

디지털 전환 위한 데이터 가치 실현 데이터 인프라 준비는 이렇게!

전 세계적으로 데이터가 폭증하면서 기업들의 고민이 깊어지고 있다. 복잡한 데이터로 인한 문제가 가장 시급하게 해결해야 할 과제이며 비즈니스 지속가능성, 데이터 보호, 거버넌스 관련 윤리 조건과 규제도 늘고 있다.

현대적인 데이터 인프라를 구축하면 이런 문제를 쉽게 해결할 수 있지만, 모든 기업이 이를 적용하기는 쉽지 않다. 데이터가 주도하는 세상에서 기업 지속가능성의 토대가 되어줄 데이터 인프라를 어떻게 준비해야 하는지 알아본다.

데이터는 최고의 비즈니스 자산

데이터 가치 실현을 원한다면 이것만은 해결해야!

데이터 인프라 확장성 부족

데이터의 양이 이미 감당할 수 있는 수준을 넘어섰으며, 폭증하는 데이터가 어디에 사용될 것인지 파악하는 것조차 쉽지 않음

랜섬웨어 공격에 대한 복원성 부족

보안은 온프레미스뿐 아니라 퍼블릭/프라이빗 클라우드에서도 가장 중요한 이슈

데이터 가치 실현, 핵심은 데이터 현대화

만반의 준비를 갖추고 있는 기업들도 서드파티 벤더 및 파트너의 도움이 필요하다는 데 동의

‘관리해야 할 데이터, 한계치를 넘었다!’

히타치 벤틀라가 12개국의 대기업 의사결정권자 1,288명을 대상으로 인터뷰를 진행한 바에 따르면, 2023년 현재 전 세계 기업들이 보유한 데이터는 평균 35PB로 추정된다. 2025년에는 이 수치가 65PB에 달할 것으로 보이며, 아시아 지역의 일부 기업은 최대 80PB까지 될 것으로 전망하기도 한다. 80PB는 현재 NASA가 보유하고 있는 지구과학 관련 전체 데이터보다 큰 규모다. 데이터 인프라와 관련해 기업들이 현재 처한 상황을 구체적으로 살펴보자.

데이터가 증가하면 데이터 관리와 스토리지 비용도 증가한다. 데이터의 증가는 단순히 양적인 증가로 끝나는 것이 아니다. 이에 맞물려 복잡성, 소스, 포맷도 함께 증가하기 때문이다. 조사 대상 기업의 61%가 ‘관리해야 할 데이터의 양이 이미 한계치를 넘었다’고 답했으며, 75%는 ‘현재의 데이터 인프라가 향후 2년 이내에 전체 데이터 수요를 충족시킬 수 있을 정도로 확장성이 부족하다’고 말했다. 약 30%의 기업들만이 향후 2년 이내에 비정형 IoT 및 감시 데이터 관련 투자를 26~50%까지 늘릴 예정이라고 밝혔다.

기업들은 현재 사용 중인 데이터 인프라가 급변하는 기술 환경에 충분히 대응하기 어렵다는 데 대부분 공감했다. 디지털 전환 시대에 대비한 데이터 인프라 업데이트와 현대화에 대한 필요성을 인지하고 있는 것이다.

현재 전 세계 데이터센터에서 배출되는 온실가스(GHG)는 전체 온실가스 배출량의 최대 3.7%에 달한다. 데이터센터는 냉각을 위해 엄청난 양의 물을 사용하며, 2020년 한 해에만 300미터톤¹⁾에 달하는 CO2를 배출했다. 그러나 조사 기업 중 약 1/4 정도는 지속가능성 목표를 갖고 있지 않다고 말했으며, 지속가능성 목표를 갖고 있다고 답한 기업들의 절반 이상도 다크 데이터까지는 범주에 넣고 있지 않다고 밝혔다. 그렇다면 지속 가능하고 환경친화적인 기술과 데이터 운영의 책임은 누구에게 있을까? 응답 기업의 절반 정도는 기업의 리더를 꼽았으나, 서드파티 벤더와 컨설턴트를 꼽은 기업도 적지 않았다.

응답 기업의 절반 정도가 데이터센터 리소스의 에너지 절감 목표와 소비량도 함께 측정하고 있는 것으로 조사되었다. 그러나 10곳 중 1곳은 에너지 절감 목표는 있지만 소비량은 측정하지 않고 있으며, 10곳 중 2곳은 에너지 절감 목표 없이 소비량만 측정한다고 답했다. 지속가능성 목표를 염두에 두고 인프라 폐기를 관리하는 기업은 73%에 달했으며, 76%는 현재 운영 중인 데이터 인프라 솔루션이 환경친화적이라고 생각하고 있었다.

보안 부문에서는 어떨까? 데이터 관련 보안과 통제가 강화되면 사용자는 다소 귀찮다고 느낄 수 있다. 그러나 현실적으로 이를 피할 수 있는 방법은 없다. 응답 기업의 절반 이상은 ‘데이터 침해를 제때 감지하지 못해 발생할 수 있는 데이터 유출’을 가장 우려하고 있었다. 최근 랜섬웨어 공격이 진화를 거듭하면서, 랜섬 비용에 대한 지급 요구뿐 아니라 비용을 지불하지 않을 시 데이터를 유포하거나 판매하겠다는 위협까지 더해지고 있다.

1) 미터톤(meter ton): 미터법에 따른 질량의 단위로, 1미터톤은 1kg의 1,000배에 해당한다.

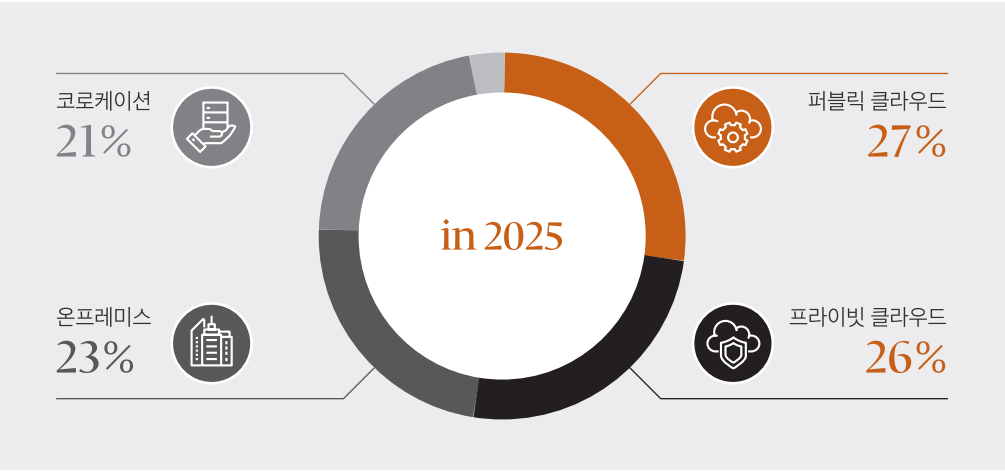
데이터 인프라에서 가장 중요한 기준은 손상을 입지 않는 것이다. 설령 재해 등으로 인해 데이터 손실이 발생하더라도 100% 복구할 수 있어야 한다. 그러나 이번 조사에 응한 기업의 22%는 중요한 데이터를 백업하지 않고 있다고 답했으며, 18%는 데이터 손상으로 인해 지난 2년간 해당 데이터에 액세스하지 못했다고 밝혔다. 가장 취약한 보안 지점은 '직원'이다. 직원들이 보안 정책을 따른다고 응답한 비율은 29%에 불과했다.

최적의 데이터 인프라를 위한 현실적인 대안 모색

지난 몇 년간 기업들은 클라우드를 데이터 인프라 문제를 해결할 수 있는 솔루션으로 생각해 왔다. 지속가능성, 보안, 접근성 문제를 클라우드 공급 벤더가 도맡아 주기를 기대하면서 말이다. 그러나 이런 기대와 달리 클라우드는 대체재로 100% 충분하지 않다. 복원성, 확장성, 유연성 부문에서 클라우드의 이점은 분명하지만, 반대로 이런 점 때문에 기업의 데이터 인프라에 또 다른 문제가 생길 수 있다. 적절한 데이터옵스(DataOps) 솔루션을 갖고 있지 않으면, 클라우드와 함께 구축된 온프레미스 리소스나 애플리케이션 환경에 대한 모니터링, 관리 등이 어려울 수 있기 때문이다. 원격 업무 확산으로 인해 사이버 공격 지점이 증가하는 것도 어려움이 커지는 요인이다.

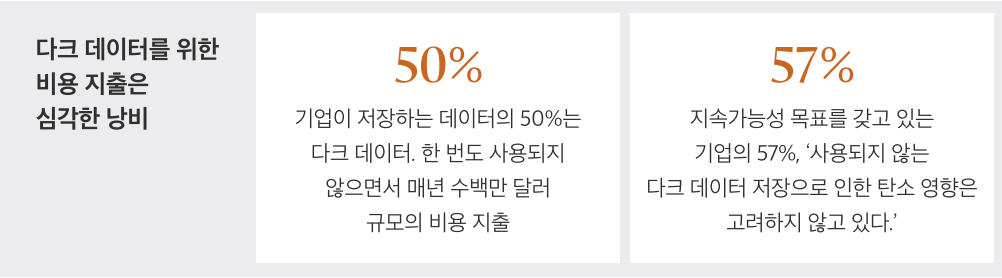
기업들의 응답 현황을 보면, 디지털 전환 측면에서는 44%가 클라우드를, 보안 측면에서는 32%가 온프레미스를 선호한다고 답했다. 원격 업무의 증가 등 최근 직면하게 된 복잡한 비즈니스 문제들을 해결할 방안으로 내부 인프라가 나올지, 외부 인프라가 나올지를 두고 기업들 사이에서 의견이 갈리고 있는 것으로 보인다.

↓ 2025년 플랫폼별 데이터센터 워크로드 비중



데이터 저장 목적과 관련해서는, 많은 기업이 전보다 더 많은 양의 데이터를 수집하고 있지만 데이터의 잠재력을 충분히 인식하고 있지 못한 것으로 나타났다. 응답 기업의 75%가 '만일의 경우를 대비해' 모든 데이터를 저장한다고 밝혔다. 수집, 처리, 저장만 할 뿐 단 한 번도 활용되지 않고 있는 다크 데이터는 평균 17PB로 추정되며, 다크 데이터가 가장 많은 분야는 교통, 에너지, 설비, 헬스케어, 공공 분야였다.

데이터 활용 분야 역시 마찬가지다. 응답 기업의 25%가 사용자가 자신이 원하는 데이터가 이미 존재한다는 사실을 모를 수 있다고 답했으며, 특히 소매와 기술 부문에서 이런 경향이 두드러졌다. 사용자가 데이터의 존재를 알고 있다 하더라도, 시스템 간 연계가 원활하지 않으면 필요한 데이터를 활용하기가 쉽지 않다. 놀랍게도 무려 80%의 기업이 필요한 데이터에 액세스하지 못하고 있다고 답했다. 그 이유로 '시스템에 통합된 데이터가 사용자가 필요로 하는 데이터의 절반에도 미치지 못해서', '기존의 데이터를 적절한 시스템에서 이용할 수 없어서' 등이라고 답했다. 필요한 데이터에 제대로 액세스하기 어렵다고 답한 기업도 29%나 되었다.



하이브리드 미래에 대한 기대가 커지면서 시스템 간 원활한 연계가 중요한 이슈로 대두되고 있다. 60% 이상의 응답자가 내·외부 데이터 인프라 간 격차를 해소하기 위해 시스템 간 통합이 필요하다고 밝혔으며, 특히 공공 부문 응답률이 높았다. 외부 전문가에 대한 의존도도 낮지 않아, 데이터 통합 사이버 보안을 위해 서드파티의 도움이 필요하다고 답했다.

지난 2년 동안 스토리지 손상으로 인해 데이터 손실을 경험한 적이 있는 기업은 18%로 조사되었으며, 중요한 데이터가 백업되지 않고 있다는 사실을 알고 있는 기업도 22%에 달했다. 과도한 보안 정책으로 인해 사용자가 데이터에 접근이 용이하지 않고, 보안 조치로 인해 데이터 액세스가 복잡하고 시간도 많이 소요된다고 답변한 기업도 많았다. 클라우드 스토리지 플랫폼을 사용 중인 응답 기업의 절반 가까이는 보안에 대한 우려 사항으로 규제 관련 이슈를 꼽기도 했다.

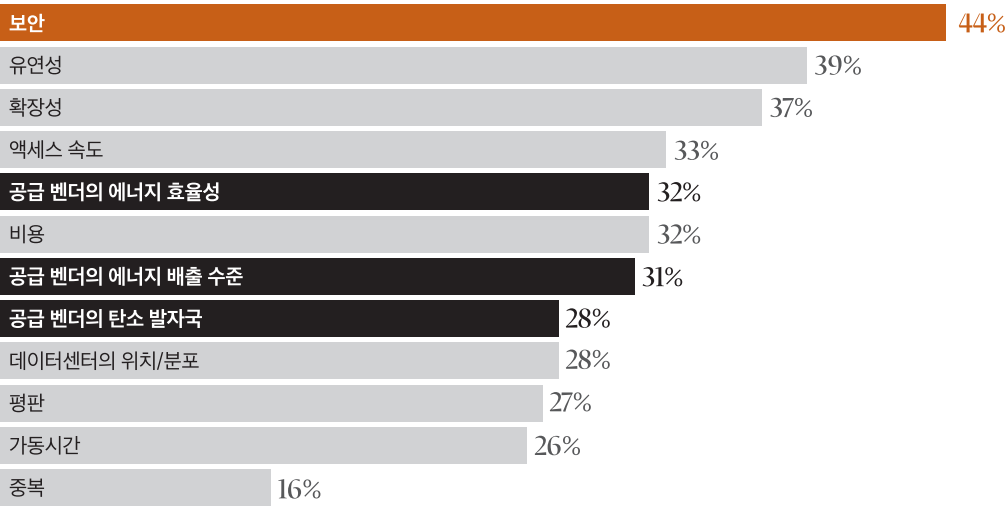
완벽한 데이터센터, 서드파티 벤더 도움 절실

많은 기업은 더 많은 비즈니스 가치를 제공하는, 유연성과 확장성을 갖춘 데이터 인프라 현대화가 이뤄져야 디지털 전환에 박차를 가할 수 있다고 인식하고 있지만, 실상은 그렇지 못한 것으로 나타났다. 데이터 관리의 어려움을 호소했으며, 수집·저장되는 모든 데이터를 처리할 수 있는 인력 자체가 전무하다는 답변도 절반 가까이 됐다. 효율적인 데이터 인프라 관리를 위해 서드파티 벤더의 도움이 필요하며, 전문적인 서드파티 벤더와 함께 환경 문제에도 적극적으로 대응할 수 있기를 기대한다고 답변한 기업도 각각 30%가 넘었다.

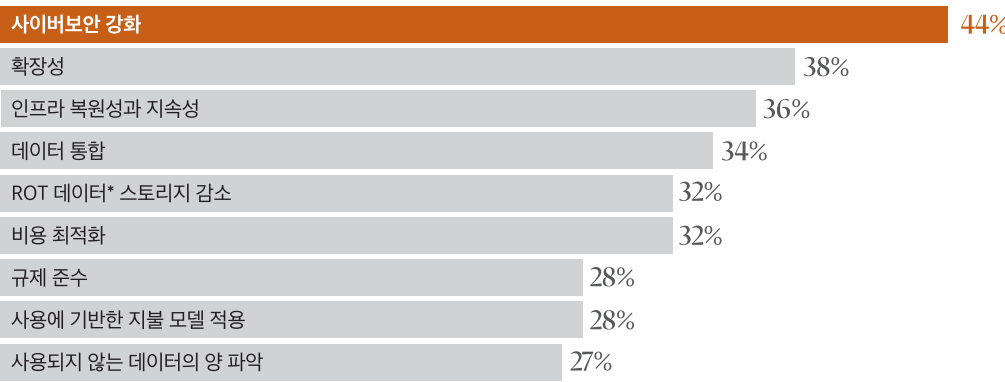
그렇다면, 서드파티 벤더를 선정할 때 고려해야 할 사항은 무엇일까? 지속가능성에 대한 관심이 높아지면서 서드파티 데이터 솔루션 공급 벤더를 선정할 때도 환경적 요인을 중요하게 고려하는 것으로 조사되었다. 데이터 스토리지 솔루션 공급 벤더 선정 시 가장 중요한 고려 요소로는 보안, 유연성, 확장성 순으로 나타났다.

보안과 관련해서는, 대부분의 응답자가 인프라 설계 시점부터 효과적인 통합 보안이 필요하다고 인지하고 있었다. 서드파티 벤더가 사이버 보안을 담당하는 기업은 절반 가까이 되며, 중국, 인도 네시아는 서드파티 벤더 의존도가 더 높았다. 또 20% 정도의 응답 기업이 데이터 보안을 100% 아웃소싱한다고 답했으며, 43%는 외부 데이터 보안 서비스 또는 플랫폼을 추가해 사용하고 있다고 밝혔다.

↓ 데이터 스토리지 솔루션 공급업체 선정 시 가장 중요하게 고려하는 요소는?

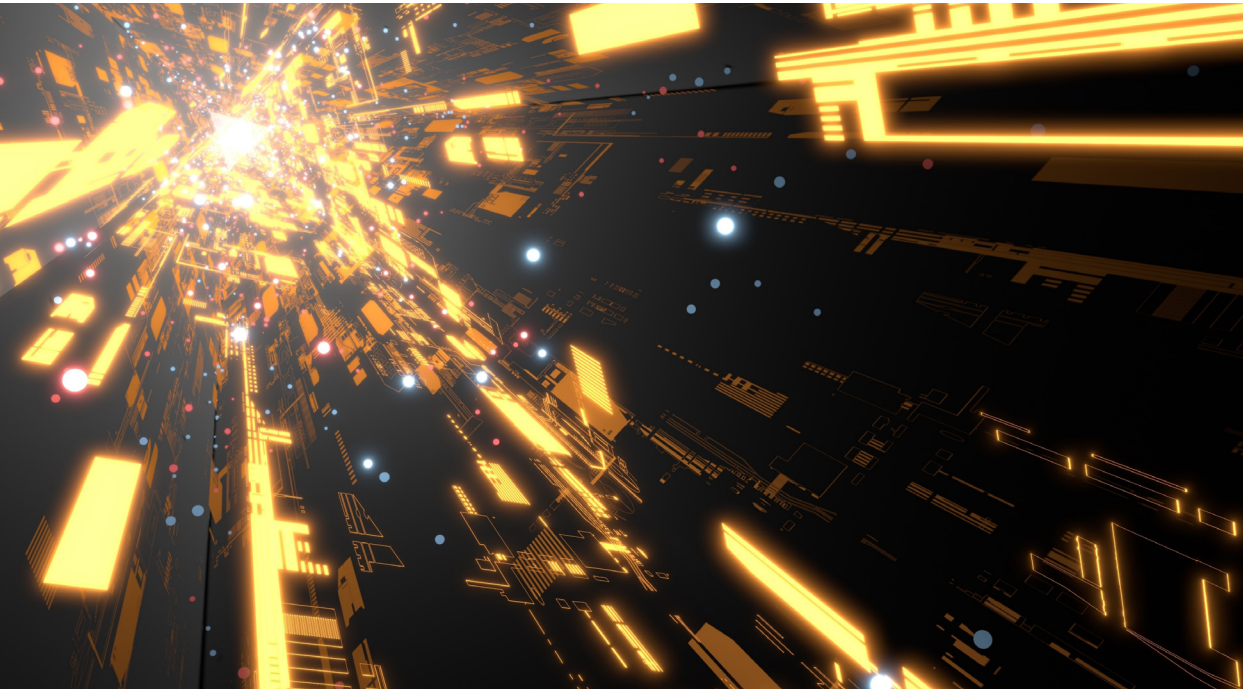


↓ 서드파티 벤더의 도움이 필요한 인프라 관리 분야



* ROT Data : 중복되고, 오래되고, 중요하지 않은 (Redundant, Obsolete, Trivial) 데이터

현대적인 데이터 인프라는 복잡성을 줄여줄 뿐만 아니라 지속가능성과 복원성을 확보할 수 있는 토대가 된다. 처음부터 지속가능성과 복원성을 염두에 두고 설계되었기 때문이다. 제대로 된 서드파티 벤더라면 충분한 경험과 전문지식을 기반으로 효율적인 데이터센터 인프라 및 솔루션을 책임 있게 설계하고 운영할 수 있다. 조사 기업들의 대부분은 환경 관련 이슈에도 적극적으로 도와줄 수 있는 공급 벤더의 전문성을 기대하고 있었다.



데이터 가치 실현을 위한 최적의 전략

효성인포메이션시스템과 히타치 벤틀라는 비즈니스 리더들이 기술과 인력, 그리고 데이터 사일로를 유발하는 환경 등의 문제를 해결할 수 있는 데이터 가치 실현 방안에 대한 몇 가지 가이드라인을 제안한다. 데이터 관련 정책은 기업의 데이터 전략과 사용자, 그리고 기술 간 융합이 잘 되도록 정기적으로 수정되어야 한다.

01 | 인프라 현대화 목적에 부합하지 않는 레거시 시스템은 디지털 전환에 걸림돌이 될 뿐이다. 이번 조사 결과, 31%의 기업이 기존 시스템을 여전히 사용하고 있으며, 현재의 데이터양을 처리할 수 있을 정도로 인프라가 충분치 않은 기업도 29%였다. 유연성과 성능 효율성을 제공하는 기술이 적용되고, 사용한 만큼 비용을 지불하는 현대적인 데이터 인프라를 도입하면 필요에 따라 유연하게 인프라를 확장할 수 있다.

02 | 사용자 권한 강화 현대적인 인프라 시스템은 데이터를 기반으로 신속한 의사결정을 지원한다. 이때 외부 보안 승인 프로세스와 액세스 제약이 걸림돌이 돼선 안 된다. 그러나 31%의 기업이 지나치게 오래 걸리는 보안 액세스 승인 절차로 인해 데이터 액세스에 어려움을 겪고 있다고 말했다. 보안과 액세스 프레임워크를 만들고, 가이드라인에 맞춰 이를 책임질 직원을 교육해야 한다.

03 | 낭비 요인 최소화 데이터를 검색할 때 한 번도 사용하지 않은 ‘불필요한’ 데이터가 지나치게 많다고 답한 기업이 28%에 달했지만, ROT(중복, 오래되거나 중요하지 않은) 데이터를 삭제해 데이터의 가치를 극대화하기 위해 노력하는 기업도 적지 않은 것으로 조사되었다. 28%가 ROT 데이터를 저장하느라 비용이 낭비되고 있다고 생각하고 있으며, 이러한 현상은 데이터 관리 시스템의 개수가 많을수록 두드러졌다.

04 | 시스템 통합 사용자가 데이터에 액세스하는 데 필요한 시스템의 수를 줄여야 한다. 가장 좋은 것은 기업의 전체 데이터 인프라를 하나의 포털에서 사용할 수 있어야 한다. 그러나 현실은 데이터 액세스를 위해 기업들이 사용하고 있는 시스템이 ‘평균 5개’에 달했다.

05 | 제약 요인 점검 보안 정책에 대한 정기적인 점검이 필요하다. 사용자가 필요한 데이터에 액세스하는 것을 불필요하게 제한하고 있지 않는지 늘 확인해야 한다. 응답 기업의 26%가 접근 권한에 문제가 있다고 답했다.

06 | 지능형 자동화 활용 효과적으로 데이터 스토리지를 최소화할 수 있도록 향상된 톨과 AI를 활용하기를 권한다. 76%가 현재 이렇게 하고 있다고 답했으나 기술은 끊임없이 변화하므로 현재 사용 중인 톨도 정기적으로 검토하고 업데이트해야 한다.

07 | 데이터 액세스 교육 서로 다른 시스템에서 데이터에 액세스하는 방법을 교육해야 한다. 27%의 기업이 데이터 액세스를 가로막는 요인으로 교육과 지식 부족을 꼽았다.

08 | 지식 공유 지식의 공유와 전달 방식을 간소화하는 것이 좋다. 32%의 기업이 중요한 업무를 담당한 직원들이 회사를 떠날 때 지식을 제대로 공유하지 않아 데이터의 가치가 하락했다고 답했다.

09 | 투명한 정보 제공 사용자가 자신이 액세스하는 데이터의 품질을 신뢰하지 않는다고 답한 기업이 23%에 달했다. 데이터 수집, 관리, 처리 방식과 관련해 사용자에게 투명한 정보를 제공해야 한다. 그래야 사용자가 데이터의 품질을 정확하게 평가하고, 필요에 맞게 사용할 수 있다.

10 | 데이터 카탈로그 데이터 카탈로그는 기업이 보유한 모든 데이터에 대한 정보 지도를 제공한다. 또한 투명성을 기반으로 데이터의 신뢰를 높여주므로 사용자가 데이터를 쉽게 검색하고 이해할 수 있다. 이렇게 하면 필요할 때 사용 가능한 데이터를 한눈에 파악할 수 있다.

11 | 처리 자동화 문서화와 권한을 지원할 수 있도록 향상된 톨을 사용해야 한다. 전문가 시스템을 이용하면 사용자가 데이터 활용을 극대화하고, 데이터 사용 또는 처리 시 조직의 전략적 목표를 침해하지 않을 수 있다.

12 | 인프라 모니터링 사용자의 요구가 변화되는 것처럼 데이터 환경도 끊임없이 변화한다. 따라서 데이터와 사용자의 수요, 잠재력과 더불어 조직의 전략적 목표를 달성하려면 데이터 정책도 끊임없이 수정해야 한다. 정책에 대한 정기적인 재점검은 필수다.

*출처: Modern Data Infrastructure Dynamics, www.hitachivantara.com, 2023년 7월