이터 처리 파이프라인, 저장소, 분석 엔진의 대대적 확장과 표준화가 요구됩니다. 예를 들어, 제조 현장에서는 영상 기반 장애 감지 시스템과 기존 로그/메트릭 분석 시스템이 통합되어 운영 효율성이 크게 향상되는 사례가 등장하고 있습니다.

2028년은 '거버넌스 정비 시점'으로, 자율 액션 정책, 감사로그, 롤백 자동화 등 자동화된 운영의 신뢰성과 투명성을 보장할 수 있는 거버넌스 표준을 반드시 수립해야 합니다. 이 기준이 미비할 경우, 자율화의 이점이 오히려 운영 리스크로 전환될 수 있으므로, IT 의사결정자는 기술 도입과 함께 거버넌스 체계의 성숙도를 반드시 병행 평가해야 합니다.



[그림 4] 2026~2028 시점별 변곡점 — Agentic GA · LLM 옴저버빌리티 융합 · 자율 Remediation|802

## 9.2 5대 전략적 합의 — BYO LLM+MCP·SKU 통합 콘솔·Agentic 단계적 도달·CSP-Native vs 독립·국산 골든 타임

본 절에서는 앞서 제시한 시장 전망과 기술 분석을 바탕으로, IT 의사결정자가 실질적으로 활용할 수 있는 5대 전략적 합의를 도출합니다. 각 합의는 장기적 도입 전략, RFP 작성, 벤더 선정, 사내 거버넌스 체계 수립 등 실무 단계에서 직접 연결될 수 있는 핵심 메시지로 구성되어 있습니다. 특히 한국 시장의 특수성과 글로벌 트렌드가 교차하는 지점에서, 어떠한 전략적 선택이 필요한지 명확히 제시함으로써, 조직의 경쟁력 강화와 리스크 최소화에 실질적인 도움을 드리고자 합니다.

### 9.2.1 합의 1·2·3 — 벤더 종속 회피의 새 표준(BYO LLM + MCP) / SKU 통합 콘솔 / Agentic 자율 운영의 단계적 도달

BYO LLM(Bring Your Own LLM)과 MCP(Model Context Protocol) 지원은 이제 벤더 종속 회피와 데이터 주권 확보를 위한 새로운 표준으로 자리잡고 있습니다. 기존에는 ServiceNow Now LLM, Dynatrace Azure OpenAI 등 특정 LLM 벤더에 종속된 구조가 도입 장벽으로 작용했으나, 최근에는 BYO LLM과 MCP를 지원하는 솔루션이 시장에서 경쟁력을 갖추고 있습니다. 실제로 국내 공공·금융권에서는 데이터 주권과 보안 요건을 충족하기 위해 BYO LLM 및 MCP 호환성 여부를 최우선 평가 항목으로 삼고 있으며, 이는 글로벌 기업의 도입 사례에서도 동일하게 확인되고 있습니다.

SKU(기능별 라이선스) 기반의 기능 분리 모델은 가격 정책 측면에서는 여전히 유지될 수 있으나, 실제 사용자 경험은 AI 기반 통합 콘솔로 급속히 수렴하고 있습니다. 장애 발생 시 여러 콘솔을 오가야 하는 구조는 운영 효율성과 사용자 만족도 측면에서 한계가 명확하며, 단일 콘솔에서 자연어 질의(NL)부터 진단, 조치, 검증까지 일관된 흐름을 제공하는 플랫폼이 시장 표준으로 자리잡고 있습니다. 예를 들어, 글로벌 제조사와 금융기관은 통합 콘솔 도입을 통해 트러블슈팅 속도와 데이터 일관성을 크게 개선한 바 있습니다. 평가 기준 역시 SKU 수가 아니라 트러블슈팅 단일 흐름의 완결성으로 이동하고 있습니다.

Agentic 워크플로 기반의 자율 운영은 2026년 Read+Reason(읽고 추론) 단계의 GA, 2027년 Act(실행) 단계의 보편화, 2028년 일부 도메인에서의 unsupervised act(무감독 자동 실행)로 단계적으로 진화할 전망입니다. 의사결정자는 도입 시점별로 자율화 단계의 깊이, 인간 승인 계

이팅, 감사 추적성 등 거버넌스 체계의 성숙도를 반드시 함께 평가해야 하며, 실제로 글로벌 선도 기업들은 자율화 단계별로 거버넌스 기준을 차별화하여 운영 리스크를 최소화하고 있습니다. 예를 들어, 2026년에는 인간 승인 중심의 자율화, 2027년 이후에는 부분적 무감독 자동화로 진화하는 사례가 점차 늘어나고 있습니다.

## 9.2.2 합의 4·5 — CSP-Native vs 독립 벤더 균형 동요 / 국산 오피버빌리티의 골든 타임

멀티클라우드 및 하이브리드 환경에서는 독립 벤더(Elastic, Splunk, IBM Instana 등)가 여전히 우위를 점하고 있으며, Microsoft Arc와 같은 예외를 제외하면 CSP-Native(클라우드 사업자 자체 솔루션)는 주로 단일 클라우드 또는 중간 규모 시장에서만 경쟁력을 갖습니다. 한국 시장의 경우 멀티클라우드(K-CSP + 외산 CSP ‘하’ 등급) 비중이 높아 독립 벤더 또는 국산 솔루션의 우위가 지속될 전망입니다. 실제로 국내 대형 금융기관과 공공기관은 멀티클라우드 환경에서 독립 벤더 솔루션을 선호하는 경향이 강하게 나타나고 있으며, 이는 데이터 주권, 규제 준수, 기술적 유연성 측면에서 중요한 전략적 선택이 되고 있습니다. 의사결정자는 사내 클라우드 구조와 시장 트렌드를 동시에 고려해 균형 잡힌 전략을 수립해야 하며, 단기적 비용 절감보다는 장기적 확장성과 규제 대응력을 우선시해야 합니다.

글로벌 5강 벤더가 한국 데이터센터, 한국어 sLLM, 국내 인증을 모두 갖추기 전 1~2년이 국산 오피버빌리티·AIOps 솔루션의 시장 선점 골든 타임입니다(8.4 참고). 이 기간 내에 BYO LLM, MCP, AIOps 거버넌스 등 5축을 동시 충족하는 국산 솔루션이 공공·금융 핵심 인프라 시장에서 표준 위치를 확보할 수 있습니다. 실제로 최근 2~3년간 국산 솔루션의 도입 사례가 증가하고 있으며, 정부 및 공공기관의 RFP에서도 국산 솔루션에 대한 우대 정책이 강화되고 있습니다. IT 의사결정자는 이 시간 창을 놓치지 않고, 국산 솔루션의 도입, 검토, PoC를 적극 추진해야 하며, 이를 통해 조직의 데이터 주권 확보와 규제 대응력을 극대화할 수 있습니다.

## 9.3 의사결정 체크리스트 — 7개 평가 항목으로 도입 결정하기

본 절에서는 앞서 분석한 모든 내용을 7개 평가 항목의 의사결정 체크리스트로 체계화하여 제공합니다. 이 체크리스트는 RFP 작성, Vendor Scorecard 평가, 법무·거버넌스 검토 등 실무 단계에서 즉시 복사·활용할 수 있도록 설계되었습니다. 각 항목은 OK/조건부/NG 3단계 판정

기준과 함께, 본 백서의 해당 장 인덱스를 명확히 제시하여 실질적인 도입 의사결정에 직접 활용할 수 있습니다. 실제로 국내외 대형 IT 조직에서는 이와 유사한 체크리스트를 활용하여 벤더 선정의 투명성과 객관성을 높이고 있습니다.

### 9.3.1 의사결정 체크리스트 7개 항목 — BYO LLM · MCP · Self-hosted · 자체 ML RCA · 인간 승인 게이팅 · 한국 거버넌스 · OTel GenAI

의사결정 체크리스트는 1~8장의 모든 분석을 압축한 7개 평가 항목으로 구성됩니다. 각 항목은 RFP 응답 양식·Vendor Scorecard·법무 검토 단계에 그대로 사용할 수 있도록 OK / 조건부 / NG의 3단계 판정 기준과 본 백서 해당 장 인덱스를 함께 제시합니다(S04, S27, S29, S41, S42, S43).

7개 체크리스트 × 평가 질문 × 판정 기준 × 본 백서 인덱스 표

#	체크리스트 항목	평가 질문(RFP 응답 강제 항목)	OK	조건부	NG	본 백서 인덱스
1	BYO LLM 지원	벤더가 한국어 sLLM(KT·네이버·SKT 등) 또는 정부 클라우드 LLM 결합을 1건 이상 사례로 제시할 수 있는가? 사용 LLM을 고객이 직접 지정·교체할 수 있는가?	공식 지원 + 결합 사례 1건 이상	로드맵상 지원 명시	벤더 고정 LLM 전용	3.3, 8.2
2	MCP 서버/클라이언트 호환성	MCP(Model Context Protocol) 서버·클라이언트 호환을 공식 문서에 명시했는가?	공식 지원 명시	3rd-party 어댑터 또는 로드맵	미지원·미 명시	3.4, 8.1.1, 8.2
3	Self-hosted + 자체 ML 기반 RCA	온프레미스/프라이빗 환경에서, 외부 LLM을 경유하지 않는 자체 ML 기반 RCA가 동작함을 데모로 입증할 수 있는가?	완전 지원·데모 가능	일부 기능 또는 로드맵	SaaS 전용 또는 LLM 의존	7.2, 8.1.3
4	Agentic 워크플로 인간 승인 게이팅·감사로그·롤백	에이전트의 액션 단계마다 인간 승인 게이팅, 감사로그 자동 기록, 롤백 자동화가 표준 기능으로 제공되는가?	3개 모두 표준 제공	일부 기능	미지원	9.1.3

5	한국 거버넌스 총족(N²SF·금융 AI 가이드라인·PIPA 4단계·CSAP)	한국 데이터센터 + 한국어 NL + 국정원 N²SF + 금융 AI 가이드라인 + PIPA 4단계 분류를 동시에 통과한 사례가 있는가?	동시 통과 사례 보유	일부 통과 또는 한국형 변형 제공	미해당	7.1, 7.2, 7.3, 8.3
6	OpenTelemetry GenAI Semantic Conventions 채택 로드맵	OTel GenAI 의미 규약 채택 로드맵을 12개월 이내로 공개했는가?	채택 완료 또는 12개월 이내 로드맵	로드맵 모호	미공개	9.1.2, 8.2
7	LLM 토큰 비용 직접 계약(벤더 재과금 금지)	벤더가 LLM 토큰 사용료를 자사 마진을 얹어 재과금하지 않고 고객이 LLM 공급자와 직접 계약하도록 허용하는가?	고객 직접 계약 허용	일부 옵션	벤더 재과금 강제	3.3, 9.2.1

표의 7개 항목은 다음과 같은 의사결정 시사점을 갖습니다. 첫째, 항목 1·2·5는 한국 시장 진입  
가능성의 게이팅 조건으로, 모두 NG이면 도입 자체를 보류하는 것이 합리적입니다(7장 4대 규제  
제약 분석과 정합). 둘째, 항목 3·4는 폐쇄망 운영 안정성과 거버넌스 리스크의 분산 조건이며, 항목  
6은 향후 12~24개월 내 표준화 흐름과의 정합성 확보를 위한 전향적 조건입니다. 셋째, 항목 7은  
다년 계약의 총비용을 통제하는 비용 거버넌스 조건으로, 벤더 락인 회피의 핵심 수단입니다(S04).  
7개 항목 모두에 대해 OK 응답을 받지 못하더라도, 조건부 응답의 구체적 로드맵과 책임자 명시를  
RFP에서 강제하면 도입 후 분쟁 리스크를 크게 줄일 수 있습니다.

### 9.3.2 체크리스트 사용법 — RFP · Vendor Scorecard · 법무 검토 3개 단계 적용

7개 체크리스트를 실무에 적용하는 표준 절차는 RFP 작성 → Vendor Scorecard 채점 → 법무·거  
버넌스 검토의 3단계로 분해됩니다. 본 항은 의사결정자가 본 백서를 닫고 즉시 사내 평가 위원회에  
가져갈 수 있는 3-step 가이드를 제공합니다(S04).

#### 7개 체크리스트 × 3개 적용 단계 × 적용 방법 매트릭스

#	체크리스트 항목	① RFP 작성 단계 — 필수 응답 항목으로 명시	② Vendor Scorecard — 채점 가중치	③ 법무·거버넌스 검토 — 통과 게이팅 항목
1	BYO LLM 지원	“한국어 sLLM·정부 클라우드 LLM 결합 사례 1건 이상 첨부” 강제	가중치 高(데이터 주권 직결)	개인정보 영향평가서·LLM 위탁 계약서 검토
2	MCP 서버/클라이언트 호환성	“MCP 공식 지원 명시 문서 첨부” 강제	가중치 中(향후 확장성)	위탁 처리 위탁사 변경 가능성 검토
3	Self-hosted + 자체 ML 기반 RCA	“외부 LLM 차단 환경 데모 진행” 강제	가중치 高(폐쇄망 필수)	망분리 가이드라인 적합성 검토(S41)
4	Agentic 인간 승인 게이팅·감사로그·롤백	“승인 게이팅 화면·감사로그 샘플 제출” 강제	가중치 高(운영 리스크)	내부통제·감사 위원회 정책 검토
5	한국 거버넌스 동시 통과	“N²SF·금융 AI 가이드라인·PIPA 4단계 통과 인증서 제출” 강제	게이팅(통과 못하면 NG)	법무 게이팅 항목(필수 통과)(S42, S43)
6	OTel GenAI 채택 로드맵	“12개월 이내 채택 로드맵 문서 제출” 강제	가중치 中(표준 적합성)	표준 변경 시 재계약 조항 검토
7	LLM 토큰 직접 계약	“고객이 LLM 공급자와 직접 계약 가능 여부 명시” 강제	가중치 中(비용 거버넌스)	3rd party 계약 분리 검토

매트릭스의 사용 방법은 다음과 같습니다. RFP 작성 단계에서는 7개 항목을 모두 “필수 응답 항목”으로 강제해 벤더가 무응답·일반화된 답변을 제출하지 못하도록 막고, Vendor Scorecard 채점 단계에서는 가중치 高 항목(1·3·4) 중 하나라도 NG일 경우 평가 자체에서 탈락시키는 게이팅 룰을 적용합니다. 법무·거버넌스 검토 단계에서는 항목 5(한국 거버넌스 동시 통과)를 게이팅으로 두고, 통과하지 못한 벤더는 정식 계약 단계로 진행하지 않습니다(S42, S43). 이 3단계 가이드는 본 백서 1~8장의 모든 분석을 한 장의 평가 양식으로 압축한 결과이며, 실제 RFP 양식·Vendor Scorecard 시트로 그대로 옮겨 사용할 수 있도록 설계되었습니다.

본 절을 닫는 권고는 다음 한 줄로 요약됩니다 — **“이번 분기 안에 사내 평가 위원과 본 백서를 공유하고, 7개 항목을 RFP 양식에 즉시 추가하라.”** 본 백서의 모든 분석은 이 한 줄의 액션을 위해 작성되었으며, 7개 체크리스트가 RFP에 명시되는 순간 한국 시장 AI 오피저빌리티 도입의 의사결정 품질이 측정 가능한 기준 위에 올라가게 됩니다.

## 부록: 전체 출처 목록

본 부록은 본 백서의 모든 분석, 비교, 표, 인용의 근거가 되는 공식 외부 출처 목록을 체계적으로 제시합니다. 각 출처는 벤더 공식 문서, 글로벌 리서치 기관 보고서, 제품 릴리스 노트, 기술 블로그, 국내외 주요 언론 기사, 오픈소스 프로젝트 문서 등으로 구성되어 있습니다. 출처 목록은 백서의 신뢰성과 투명성을 보증하며, 독자가 각 장의 주요 주장과 데이터의 근거를 직접 확인할 수 있도록 설계되었습니다. 인용 표기 방식은 S01~S78의 식별자를 활용하였으며, 각 장의 메타데이터 블록과 연계되어 자동으로 참조됩니다. 아래에 전체 출처 목록을 기술합니다.

S01:<https://www.gartner.com/en/documents/6688834>(Gartner MQ Observability 2025)

Gartner의 Magic Quadrant for Observability Platforms 2025년 판 공식 보고서입니다. 글로벌 옴저버빌리티 시장의 리더, 비저너리, 니치 플레이어, 챌린저 구분과 각 벤더의 강점, 약점, 시장 변화 요인을 종합적으로 분석합니다. 본 백서의 글로벌 벤더 비교, 시장 변천사, 평가 기준 진화 등 다수의 핵심 논거에 직접적인 근거로 활용되었습니다.

S02:<https://www.gartner.com/en/documents/6755734>(Gartner Hype Cycle Monitoring 2025)

Gartner Hype Cycle for Monitoring and Observability 2025년 판 공식 문서로, GenAI, Agentic AI, AIOps, LLM Observability 등 신기술의 시장 성숙도와 기대치 변화를 시계열로 분석합니다. Agentic AI가 정점(Peak)에 도달하고 GenAI가 환멸기(Trough)에 진입한 사실 등, 기술 도입 시점 결정의 근거로 사용되었습니다.

S03:<https://www.gartner.com/en/newsroom/press-releases/2025-08-05-gartner-hype-cycle-identifies-top-ai-innovations-in-2025>

Gartner 공식 보도자료로, 2025년 Hype Cycle에서 AI Agents, Generative AI, AIOps 등 주요 기술의 위치와 시장 변동성을 요약합니다. 2028년 이후 자율 remediation, 멀티모달 옴저버빌리티 등 미래 방향성 근거로 인용되었습니다.

S04:<https://www.forrester.com/report/the-forrester-wave-tm-aiops-platforms-q2-2025/RES182220>

Forrester Wave™: AIOps Platforms Q2 2025 공식 리포트로, Dynatrace, ScienceLogic, Datadog 등 글로벌 리더의 평가 기준, Agentic AI 통합 깊이, GenAI 평준화 현상 등을

분석합니다. 본 백서의 의사결정 체크리스트, 전략적 함의, 한국 시장 진입 조건 등 전반에 걸쳐 핵심 인용으로 사용되었습니다.

S05:<https://my.idc.com/getdoc.jsp?containerId=US54116226>(IDC AIOps 2026)

IDC MarketScape: Worldwide AIOps Platforms 2026 평가 보고서로, 17개 벤더의 전략, 제품 기능, 시장 리더십을 비교합니다. New Relic의 리더 등재, 평가축 진화 등 IDC 기준의 변화 근거로 활용되었습니다.

S06:<https://my.idc.com/getdoc.jsp?containerId=US53004325>(IDC Observability 2025)

IDC MarketScape: Worldwide Observability Platforms 2025 공식 리포트입니다. 26개 벤더의 데이터 플레인 폭, 분석 깊이, 리더·메이저 플레이어 구분 등 평가 결과가 포함되어 있습니다. Elastic의 리더 등재, HPE OpsRamp의 Major Player 등재 등 주요 데이터의 출처입니다.

S07:<https://www.datadoghq.com/about/latest-news/press-releases/datadog-launches-bits-ai-sre-agent-to-resolve-incidents-faster/>

Datadog 공식 보도자료로, Bits AI SRE Agent의 글로벌 GA(General Availability) 발표, 주요 기능, MTTR 단축 효과, AI 기반 트러블슈팅 자동화 사례 등이 상세히 기술되어 있습니다. AI 기반 트러블슈팅 즉각 종결, KPI 변화, 한국 시장 도입 한계 등 다수의 논거에 활용되었습니다.

S08:<https://www.datadoghq.com/product/ai/bits-ai-sre/>

Datadog Bits AI SRE 제품 공식 문서입니다. AI SRE Agent의 기능 상세, conclusive investigation 평균 시간 단축, NL 질의→결론까지의 흐름, 평가 지표 등이 포함되어 있습니다. Time-to-Resolution KPI 정의, AI 도입 전후 비교 등 실증 데이터의 근거입니다.

S09:<https://www.dynatrace.com/news/blog/announcing-general-availability-of-davis-copilot-your-new-ai-assistant/>

Dynatrace Davis CoPilot GA 공식 블로그로, Davis CoPilot의 기능, Azure OpenAI 고정 모델 구조, Agentic 워크플로, 평가 기준 변화 등 핵심 내용을 제공합니다. 고정 LLM 진영의 대표 사례로 인용되었습니다.

S10:<https://www.dynatrace.com/news/press-release/forrester-wave-aiops-platforms-q2-2025/>

Dynatrace의 Forrester Wave AIOps Platforms Q2 2025 등재 공식 보도자료입니다. 평가 항목, 만점 획득 내역, 전략적 강점 등 리더십 근거로 활용되었습니다.

S11:<https://docs.newrelic.com/whats-new/2025/06/whats-new-06-04-nrai/>

New Relic NRAI(Neural Relic AI) 공식 업데이트 문서로, OpenAI 기반 AI 기능, GA 시점, 과금 구조 등 SaaS 진영의 AI 기능 비교 근거입니다.

S12:<https://newrelic.com/blog/news/new-relic-named-a-leader-in-the-2026-idc-marketscape-for-worldwide-aiops>

New Relic의 IDC MarketScape AIOps 2026 리더 등재 공식 블로그로, 평가 기준, 리더십 근거, 시장 전략 등 IDC 보고서의 주요 내용을 요약합니다.

S13:[https://www.splunk.com/en\\_us/blog/conf-splunklive/supercharge-insights-with-generative-ai-splunk-ai-assistant-for-spl-is-now-ga.html](https://www.splunk.com/en_us/blog/conf-splunklive/supercharge-insights-with-generative-ai-splunk-ai-assistant-for-spl-is-now-ga.html)

Splunk AI Assistant for SPL 공식 GA 발표 블로그입니다. 자체 호스팅 OSS Transformer 기반, NL→쿼리 흐름, 한국 진입 적합도 등 비교 분석에 활용되었습니다.

S14:[https://www.splunk.com/en\\_us/blog/artificial-intelligence/scaling-anomaly-detection-with-mltk-5-5.html](https://www.splunk.com/en_us/blog/artificial-intelligence/scaling-anomaly-detection-with-mltk-5-5.html)

Splunk MLTK(머신러닝 툴킷) 5.5 이상 탐지 기능 공식 블로그로, 자체 ML 기반 이상 탐지의 즉시 도입 가능성, 한국 규제 환경 적합성의 주요 근거입니다.

S15:[https://www.splunk.com/en\\_us/newsroom/press-releases/2025/cisco-data-fabric-transforms-machine-data-into-ai-ready-intelligence.html](https://www.splunk.com/en_us/newsroom/press-releases/2025/cisco-data-fabric-transforms-machine-data-into-ai-ready-intelligence.html)

Splunk(Cisco) Data Fabric 공식 보도자료로, 머신데이터의 AI 활용, 통합 데이터 플레인 등 차세대 오피버빌리티 인프라의 방향성을 제시합니다.

S16:<https://www.elastic.co/observability/ai-assistant>

Elastic AI Assistant for Observability 공식 문서로, BYO LLM 지원, Self-hosted 옵션, 한국형 변형 필요성 등 엔터프라이즈 진영의 대표 사례입니다.

S17:<https://www.businesswire.com/news/home/20251114859815/en/Elastic-Named-a-Leader-in-2025-IDC-MarketScape-for-Worldwide-Observability-Platforms>

Elastic의 IDC MarketScape Observability 2025 리더 등재 공식 보도자료로, 평가 기준, 시장 전략, 리더십 근거 등 IDC 기준의 변화에 활용됩니다.

S18:<https://grafana.com/docs/grafana-cloud/alerting-and-irm/sift/>

Grafana Sift 공식 문서로, LLM 플러그인, BYO LLM, Self-hosted 구조 등 엔터프라이즈 진영의 AI 기능 비교 근거입니다.

S19:<https://www.servicenow.com/products/itom.html>

ServiceNow ITOM 공식 제품 페이지로, Now Assist for ITOM의 AI 기능, LLM 고정 구조, 한국 시장 적합도 등 ITSM 진영의 대표 사례입니다.

S20:<https://www.bmc.com/it-solutions/bmc-helixgpt.html>

BMC HelixGPT 공식 제품 문서로, BYO LLM 지원, Self-hosted 옵션, 한국형 변형 필요성 등 엔터프라이즈 진영의 대표 사례입니다.

S21:<https://www.ibm.com/products/instana>

IBM Instana 공식 제품 페이지로, watsonx 결합, BYO LLM, Self-hosted 지원, 한국 진입 적합도 등 엔터프라이즈 진영의 대표 사례입니다.

S22:<https://aws.amazon.com/about-aws/whats-new/2025/06/ga-accelerate-troubleshooting-amazon-cloudwatch-investigations/>

AWS CloudWatch Investigations 2025-06 GA 공식 발표로, AWS DevOps Agent, AI 기반 트러블슈팅, agent-second 과금 구조 등 CSP 진영의 AI 전략 비교 근거입니다.

S23:<https://aws.amazon.com/about-aws/whats-new/2026/03/aws-devops-agent-generally-available/>

AWS DevOps Agent 2026-03 GA 공식 발표로, agent-second 과금, 글로벌 GA 시점, 한국 리전 가용성 등 CSP 진영의 AI 전략 비교 근거입니다.

S24:<https://techcommunity.microsoft.com/blog/appsonazureblog/announcing-general-availability-for-the-azure-sre-agent/4500682>

Microsoft Azure SRE Agent GA 공식 블로그로, AAU 과금, 멀티클라우드 지원, Azure Arc 연계 등 CSP 진영의 AI 전략 비교 근거입니다.

S25:<https://cloud.google.com/products/gemini/cloud-assist>

Google Cloud Gemini Cloud Assist 공식 문서로, Preview/GA 현황, AI 기능, 한국 리전 가용성 등 CSP 진영의 AI 전략 비교 근거입니다.

S26:<https://blogs.oracle.com/observability/latest-enhancements-we-rolled-early-2025-in-oci-ops-insights>

Oracle Cloud Ops Insights 2025 공식 블로그로, OCI GenAI, Agent Hub, DB 워크로드 특화 AI 기능 등 CSP 진영의 AI 전략 비교 근거입니다.

S27:<https://www.anthropic.com/news/model-context-protocol>

Anthropic Model Context Protocol(MCP) 공식 발표로, 오피저버빌리티 도구의 LLM 에이전트 표준 입력 채널화, MCP 표준화 타임라인 등 차세대 AI 오피저버빌리티의 핵심 근거입니다.

S28:<https://modelcontextprotocol.io/>

Model Context Protocol 공식 사이트로, MCP의 사양, 지원 벤더, 표준화 현황 등 MCP 기반 에이전트 오피저버빌리티의 기술 근거입니다.

S29:<https://opentelemetry.io/docs/specs/semconv/gen-ai/>

OpenTelemetry GenAI Semantic Conventions 공식 문서로, LLM Observability, MCP 연동, 프롬프트/응답/토큰/비용/평가 등 LLM 트레이스 표준화의 근거입니다.

S30:<https://www.solarwinds.com/database-performance-analyzer/use-cases/ai-database-monitoring>

SolarWinds Database Performance Analyzer AI 기능 공식 문서로, AI Query Assist, ML 이상 탐지, DB 모니터링 AI 분화 매트릭스의 근거입니다.

S31:<https://blog.quest.com/product-post/alarms-ai-is-here-in-foglight-for-databases-v-8/>

Quest Foglight 8/8.1 AI Alarms 공식 블로그로, 생성형 LLM 탑재, 모델 비공개, DB 모니터링 AI 분화 매트릭스의 근거입니다.

S32:<https://docs.percona.com/percona-monitoring-and-management/3/release-notes/3.5.0.html>

Percona PMM 3.5.0 공식 릴리스 노트로, MCP 외부 LLM 연동, OSS+BYO LLM+Self-managed 구조, 한국 금융권 적합도 근거입니다.

S33:<https://www.datadoghq.com/product/database-monitoring/>

Datadog Database Monitoring 공식 문서로, 쿼리 회귀 자동 탐지, Bits AI Dev Agent, DB 모니터링 AI 분화 매트릭스의 근거입니다.

S34:<https://www.pagerduty.com/blog/product/product-launch-2025-h2/>

PagerDuty SRE Agent 공식 제품 블로그로, 외부 LLM 고정, 한국 도입 불가 기능군의 대표 사례입니다.

S35:<https://incident.io/ai-sre>

incident.io AI SRE 공식 문서로, AI 기반 인시던트 대응, SRE 자동화, 글로벌 AI 모니터링 기능 비교 근거입니다.

S36:<https://jennifersoft.com/ko/blog/tech/2025-12-22-jenniferai-jenniferinsights/>

제니퍼소프트 JenniferAI 공식 기술 블로그로, 브라우저 기반 LLM, 클라이언트 사이드 추론, 폐쇄망 대응 방식의 대표 사례입니다.

S37:<https://whatap.io/ko/blog/ai-business-observability-it-operations-2025>

와탭 AI 오피저버빌리티 공식 블로그로, OpsLake, 데이터 레이크 기반 AI, 국내 벤더 AI 차별화 근거입니다.

S38:<https://www.seoulfn.com/news/articleView.html?idxno=618704>(엑셈 eX-  
emble)

엑셈 eXemble 관련 언론 기사로, 자체 LLMOps, MCP 연동, 국내 벤더 AI 차별화 근거입니다.

S39:<https://www.pinpointnews.co.kr/news/articleView.html?idxno=327086>(Zenius)

브레인즈컴퍼니 Zenius AI Agent 관련 언론 기사로, RAG 기반 AI, 폐쇄망 대응, 국내 벤더 AI 차별화 근거입니다.

S40:<https://ddaily.co.kr/page/view/2025041610500725990>(LG CNS)

LG CNS AIOps 관련 보도자료로, Agentic AI 생태계, LLMOps+MLOps 통합, 거버넌스 내재화 등 SI 진영의 AI 차별화 근거입니다.

S41:<https://www.fsc.go.kr/no010101/86080>(금융위 망분리)

금융위원회 망분리 정책 관련 공식 문서로, 망분리 단계적 허용, SaaS-AI 도입 규제, 한국 시장 구조적 제약 근거입니다.

S42:<https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=BS074&mCode=C020010000&nttId=11410>(PIPC 생성형 AI)

개인정보보호위원회(PIPC) 생성형 AI 안내서 공식 문서로, PIPA 4단계 분류, 텔레메트리 PII 외부 LLM 전송 규제, 한국 시장 구조적 제약 근거입니다.

S43:<https://www.etnews.com/20250317000203>(국정원 N<sup>2</sup>SF)

전자신문 국정원 N<sup>2</sup>SF 관련 기사로, 비금융 공공의 업무망 중요도 차등 적용, 외산 SaaS-AI 차단 근거입니다.

S44:<https://www.mordorintelligence.com/industry-reports/observability-market>

Mordor Intelligence 옵저버빌리티 시장 리포트로, 글로벌 시장 규모, 성장률, 주요 트렌드 등 시장 분석의 보조 근거입니다.

S45:<https://www.fortunebusinessinsights.com/aiops-market-109984>

Fortune Business Insights AIOps 시장 리포트로, AIOps 시장 성장, 벤더별 점유율 등 시장 분석의 보조 근거입니다.

S46:<https://www.researchnester.com/reports/observability-tools-and-platforms-market/8139>

Research Nester 옵저버빌리티 도구 시장 리포트로, 시장 규모, 성장 요인, 경쟁 구도 등 시장 분석의 보조 근거입니다.

S47:<https://www.networkworld.com/article/4032218/in-crowded-observability-market-gartner-calls-out-ai-capabilities-cost-optimization-devops-integration.html>

NetworkWorld Gartner MQ Observability 2025 분석 기사로, 리더 8사, 평가 기준 변화, AI 기능 비중 등 시장 동향의 보조 근거입니다.

S48:<https://www.comworld.co.kr/news/articleView.html?idxno=51321>(엑셈)

엑셈 AI 옵저버빌리티 관련 국내 기사로, 자체 LLMOps, MCP 연동, 국내 벤더 AI 차별화 근거입니다.

S49:<https://www.newstheai.com/news/articleView.html?idxno=9871>(와탭)

와탭 AI 옵저버빌리티 관련 국내 기사로, OpsLake, 데이터 레이크 기반 AI, 국내 벤더 AI 차별화 근거입니다.

S50:<https://www.dailysecu.com/news/articleView.html?idxno=200759>(Zenius)

브레인즈컴퍼니 Zenius AI Agent 관련 국내 기사로, RAG 기반 AI, 폐쇄망 대응, 국내 벤더 AI 차별화 근거입니다.

S51:<https://www.globenewswire.com/news-release/2025/12/09/3202429/0/en/Application-Performance-Monitoring-Market-Set-for-Strong-Growth-to-USD-25-24-Billion-by-2033-Driven-by-Digital-Transformation-and-Cloud-Native-Complexity-Research-by-SNS-Insider.html>

SNS Insider APM 시장 성장 리포트로, 글로벌 APM 시장 규모 및 성장률 등 시장 분석의 보조 근거입니다.

S52:<https://docs.newrelic.com/docs/agentic-ai/mcp/overview/>

New Relic MCP 공식 문서로, MCP 표준화, Agentic AI 연동, 옴저버빌리티 도구의 LLM 에이전트 채널화 근거입니다.

S53:[https://docs.datadoghq.com/llm\\_observability/monitoring/cost/](https://docs.datadoghq.com/llm_observability/monitoring/cost/)

Datadog LLM Observability 공식 문서로, LLM 트레이스, 프롬프트/응답/토큰/비용/평가, 단일 데이터 플레인 등 LLM Observability의 대표 사례입니다.

S54:<https://www.techtarget.com/searchitoperations/news/366630519/Under-Cisco-Splunk-AI-roadmap-tees-up-pricing-overhaul>

TechTarget Splunk AI 로드맵 및 가격 구조 변화 기사로, AI 기능별 과금, SKU 통합, 글로벌 벤더 가격 정책 변화 근거입니다.

S55:<https://www.datadoghq.com/blog/dash-2025-new-feature-roundup-keynote/>

Datadog Dash 2025 신기능 공식 블로그로, Bits AI SRE, Monitoring as Code, NL 질의, 콘솔 통합 등 최신 기능의 근거입니다.

S56:<https://learn.microsoft.com/en-us/azure/sre-agent/pricing-billing>

Microsoft Azure SRE Agent 공식 과금 문서로, AAU 과금 구조, agent-second 단위 등 CSP 진영의 과금 모델 비교 근거입니다.

S57:<https://aws.amazon.com/devops-agent/pricing/>

AWS DevOps Agent 공식 과금 문서로, agent-second 단위 과금, 월 추정 비용 시뮬레이션 등 CSP 진영의 과금 모델 비교 근거입니다.

S58:<https://www.computerweekly.com/news/366627660/Datadog-doubles-down-on-APAC-targets-faster-growth>

ComputerWeekly Datadog APAC 전략 기사로, 글로벌 벤더의 아시아 시장 진입 전략, 한국 시장 진입 한계 등 분석의 보조 근거입니다.

S59:<https://code.visualstudio.com/blogs/2025/04/07/agentMode>

VS Code Agent Mode 공식 블로그로, MCP 기반 LLM Agent 연동, 에이전틱 옴저버빌리티 허브의 대표 사례입니다.

S60:<https://www.samsungsds.com/kr/insights/2025-dx-prospects.html>

삼성SDS 2025 DX 전망 공식 문서로, Brightics AI, Fabrix, Brity Copilot 등 SI 진영의 AI 차별화 근거입니다.

S61:<https://www.lawtimes.co.kr/news/articleView.html?idxno=215867>

법률신문 망분리·SaaS 규제 관련 기사로, 금융위·국정원 정책 변화, 한국 시장 구조적 제약의 보조 근거입니다.

S62:<https://www.atlassian.com/software/jira/ai>(Atlassian Rovo)

Atlassian Rovo AI 공식 제품 페이지로, 외부 LLM 고정, 한국 도입 불가 기능군의 대표 사례입니다.

S63:<https://docs.firehydrant.com/docs/ai-powered-incident-management>

FireHydrant AI-Powered Incident Management 공식 문서로, AI 기반 인시던트 대응, 글로벌 AI 모니터링 기능 비교 근거입니다.

S64:<https://www.pagerduty.com/blog/product/product-launch-enhancements-to-pagerduty-operations-cloud-2025-h1/>

PagerDuty Operations Cloud 2025 H1 공식 블로그로, AI 기반 인시던트 자동화, 글로벌 AI 모니터링 기능 비교 근거입니다.

S65:<https://newsroom.cisco.com/c/r/newsroom/en/us/a/y2024/m03/cisco-completes-acquisition-of-splunk.html>

Cisco의 Splunk 인수 공식 보도자료로, Splunk AI 전략, 글로벌 벤더 시장 재편 근거입니다.

S66:<https://docs.dynatrace.com/docs/discover-dynatrace/platform/davis-ai/copilot/copilot-data-privacy>

Dynatrace Davis CoPilot 데이터 프라이버시 공식 문서로, Azure OpenAI 고정, BYO LLM 부재, 한국 도입 한계 근거입니다.

S67:<https://docs.dynatrace.com/docs/discover-dynatrace/platform/davis-ai>

Dynatrace Davis AI 공식 문서로, Causal RCA, Agentic 워크플로, 단일 흐름 4계층 등 글로벌 벤더 AI 기능 비교 근거입니다.

S68:<https://www.dynatrace.com/news/press-release/dynatrace-expanding-davis-hypermodal-ai/>

Dynatrace Davis Hypermodal AI 공식 보도자료로, Causal AI, Agentic AI 결합, 글로벌

벤더 AI 차별화 근거입니다.

S69:<https://docs.newrelic.com/docs/security/security-privacy/compliance/fedramp-compliant-endpoints/>

New Relic FedRAMP 준수 공식 문서로, Self-hosted 옵션 부재, 한국 폐쇄망 적합도 한계 근거입니다.

S70:<https://www.ibm.com/watsonx>

IBM watsonx 공식 제품 페이지로, BYO LLM, 자체 호스팅, 한국 진입 적합도 등 엔터프라이즈 진영의 대표 사례입니다.

S71:<https://www.ibm.com/granite>

IBM Granite Foundation Model 공식 문서로, watsonx LLM 선택 옵션, 엔터프라이즈 AI 플랫폼 근거입니다.

S72:<https://grafana.com/grafana/plugins/grafana-llm-app/>

Grafana LLM 플러그인 공식 문서로, BYO LLM, Self-hosted, 한국형 변형 필요성 등 엔터프라이즈 진영의 대표 사례입니다.

S73:<https://github.com/grafana/mcp-grafana>

Grafana MCP 연동 공식 GitHub 저장소로, MCP 표준화, 에이전틱 오피저빌리티 허브 구현 근거입니다.

S74:<https://azure.microsoft.com/en-us/products/sre-agent/>

Azure SRE Agent 공식 제품 페이지로, 멀티클라우드 지원, 단일 콘솔, CSP 진영의 AI 전략 비교 근거입니다.

S75:<https://www.etnews.com/20251212000235>(CSAP 보도)

전자신문 CSAP 제도 변화 관련 기사로, 보안적합성 검증, 외산 SaaS-AI 차단 지속 등 한국 시장 구조적 제약 근거입니다.

S76:<https://www.oracle.com/autonomous-database/features/>

Oracle Autonomous Database 공식 기능 문서로, Self-tuning, Select AI, Ask EM, DB 모니터링 AI 분화 매트릭스의 근거입니다.

S77:<https://www.mongodb.com/docs/atlas/performance-advisor/access-with-mcp/>

MongoDB Atlas Performance Advisor + MCP 공식 문서로, MCP Server, BYO LLM,

클라우드 매니지드 진영의 대표 사례입니다.

S78:<https://blog.quest.com/product-post/stop-reacting-start-preventing-how-foglight-8-1-puts-you-ahead-of-database-problems/>

Quest Foglight 8.1 AI 기능 공식 블로그로, 생성형 LLM 비공개, DB 모니터링 AI 분화 매트릭스의 근거입니다.

---

## Appendix

### References

1. AWS. (2025). “CloudWatch Investigations GA.”<https://aws.amazon.com/about-aws/whats-new/2025/06/ga-accelerate-troubleshooting-amazon-cloudwatch-investigations/>
2. AWS. (2025). “GA: Accelerate Troubleshooting with Amazon CloudWatch Investigations.”<https://aws.amazon.com/about-aws/whats-new/2025/06/ga-accelerate-troubleshooting-amazon-cloudwatch-investigations/>
3. AWS. (2026). “AWS DevOps Agent GA.”<https://aws.amazon.com/about-aws/whats-new/2026/03/aws-devops-agent-generally-available/>
4. AWS. (2026). “AWS DevOps Agent”.<https://aws.amazon.com/about-aws/whats-new/2026/03/aws-devops-agent-generally-available/>
5. Anthropic. (2024). “Introducing the Model Context Protocol (MCP).”<https://www.anthropic.com/news/model-context-protocol>
6. Anthropic. (2024). “Model Context Protocol(MCP)”.<https://www.anthropic.com/news/model-context-protocol>
7. Anthropic. (2024). “Model Context Protocol.”<https://www.anthropic.com/news/model-context-protocol>
8. BMC Software, Inc. (2025). “BMC HelixGPT.”<https://www.bmc.com/it-solutions/bmc-helixgpt.html>

9. BMC. (2025). “BMC HelixGPT” .<https://www.bmc.com/it-solutions/bmc-helix-gpt.html>
10. Brainscompany. (2025). “Zenius AI Agent” .<https://www.pinpointnews.co.kr/news/articleView.html?idxno=327086>
11. Comworld. (2025). “엑셈 AI 오피서빌리티” .<https://www.comworld.co.kr/news/articleView.html?idxno=51321>
12. Dailysecu. (2025). “Zenius AI Agent” .<https://www.dailysecu.com/news/articleView.html?idxno=200759>
13. Datadog, Inc. (2025). “Datadog Launches Bits AI SRE Agent to Resolve Incidents Faster.”<https://www.datadoghq.com/about/latest-news/press-releases/datadog-launches-bits-ai-sre-agent-to-resolve-incidents-faster/>
14. Datadog. (2025). “Bits AI SRE.”<https://www.datadoghq.com/product/ai/bits-ai-sre/>
15. Datadog. (2025). “DASH 2025 New Feature Roundup.”<https://www.datadoghq.com/blog/dash-2025-new-feature-roundup-keynote/>
16. Datadog. (2025). “Database Monitoring” .<https://www.datadoghq.com/product/database-monitoring/>
17. Datadog. (2025). “Datadog Bits AI SRE” .<https://www.datadoghq.com/about/latest-news/press-releases/datadog-launches-bits-ai-sre-agent-to-resolve-incidents-faster/>
18. Datadog. (2025). “Datadog Launches Bits AI SRE Agent to Resolve Incidents Faster.”<https://www.datadoghq.com/about/latest-news/press-releases/datadog-launches-bits-ai-sre-agent-to-resolve-incidents-faster/>
19. Datadog. (2025). “LLM Observability Monitoring Cost.”[https://docs.datadoghq.com/llm\\_observability/monitoring/cost/](https://docs.datadoghq.com/llm_observability/monitoring/cost/)
20. Datadog. (2025). “LLM Observability” .[https://docs.datadoghq.com/llm\\_observability/monitoring/cost/](https://docs.datadoghq.com/llm_observability/monitoring/cost/)
21. Dynatrace Docs. (2025). “Copilot Data Privacy.”<https://docs.dynatrace.com/docs/discover-dynatrace/platform/davis-ai/copilot/copilot-data-privacy>

22. Dynatrace News. (2025). “Dynatrace named a Leader in The Forrester Wave™: AIOps Platforms, Q2 2025”. <https://www.dynatrace.com/news/press-release/forrester-wave-aiops-platforms-q2-2025/>
23. Dynatrace, Inc. (2024). “Announcing General Availability of Davis CoPilot, Your New AI Assistant.” <https://www.dynatrace.com/news/blog/announcing-general-availability-of-davis-copilot-your-new-ai-assistant/>
24. Dynatrace. (2025). “Dynatrace expanding Davis hypermodal AI.” <https://www.dynatrace.com/news/press-release/dynatrace-expanding-davis-hypermodal-ai/>
25. Elastic N.V. (2025). “AI Assistant for Observability.” <https://www.elastic.co/observability/ai-assistant>
26. Elastic. (2025). “Elastic Named a Leader in 2025 IDC MarketScape for Worldwide Observability Platforms” <https://www.businesswire.com/news/home/20251114859815/en/Elastic-Named-a-Leader-in-2025-IDC-MarketScape-for-Worldwide-Observability-Platforms>
27. Elastic. (2025). “Elastic Observability AI Assistant” <https://www.elastic.co/observability/ai-assistant>
28. Forrester. (2025). “The Forrester Wave™: AIOps Platforms, Q2 2025” <https://www.forrester.com/report/the-forrester-wave-tm-aiops-platforms-q2-2025/RES182220>
29. Forrester. (2025). “The Forrester Wave™: AIOps Platforms, Q2 2025.” <https://www.forrester.com/report/the-forrester-wave-tm-aiops-platforms-q2-2025/RES182220>
30. Gartner Press Release. (2025). <https://www.gartner.com/en/newsroom/press-releases/2025-08-05-gartner-hype-cycle-identifies-top-ai-innovations-in-2025>
31. Gartner. (2025). “Gartner Hype Cycle identifies top AI innovations in 2025.” <https://www.gartner.com/en/newsroom/press-releases/2025-08-05-gartner-hype-cycle-identifies-top-ai-innovations-in-2025>

32. Gartner. (2025). “Hype Cycle for Monitoring and Observability 2025”.<https://www.gartner.com/en/documents/6755734>
33. Gartner. (2025). “Hype Cycle for Monitoring and Observability.”<https://www.gartner.com/en/documents/6755734>
34. Gartner. (2025). “Magic Quadrant for Observability Platforms 2025”.<https://www.gartner.com/en/documents/6688834>
35. Gartner. (2025). “Magic Quadrant for Observability Platforms.”<https://www.gartner.com/en/documents/6688834>
36. Google Cloud. (2025). “Gemini Cloud Assist.”<https://cloud.google.com/products/gemini/cloud-assist>
37. Grafana Labs (2025). “Grafana LLM App.”<https://grafana.com/grafana/plugins/grafana-llm-app/>
38. Grafana Labs. (2025). “Grafana LLM 플러그인”.<https://grafana.com/grafana/plugins/grafana-llm-app/>
39. IBM Corp. (2025). “IBM Instana.”<https://www.ibm.com/products/instana>
40. IBM Corp. (2025). “IBM watsonx.”<https://www.ibm.com/watsonx>
41. IBM. (2025). “watsonx”.<https://www.ibm.com/watsonx>
42. IDC. (2025). “MarketScape: Worldwide Observability Platforms 2025”.<https://my.idc.com/getdoc.jsp?containerId=US53004325>
43. IDC. (2026). “MarketScape: Worldwide AIOps Platforms 2026”.<https://my.idc.com/getdoc.jsp?containerId=US54116226>
44. Jennifersoft. (2025). “JenniferAI: 브라우저 기반 LLM”.<https://jennifersoft.com/ko/blog/tech/2025-12-22-jenniferai-jenniferinsights/>
45. LG CNS. (2025). “AIOps 플랫폼 구축 사례”.<https://ddaily.co.kr/page/view/2025041610500725990>
46. Microsoft. (2025). “Azure Arc.”<https://azure.microsoft.com/en-us/products/sre-agent/>
47. Microsoft. (2025). “Azure SRE Agent GA.”<https://techcommunity.microsoft.com/blog/appsonazureblog/announcing-general-availability-for-the-azure>

-sre-agent/4500682

48. ModelContextProtocol.io. (2025). “MCP Specification.”<https://modelcontextprotocol.io/>
49. MongoDB. (2025). “Accessing Performance Advisor with MCP” .<https://www.mongodb.com/docs/atlas/performance-advisor/access-with-mcp/>
50. NetworkWorld. (2025). “In crowded observability market, Gartner calls out AI capabilities, cost optimization, DevOps integration” .<https://www.networkworld.com/article/4032218/in-crowded-observability-market-gartner-calls-out-ai-capabilities-cost-optimization-devops-integration.html>
51. New Relic Docs. (2025). “FedRAMP Compliant Endpoints.”<https://docs.newrelic.com/docs/security/security-privacy/compliance/fedramp-compliant-endpoints/>
52. New Relic, Inc. (2025). “What’s New: NRAI.”<https://docs.newrelic.com/whats-new/2025/06/whats-new-06-04-nrai/>
53. New Relic. (2026). “New Relic Named a Leader in the 2026 IDC MarketScape for Worldwide AIOps” .<https://newrelic.com/blog/news/new-relic-named-a-leader-in-the-2026-idc-marketscape-for-worldwide-aiops>
54. NewstheAI. (2025). “와탭 AI 옵저버빌리티” .<https://www.newstheai.com/news/articleView.html?idxno=9871>
55. OpenTelemetry. (2024). “Semantic Conventions for GenAI.”<https://opentelemetry.io/docs/specs/semconv/gen-ai/>
56. OpenTelemetry. (2025). “GenAI Semantic Conventions” .<https://opentelemetry.io/docs/specs/semconv/gen-ai/>
57. OpenTelemetry. (2025). “GenAI Semantic Conventions.”<https://opentelemetry.io/docs/specs/semconv/gen-ai/>
58. Oracle. (2025). “Autonomous Database Features” .<https://www.oracle.com/autonomous-database/features/>
59. Oracle. (2025). “OCI Ops Insights Enhancements.”<https://blogs.oracle.com/observability/latest-enhancements-we-rolled-early-2025-in-oci-ops-ins>

## ights

60. PIPC. (2025). “생성형 AI 개발·활용 안내서”. <https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=BS074&mCode=C020010000&nttId=11410>
61. PagerDuty. (2025). “PagerDuty SRE Agent”. <https://www.pagerduty.com/blog/product/product-launch-2025-h2/>
62. Percona. (2025). “Percona Monitoring and Management 3.5.0 Release Notes”. <https://docs.percona.com/percona-monitoring-and-management/3/release-notes/3.5.0.html>
63. Quest. (2025). “Alarms AI is Here in Foglight for Databases v.8”. <https://blog.quest.com/product-post/alarms-ai-is-here-in-foglight-for-databases-v-8/>
64. S01:<https://www.gartner.com/en/documents/6688834>
65. S02:<https://www.gartner.com/en/documents/6755734>
66. S03:<https://www.gartner.com/en/newsroom/press-releases/2025-08-05-gartner-hype-cycle-identifies-top-ai-innovations-in-2025>
67. S04:<https://www.forrester.com/report/the-forrester-wave-tm-aiops-platforms-q2-2025/RES182220>
68. S05:<https://my.idc.com/getdoc.jsp?containerId=US54116226>
69. S06:<https://my.idc.com/getdoc.jsp?containerId=US53004325>
70. S07:<https://www.datadoghq.com/about/latest-news/press-releases/datadog-launches-bits-ai-sre-agent-to-resolve-incidents-faster/>
71. S08:<https://www.datadoghq.com/product/ai/bits-ai-sre/>
72. S09:<https://www.dynatrace.com/news/blog/announcing-general-availability-of-davis-copilot-your-new-ai-assistant/>
73. S10:<https://www.dynatrace.com/news/press-release/forrester-wave-aiops-platforms-q2-2025/>
74. S11:<https://docs.newrelic.com/whats-new/2025/06/whats-new-06-04-nrai/>
75. S12:<https://newrelic.com/blog/news/new-relic-named-a-leader-in-the-2026-idc-marketscape-for-worldwide-aiops>
76. S13:[https://www.splunk.com/en\\_us/blog/conf-splunklive/supercharge-ins](https://www.splunk.com/en_us/blog/conf-splunklive/supercharge-ins)

- ights-with-generative-ai-splunk-ai-assistant-for-spl-is-now-ga.html
77. S14:[https://www.splunk.com/en\\_us/blog/artificial-intelligence/scaling-g-anomaly-detection-with-mltk-5-5.html](https://www.splunk.com/en_us/blog/artificial-intelligence/scaling-g-anomaly-detection-with-mltk-5-5.html)
  78. S15:[https://www.splunk.com/en\\_us/newsroom/press-releases/2025/cisco-data-fabric-transforms-machine-data-into-ai-ready-intelligence.html](https://www.splunk.com/en_us/newsroom/press-releases/2025/cisco-data-fabric-transforms-machine-data-into-ai-ready-intelligence.html)
  79. S16:<https://www.elastic.co/observability/ai-assistant>
  80. S17:<https://www.businesswire.com/news/home/20251114859815/en/Elastic-Named-a-Leader-in-2025-IDC-MarketScape-for-Worldwide-Observability-Platforms>
  81. S18:<https://grafana.com/docs/grafana-cloud/alerting-and-irm/sift/>
  82. S19:<https://www.servicenow.com/products/itom.html>
  83. S20:<https://www.bmc.com/it-solutions/bmc-helixgpt.html>
  84. S21:<https://www.ibm.com/products/instanta>
  85. S22:<https://aws.amazon.com/about-aws/whats-new/2025/06/ga-accelerate-troubleshooting-amazon-cloudwatch-investigations/>
  86. S23:<https://aws.amazon.com/about-aws/whats-new/2026/03/aws-devops-agent-generally-available/>
  87. S24:<https://techcommunity.microsoft.com/blog/appsonazureblog/announcing-general-availability-for-the-azure-sre-agent/4500682>
  88. S25:<https://cloud.google.com/products/gemini/cloud-assist>
  89. S26:<https://blogs.oracle.com/observability/latest-enhancements-we-rolled-early-2025-in-oci-ops-insights>
  90. S27:<https://www.anthropic.com/news/model-context-protocol>
  91. S28:<https://modelcontextprotocol.io/>
  92. S29:<https://opentelemetry.io/docs/specs/semconv/gen-ai/>
  93. S30:<https://www.solarwinds.com/database-performance-analyzer/use-cases/ai-database-monitoring>
  94. S31:<https://blog.quest.com/product-post/alarms-ai-is-here-in-foglight-for-databases-v-8/>

95. S32:<https://docs.percona.com/percona-monitoring-and-management/3/release-notes/3.5.0.html>
96. S33:<https://www.datadoghq.com/product/database-monitoring/>
97. S34:<https://www.pagerduty.com/blog/product/product-launch-2025-h2/>
98. S35:<https://incident.io/ai-sre>
99. S36:<https://jennifersoft.com/ko/blog/tech/2025-12-22-jenniferai-jenniferinsights/>
100. S37:<https://whatap.io/ko/blog/ai-business-observability-it-operations-2025>
101. S38:<https://www.seouln.com/news/articleView.html?idxno=618704>
102. S39:<https://www.pinpointnews.co.kr/news/articleView.html?idxno=327086>
103. S40:<https://ddaily.co.kr/page/view/2025041610500725990>
104. S41:<https://www.fsc.go.kr/no010101/86080>
105. S42:<https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=BS074&mCode=C020010000&nttId=11410>
106. S43:<https://www.etnews.com/20250317000203>
107. S44:<https://www.mordorintelligence.com/industry-reports/observability-market>
108. S45:<https://www.fortunebusinessinsights.com/aiops-market-109984>
109. S46:<https://www.researchnester.com/reports/observability-tools-and-platforms-market/8139>
110. S47:<https://www.networkworld.com/article/4032218/in-crowded-observability-market-gartner-calls-out-ai-capabilities-cost-optimization-devops-integration.html>
111. S48:<https://www.comworld.co.kr/news/articleView.html?idxno=51321>
112. S49:<https://www.newstheai.com/news/articleView.html?idxno=9871>
113. S50:<https://www.dailysecu.com/news/articleView.html?idxno=200759>
114. S51:<https://www.globenewswire.com/news-release/2025/12/09/3202429/0/en/Application-Performance-Monitoring-Market-Set-for-Strong-Growth-to-U>

- [SD-25-24-Billion-by-2033-Driven-by-Digital-Transformation-and-Cloud-Native-Complexity-Research-by-SNS-Insider.html](#)
115. S52:<https://docs.newrelic.com/docs/agentic-ai/mcp/overview/>
116. S53:[https://docs.datadoghq.com/llm\\_observability/monitoring/cost/](https://docs.datadoghq.com/llm_observability/monitoring/cost/)
117. S54:<https://www.techtarget.com/searchitoperations/news/366630519/Under-Cisco-Splunk-AI-roadmap-tees-up-pricing-overhaul>
118. S55:<https://www.datadoghq.com/blog/dash-2025-new-feature-roundup-keynote/>
119. S56:<https://learn.microsoft.com/en-us/azure/sre-agent/pricing-billing>
120. S57:<https://aws.amazon.com/devops-agent/pricing/>
121. S58:<https://www.computerweekly.com/news/366627660/Datadog-doubles-down-on-APAC-targets-faster-growth>
122. S59:<https://code.visualstudio.com/blogs/2025/04/07/agentMode>
123. S60:<https://www.samsungsds.com/kr/insights/2025-dx-prospects.html>
124. S61:<https://www.lawtimes.co.kr/news/articleView.html?idxno=215867>
125. S62:<https://www.atlassian.com/software/jira/ai>
126. S63:<https://docs.firehydrant.com/docs/ai-powered-incident-management>
127. S64:<https://www.pagerduty.com/blog/product/product-launch-enhancements-to-pagerduty-operations-cloud-2025-h1/>
128. S65:<https://newsroom.cisco.com/c/r/newsroom/en/us/a/y2024/m03/cisco-completes-acquisition-of-splunk.html>
129. S66:<https://docs.dynatrace.com/docs/discover-dynatrace/platform/davis-ai/copilot/copilot-data-privacy>
130. S67:<https://docs.dynatrace.com/docs/discover-dynatrace/platform/davis-ai>
131. S68:<https://www.dynatrace.com/news/press-release/dynatrace-expanding-davis-hypermodal-ai/>
132. S69:<https://docs.newrelic.com/docs/security/security-privacy/compliance/fedramp-compliant-endpoints/>

133. S70:<https://www.ibm.com/watsonx>
134. S71:<https://www.ibm.com/granite>
135. S72:<https://grafana.com/grafana/plugins/grafana-llm-app/>
136. S73:<https://github.com/grafana/mcp-grafana>
137. S74:<https://azure.microsoft.com/en-us/products/sre-agent/>
138. S75:<https://www.etnews.com/20251212000235>
139. S76:<https://www.oracle.com/autonomous-database/features/>
140. S77:<https://www.mongodb.com/docs/atlas/performance-advisor/access-with-mcp/>
141. S78:<https://blog.quest.com/product-post/stop-reacting-start-preventing-how-foglight-8-1-puts-you-ahead-of-database-problems/>
142. SamsungSDS. (2025). “Brightics AI, Fabrix, Brity Copilot”.<https://www.samsungsds.com/kr/insights/2025-dx-prospects.html>
143. ServiceNow, Inc. (2025). “ITOM.”<https://www.servicenow.com/products/itom.html>
144. SolarWinds. (2025). “AI Database Monitoring Use Cases”.<https://www.solarwinds.com/database-performance-analyzer/use-cases/ai-database-monitoring>
145. Splunk, Inc. (2024). “Supercharge Insights with Generative AI: Splunk AI Assistant for SPL Is Now GA.”[https://www.splunk.com/en\\_us/blog/conf-splunklive/supercharge-insights-with-generative-ai-splunk-ai-assistant-for-spl-is-now-ga.html](https://www.splunk.com/en_us/blog/conf-splunklive/supercharge-insights-with-generative-ai-splunk-ai-assistant-for-spl-is-now-ga.html)
146. Splunk. (2023). “Scaling anomaly detection with MLTK”.[https://www.splunk.com/en\\_us/blog/artificial-intelligence/scaling-anomaly-detection-with-mltk-5-5.html](https://www.splunk.com/en_us/blog/artificial-intelligence/scaling-anomaly-detection-with-mltk-5-5.html)
147. Splunk. (2025). “Supercharge Insights with Generative AI: Splunk AI Assistant”.[https://www.splunk.com/en\\_us/blog/conf-splunklive/supercharge-insights-with-generative-ai-splunk-ai-assistant-for-spl-is-now-ga.html](https://www.splunk.com/en_us/blog/conf-splunklive/supercharge-insights-with-generative-ai-splunk-ai-assistant-for-spl-is-now-ga.html)
148. VS Code. (2025). “Agent Mode”.<https://code.visualstudio.com/blogs/2025>

[/04/07/agentMode](#)

- 149. Whatap. (2025). “OpsLake 및 AI 챗봇”. <https://whatap.io/ko/blog/ai-business-observability-it-operations-2025>
- 150. eXem. (2025). “eXemble LLMOps 및 MCP 연동”. <https://www.seoulfn.com/news/articleView.html?idxno=618704>
- 151. 개인정보보호위원회. (2025). “생성형 AI 개발·활용 안내서”. <https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=BS074&mCode=C020010000&nttId=1410>
- 152. 금융위원회. (2024). “금융분야 망분리 개선 로드맵”. <https://www.fsc.go.kr/no010101/86080>
- 153. 로타임즈. (2025). “망분리·SaaS 도입 정책”. <https://www.lawtimes.co.kr/news/articleView.html?idxno=215867>
- 154. 전자신문. (2025). “CSAP 2026 변화 보도”. <https://www.etnews.com/20251212000235>
- 155. 전자신문. (2025). “국정원 N²SF 적용 현황”. <https://www.etnews.com/20250317000203>

## Glossary

용어	정의
망분리	업무망과 인터넷망을 물리적으로 분리하여 보안성을 강화하는 정책
에이전틱 워크플로	Agentic Workflow. LLM/AI가 도구 호출, 계획, 실행을 자율적으로 수행하는 폐쇄 루프 자동화.
AAU	Azure Automation Unit, Azure 자동화 과금 단위
agent-second	에이전트가 동작한 시간(초) 단위로 과금하는 모델
Agentic 워크플로	AI가 탐지→진단→조치→검증(폐쇄 루프)을 자율적으로 수행하는 자동화 워크플로.
Agentic AI	도구 호출, 계획, 메모리 등 복합적 기능을 통해 detect→diagnose→remediate→verify의 폐쇄 루프를 자율적으로 수행하는 AI.
Agentic AIOps	LLM 기반 도구 호출·계획·메모리 결합형 자동화, 트러블슈팅 폐쇄 루프 구현
AIOps	AI for IT Operations, IT 운영 자동화에 AI를 적용하는 기술

<b>Air-gap</b>	물리적 망분리 환경에서 외부 네트워크와 완전히 분리된 시스템.
<b>APM</b>	Application Performance Monitoring, 애플리케이션 성능 모니터링
<b>ARMS Copilot</b>	Alibaba Cloud의 AI 운영 에이전트
<b>Azure Arc</b>	멀티클라우드 및 온프레미스 자원을 Azure에서 통합 관리하는 플랫폼
<b>Bottom-up 모니터링</b>	인프라→비즈니스 영향까지 거슬러 올라가는 추적 방식
<b>BYO LLM</b>	Bring Your Own LLM. 고객이 원하는 대형 언어모델(LLM)을 직접 선택·연결하는 방식.
<b>Causal AI</b>	통계적 상관관이 아닌 인과 추론 기반의 근본 원인 분석 AI
<b>Causal RCA</b>	인과 기반 근본원인분석(Root Cause Analysis). 통계적 상관관이 아닌 인과관계에 근거한 장애 원인 분석.
<b>Configuration as Code</b>	인프라 자원(서버, 네트워크 등)의 Git 기반 코드 관리
<b>CSAP</b>	Cloud Security Assurance Program, 공공기관용 클라우드 보안 인증제
<b>CSP</b>	Cloud Service Provider, 클라우드 서비스 제공자
<b>Current Offering</b>	Forrester Wave 등에서 벤더의 현재 기능 및 제품 완성도를 평가하는 항목.
<b>GA</b>	General Availability, 정식 출시
<b>Hype Cycle</b>	Gartner가 기술의 시장 기대치와 실제 도입 타이밍을 시각적으로 표현한 곡선.
<b>LLM Observability</b>	LLM 기반 AI 앱의 품질, 성능, 비용, 보안 등을 모니터링하는 기능.
<b>Magic Quadrant(MQ)</b>	Gartner가 시장의 주요 벤더를 리더, 비전가 등 4개의 사분면으로 평가하는 리서치 보고서.
<b>MCP</b>	Model Context Protocol. LLM 에이전트와 외부 도구 간 표준화된 데이터 교환 프로토콜.
<b>Monitoring</b>	알려진 문제를 임계값, 룰 기반으로 추적하는 전통적 IT 운영 방식
<b>Monitoring as Code</b>	모니터링 정책(알람, 대시보드 등)의 Git 기반 코드 관리
<b>MTTR</b>	Mean Time To Resolution, 장애 발생~복구까지의 평균 시간
<b>N°SF</b>	국가정보원 국가정보보호 프레임워크(National Information Security Framework)
<b>NL</b>	Natural Language, 자연어 질의 및 응답
<b>NL 질의</b>	Natural Language Query, 자연어 질의.
<b>Observability</b>	외부 출력(메트릭, 로그, 트레이스 등)으로 시스템 내부 상태를 추론하는 능력
<b>OCI</b>	Oracle Cloud Infrastructure, 오라클 클라우드 인프라
<b>Outcome-centric, governed automation</b>	결과 중심의 의사결정과 거버넌스 내재화 자동화를 의미하는 2026년 IDC 평가 키워드.
<b>PII</b>	Personally Identifiable Information, 개인식별정보.
<b>PIPA</b>	개인정보보호법(Personal Information Protection Act), 대한민국 개인정보 보호의 기본법
<b>PoC</b>	Proof of Concept, 실증 검증 단계
<b>QAN</b>	Query Analytics, 쿼리 분석 도구.


<b>RAG</b>	Retrieval-Augmented Generation. 외부 데이터 검색과 LLM 생성을 결합한 AI 아키텍처.
<b>RCA</b>	Root Cause Analysis, 근본 원인 분석
<b>SaaS</b>	Software as a Service, 클라우드 기반 소프트웨어 제공 방식.
<b>Self-hosted</b>	온프레미스 또는 고객 데이터센터 내에 직접 설치·운영하는 방식.
<b>Self-managed</b>	조직이 직접 설치·운영·업그레이드하는 방식.
<b>SKU</b>	Stock Keeping Unit, 제품 기능별 라이선스 단위
<b>sLLM</b>	Specialized LLM 또는 Self-hosted LLM. 특정 도메인에 특화되거나 자체 호스팅 가능한 LLM.
<b>SRE</b>	Site Reliability Engineering, 신뢰성 엔지니어링
<b>SRE Agent</b>	Site Reliability Engineering Agent, 운영 자동화 에이전트
<b>Strategy</b>	Forrester Wave 등에서 벤더의 미래 전략, 로드맵, 혁신 가능성을 평가하는 항목.
<b>Top-down 모니터링</b>	비즈니스 KPI→인프라로 내려가는 진단 방식

## Endnotes

[1] agent-second 과금은 사용량 예측이 어렵기 때문에 실제 운영 패턴에 따라 월간 비용이 크게 달라질 수 있습니다. [2] 한국 리전 AI 기능 가용성은 CSP의 정책 변화, 신규 서비스 론칭에 따라 수시로 변동될 수 있으므로, 도입 직전 반드시 공식 공지를 확인해야 합니다. [3] 망분리 단계별 정책 변화는 금융위 및 국정원 공식 공지에 따라 수시로 변동될 수 있으므로, 실제 도입 전 최신 정책 확인이 필수입니다. [4] PIPA 4단계 분류는 2025년 PIPC 가이드라인을 기준으로 하며, 각 단계별 적법 근거는 법무팀과 사전 협의 필요. [5] CSAP 변화는 공공 클라우드와 SaaS-AI의 적용성에 차이가 있으므로, 공공기관별 세부 지침을 별도 확인해야 합니다.

# Contact Us

 [hello@cncf.co.kr](mailto:hello@cncf.co.kr)

 02-469-5426

 [www.cncf.co.kr](http://www.cncf.co.kr)

## CNF Blog

다양한 콘텐츠와 전문 지식을 통해 더 나은 경험을 제공합니다.

## CNF eBook

이제 나도 클라우드 네이티브 전문가  
쿠버네티스 구축부터 운영 완전 정복

## CNF Resource

Community Solution의 최신 정보와  
유용한 자료를 만나보세요.

