

자동차 데이터 공유의 장애/성공 요인 및 기술개발 과제

Bottleneck and Success Factors of Vehicle Data Sharing and Suggestions for Technology Development

김주성 (J.S. Kim, juskim@etri.re.kr)

지능화정책연구실 연구전문위원

ABSTRACT

Sharing vehicle data among the companies within a car ecosystem can improve driving experience, increase driver comfort, contribute to social goals such as improving road safety and lowering fuel consumption. Furthermore, by participating in the ecosystem, companies can secure long-term and sustainable new revenue-generating opportunities. In this paper, we will examine the bottleneck and success factors of data sharing, as well as the technological solutions that urgently require development for car data sharing.

KEYWORDS 기술개발과제, 데이터 공유, 성공 요인, 자동차 데이터, 장애 요인

1. 서론

최근 들어 차량에서 생성되는 데이터의 양이 증가함에 따라 제3자 서비스 제공업체는 서비스 제공을 위해 자동차 데이터에 액세스하고 사용하는 데 점점 더 많은 관심을 기울이고 있다. 서비스 제공업체에는 보험 회사, 주차장 운영자, 금융 및 차량 서비스 제공자, 도로 인프라 운영자, 차고 및 고장 서비스 업체 등이 포함되며, 그밖에 엔터테인먼트 및 여행 서비스 제공업체, 소셜 네트워크, 검색 엔진 운영자와 같은 비전통적인 자동차 업체도 포함된다.

생태계 내 참여기업 간에 데이터를 공유하면 운전 경험을 개선하고 운전자의 편안함을 높이며 제품을 최적화하고 도로 안전 개선 및 연료 소비 감소와 같은 사회적 목표에 기여할 수 있다. 또한, 생태계 참여기업은 이를 통해 장기적이고도 안정적인 새로운 수익 창출 기회를 확보할 수 있다.

하지만 아직 자동차 제조업체(OEM) 등에 축적된 다양한 데이터를 활용하는 방안은 제대로 모색되지 못하였으며, 특히 정부 차원에서도 이를 제도화하여 관리하는 방안을 제시하지 못하고 있다. 국내의 경우 자율주행 정보 기록장치 부착을 의무화하였으

* DOI: <https://doi.org/10.22648/ETRI.2022.J.370402>

* 이 논문은 '국가 지능화 R&D 경쟁력 제고를 위한 기술정책 연구'[21ZR1420] 과제 수행으로 발간된 "자율주행차/커넥티드카 데이터의 수익화 및 공유방안" 도서의 일부 내용을 재구성하여 작성되었음.

나, 자율주행차 데이터의 정의 및 분류체계조차 마련되지 않은 상태(외국의 경우에도 상황은 유사)이며, 특히 자동차 제조업체의 데이터 활용이 아직 미흡하므로 이를 효율적으로 활용하기 위한 자동차 데이터 공유 플랫폼의 구축이 필요한 실정이다[1].

이에 따라 본고에서는 자동차 데이터 가치사슬 및 자동차 데이터 공유 플랫폼을 분석하고, 데이터 공유의 부진 사유 및 성공 요인을 살펴본 다음, 데이터 공유를 위해 개발이 급히 필요한 기술 솔루션 과제를 살펴보기로 한다.

II. 자동차 데이터 액세스 방식

자동차 제조업체(OEM)가 제3자 서비스 제공업체에 자동차 데이터에 대한 액세스를 제공하는 방식에는 오프보드 액세스(Off-board Access) 방식과 차량 내 직접 액세스(Direct Access) 방식의 두 가지 방식이 있다.

OEM이 자동차 데이터를 타사 서비스에 제공하기 위해서는 차량 사용자의 개인 데이터 보호를 보

장하고, 차량의 안전하고 확실한 기능을 위협하지 않아야 하며, 자동차 제조업체로서의 책임을 훼손하지 않아야 한다. 이를 위한 균형 잡힌 대안은 차량 제조업체가 관련 데이터를 안전한 방식으로 오프보드 시설에 전달하는 것이다. 데이터에 대한 오프보드 액세스는 명확하게 정의된 기술, 데이터 보호 및 경쟁 규칙에 따라 작동하므로 제3자에게 개방적이면서도 보호된 인터페이스를 제공한다.

반면 데이터에 대한 차량 내 직접 액세스 방식은 기본적으로 제3자 서비스 제공자가 이동 중인 자동차 내부의 데이터에 직접적이고 통제되지 않은 액세스를 갖는 방식인데, 이는 차량, 운전자, 동료 승객 및 기타 도로 사용자에게 심각한 보안 및 안전 위험을 초래할 수 있다(표 1 참고)[2].

이에 비해 오프보드 시설은 다양한 인터페이스와 오프보드 데이터 저장 수단을 통해 명확하게 정의된 기술, 데이터 보호 및 경쟁 규칙에 따라 차량 데이터에 대한 액세스를 제공하여 안전, 보안 및 책임 위험을 줄일 수 있다. 또한 오프보드 시설은 자유 경쟁과 혁신을 촉진하기 위해 차량 제조업체와

표 1 차량 내 직접 액세스 방식의 문제점

문제점	주요 내용
해커 공격	<ul style="list-style-type: none"> 모든 새로운 외부 데이터 인터페이스는 잠재적인 표적 및 진입점 수를 증가시키기 때문에 차량 기능에 대한 제3자의 직접적인 액세스는 해커 공격을 용이하게 할 수 있음 - 자동차 절도나 원격 도어 잠금 해제를 가능하게 할 뿐만 아니라 사기 또는 개인 데이터 절도의 기회를 만들 수 있음
안전 관련 주요 기능의 신속한 처리 저해	<ul style="list-style-type: none"> 제동과 같은 안전에 중요한 기능이 승인되지 않은 타사 애플리케이션을 위한 차량 내 리소스 및 컴퓨팅 용량 사용으로 인해 부정적인 영향을 받을 수 있음 - 수십 개의 서비스 제공업체에 직접적이고 통제되지 않은 액세스 권한이 부여된다면 요청된 데이터의 양이나 빈도를 신속하게 처리하지 못할 수 있음
운전자 주의 산만	<ul style="list-style-type: none"> 외부 제3자가 차량의 운보드 시스템, 사용자 인터페이스 및 기능 디스플레이에 통제할 수 없는 접근을 하는 경우 운전자 주의 산만이라는 측면에서 추가적인 안전 위험이 발생할 수 있음 - 운전자 주의를 분산시킬 수 있는 애플리케이션이나 추가적 제어 장치를 통해 발생 가능
소프트웨어 오작동	<ul style="list-style-type: none"> 자동차, 밴, 트럭 또는 버스 내부의 데이터에 직접적이고 통제되지 않은 액세스를 제공하면 추가 소프트웨어 설치로 인해 심각한 오작동 또는 시스템 충돌과 같은 의도하지 않은 결과가 발생할 수 있음 - 자동차는 운전 중 문제가 발생하면 재부팅할 수 있는 PC와 같은 유형의 제품이 아니므로 안전, 보안 및 개인정보 보호에서 훨씬 더 높은 기준이 요구됨

출처 Reproduced from [2].

계약을 개별적으로 체결할 필요 없이 관심 있는 제3자 서비스 제공업체가 데이터를 쉽게 사용할 수 있도록 중립 서버를 설정하는 방식이다. OEM이 운영하거나 자금을 조달하지 않고 독립적인 당사자(Data Hub)가 운영하며, 서비스 제공업체는 차량 사용자에게 서비스를 제공하는 데 필요한 데이터에 공정하고 합리적이며 차별 없이 액세스할 수 있다. 이에 따라 고객의 선택권이 보장되어 차량 사용자는 차량 제조업체, 공인 수리업체 네트워크 또는 기타 서비스 제공업체로부터 서비스를 자유롭게 받을 수 있다. 또한, 개별 제조업체의 여러 서버를 사용하도록 강요하지 않고 하나의 서버에서 다중 브랜드 데이터 액세스를 제공함으로써 특히 중소기업의 데이터 액세스를 용이하게 할 수 있다.

III. 자동차 데이터 가치사슬 및 데이터 공유 플랫폼

제품 및 서비스에 대한 가치사슬과 유사하게 데이터의 상업적 사용에 필요한 활동으로 구성된 자동차 데이터 가치사슬은 원시 데이터에서 서비스의 데이터 적용까지 필요한 단계를 포함하고 있다. 자동차 데이터의 일반적인 가치사슬은 6개의 고유한 가치 창출 단계(데이터 생성, on-board 처리, 데이터 전송, off-board 처리, 서비스 개발, 서비스 사용)로 구성되며, 특정 이해관계자 집합이 데이터와 서비스를 사용할 수 있도록 하는 2개의 플랫폼(데이터 접속 플랫폼, 서비스 접속 플랫폼)을 포함하고 있다[3].

자동차 데이터 가치사슬 단계에 대해 구체적으로 살펴보면 센서를 통한 원시 데이터의 생성은 자동차 데이터 가치사슬의 첫 번째 단계이며, 센서 내부 처리를 거쳐 이들 데이터는 두 번째 단계인 추가적인 차량 온보드 처리를 위해 전자 제어장치로 전송된다.

처리된 데이터는 물리적 차량 시스템 이외의 데이터 사용을 위해 세 번째 단계인 중앙처리 및 저장 장치로의 전송이 필요한데, 주로 WLAN 또는 이동 통신을 통해 이루어진다.

네 번째 단계인 서버에서의 데이터 오프보드(Off-board) 처리는 앞에서 설명한 바와 같이 자동차 관련 데이터에서 경제적 가치를 창출하기 위한 핵심적인 가치사슬 단계이다. 훨씬 더 많은 양의 사용 가능한 데이터 소스와 더 높은 처리 성능을 이용하여 추가적인 데이터 분석 및 평가를 수행할 수 있으며, 이 단계에서 다른 소스의 외부 데이터(예: 인프라 데이터)를 결합할 수 있다. 다만 데이터에 대한 액세스가 OEM으로 제한되어 있어 경제적 가치를 창출할 가능성은 매우 제한적이므로 다른 기업에 대한 자동차 데이터에 대한 액세스 제공이 해당 데이터의 수익화를 위해 반드시 필요하다.

최근 들어 OEM 및 일부 제3자 기업들은 데이터를 수집하여 다운로드 프로세스에 제공하는 데이터 액세스 플랫폼을 구축하기 시작하였다. 다만 PSA의 P4D3 또는 BMW의 CarData4와 같은 OEM에서 운영하는 공개적으로 액세스 가능한 데이터 액세스 플랫폼들이 존재하지만 활동은 아직 제한적이다. 한편 국내에서도 현대자동차그룹에서 자동차 생성 데이터를 외부에서 활용 가능한 API 형태로 가공하여 파트너사에 제공하는 자동차 데이터 오픈 플랫폼을 개설 및 운영 중이다. 현대 디벨로퍼스, 기아 디벨로퍼스, 제네시스 디벨로퍼스 등 3개 플랫폼을 자체적으로 운영 중이며, 이중 현대 디벨로퍼스는 서비스 제공을 담당하는 소수의 선정된 스타트업에 데이터를 시범적으로 제공하고 있다. 일반적으로 이러한 플랫폼은 단일 OEM에서 선택한 데이터를 제공하고 해당 데이터에 대해 잘 정의된 일련의 조작을 허용한다. OEM과 달리 Otonomo나 TomTom과 같은 데이터 허브(또는 데이터 플랫폼) 기

업은 보다 광범위한 소스에서 데이터를 수집하고 수집된 데이터를 더 넓은 범위의 잠재 고객에게 공개적으로 제공할 수 있다.

보고에서는 ‘데이터 액세스 플랫폼’을 더 넓은 의미의 시장으로 정의하여 자동차 생태계의 잠재적인 모든 데이터 소스의 데이터를 사용할 수 있고, 해당 데이터에 관심이 있는 모든 당사자와 교환이 이루어지는 시장의 의미로 사용한다. 데이터는 가치사슬의 다른 단계를 따라 수익을 창출할 수 있는데, 데이터 액세스 플랫폼은 그 과정에서 가격을 형성하고 다양한 데이터 유형 및 품질에 가치를 부여하는 핵심적인 허브 시장 역할을 수행하게 된다. 데이터가 더 넓은 범위의 잠재적 구매자에게 제공되지 않으면 그 경제적 가치가 제대로 활용되지 않게 되므로, 데이터 액세스 플랫폼이 초기에는 ‘OEM 전용’으로 시작되겠지만 필연적으로 점차 더 많은 사용자에게 개방될 것으로 보인다.

데이터 액세스 플랫폼을 통해 자동차 데이터는 다양한 이해관계자가 사용할 수 있게 되며, 특히 제3자 서비스 제공업체는 이들 데이터를 활용하여 다섯 번째 단계인 최종 고객에게 어필하는 서비스 개발을 추진할 수 있다.

서비스 제공업체에 의해 개발된 서비스는 서비스 액세스 플랫폼을 통해 서비스 사용자에게 제공된다. 데이터 액세스 플랫폼과 마찬가지로 이 플랫폼은 서비스 개발자/제공자가 최종 고객에게 솔루션을 제공하는 시장을 나타낸다. 데이터 액세스 플랫폼과 서비스 액세스 플랫폼은 잠재적으로 동일한 서버에서 호스팅되거나 최소한 동일한 시장 참여기업이 운영하게 될 수도 있겠지만, 두 플랫폼은 각각 다른 목적과 다른 대상을 제공한다.

자동차 데이터 가치사슬의 마지막 단계인 여섯 번째 단계는 차량 또는 기타 장소에서 서비스(예: 고급 주차 서비스)를 실제로 사용하는 것이다. 비록 가

치사슬의 모든 단계에서 수익을 창출할 수 있지만, 이 마지막 단계는 자동차 데이터의 수익화가 가장 분명한 단계라고 할 수 있다.

IV. 자동차 데이터 공유의 장애 요인 및 성공 요인

1. 자동차 데이터 공유의 장애 요인

데이터 공유와 관련된 장애 요인은 기술적 장애, 법적 장애 및 관련 문제, 스킬 부족 및 기타 장애로 나누어 살펴볼 수 있다(표 2 참고)[4].

2. 자동차 데이터 공유의 성공 요인

데이터 공유와 관련된 성공 요인으로는 이해당사자 간의 신뢰 구축, 수요 측면의 이해, 보완재를 제공하는 파트너십, 사용자 친화성, 관련 법률과 정책의 명확성 등을 들 수 있다(표 3 참고)[4].

V. 데이터 공유를 위한 기술개발 과제

데이터 공유와 관련된 기술 솔루션은 일반적으로 데이터 공유의 위험 완화, 가치 제고 및 마찰 감소의 세 가지 범주로 나뉘는데, 각 범주별로 개발이 필요한 기술 솔루션은 표 4와 같다[5].

VI. 결론 및 시사점

보고는 기존의 정부 주도적 관점 및 전체 데이터 산업(경제) 활성화 관점에서 접근하는 하향식 접근을 수직적인 개별 산업 차원에서 구체적인 문제점 및 활성화 방안을 분석하고 이를 일반화하는 상향식 접근방법으로 보완하면 보다 실현 가능한 방안들이 모색될 수 있으리라는 기대에 기초하고 있다. 분석대상으로는 커넥티드카 및 자율주행차(ADAS

표 2 자동차 데이터 공유의 장애 요인

장애 요인	주요 내용
기술적 장애	<ul style="list-style-type: none"> • 데이터 저장, 대용량 데이터를 처리하고 최신 상태로 유지하는 능력과 관련하여 관련 비용 및 기술적 한계 • 데이터 표준화 및 상호운용성 문제 • 데이터 공유비용 문제: 클라우드 기반 플랫폼을 선택할 경우 데이터 보안을 보장하는 동시에 구조적 비용을 부담해야 함 • 데이터의 품질이 좋지 않거나 불충분하다는 한계 • 데이터 공유에 사용되는 플랫폼 아키텍처의 높은 복잡성 • 기술 솔루션 관련 신뢰 부족 문제
법적 장애 및 관련 문제	<ul style="list-style-type: none"> • 데이터 소유권 또는 데이터 세트로 합법적으로 수행할 수 있는 작업(즉, 데이터 사용)에 관한 법적 불확실성 • 공유 데이터 사용의 추적 및 통제의 어려움 • 데이터 공유의 데이터 보호에 대한 법적 요구 사항 이해 및 충족의 어려움 • 데이터 보호 요구 사항을 충족하기 위한 익명화 및 집계 보장에 추가적인 비용 발생 • 데이터 보호 및 개인정보 보호에 관한 새로운 규제의 특수성에 대한 지식과 명확성의 부족 • 특히 아시아권 국가에서 지도 데이터를 수집 및 공유하거나 판매를 제한하는 기존 규제 제한에 의한 데이터 현지화의 제약
스킬 부족	<ul style="list-style-type: none"> • 회사 내부 작업자의 데이터 기술역량 부족 및 보유한 데이터를 조작하고 이해하는 능력 부족
기타 장애	<ul style="list-style-type: none"> • 데이터 부정확성 또는 오용으로 인한 회사의 평판 훼손에 대한 두려움 • 적절한 라이선스 계약을 어떤 경우에 사용해야 하는지, 라이선스 계약을 준비하고 각 고객에게 맞게 조정하는 번거롭고 비용이 많이 드는 프로세스 및 계약상의 불확실성 • 데이터 공유로 인한 피해 발생 시 부담해야 할 비용 관련 불확실성 • 재무자원의 제약: 데이터 공유 활동을 추구하기 위해서는 초기투자 외에 지속적인 투자 필요 • 데이터의 잠재력에 대한 이해 부족: 데이터 사용자는 데이터가 비즈니스에 어떻게 도움이 될 수 있는지 알지 못하고, 이에 따라 데이터의 잠재력을 파악하지 못하는 경우가 많음

출처 Reproduced from [4].

포함)의 부상에 따라 자동차 생성 데이터가 기하급수적으로 축적되고 있는 자동차 데이터산업을 선택하였다. 특히 데이터 구축에서 데이터 활용으로 무게 중심이 이동되기 시작하면서 민간부문에서의 데이터 생태계 활성화를 위해서 반드시 필요하다고 거론되는 자동차 데이터 공유를 중심으로 조사 및 분석을 수행하였다. 분석결과에 따른 몇 가지 정책적 시사점은 다음과 같이 요약될 수 있다.

첫째, 정부에서 발표한 ‘데이터 플랫폼 활성화 방안’의 후속 시리즈로서 각 수직 영역별 데이터 플랫폼 구축 및 초기 단계 자생력 확보를 위한 지원이 필요하다. 2021년 6월에 발표된 ‘데이터 플랫폼 활성화 방안’ 등을 통해 국가 차원의 데이터 플랫폼 개선을 위한 큰 그림은 제시되었으므로 이제는 수직적 세부 영역별로 효율적 데이터 활용을 위한 데이터 플랫폼 구축 및 운영방안 모색이 필요하다고 할 수 있다[6,7]. 사회적 성과도 중요하지만 보다 가시적

인 경제적 성과를 거두기 위해서는 민간부문의 참여가 절대적으로 필요하며, 이와 관련하여 각 데이터 생태계 내에서의 데이터를 통한 수익 창출 및 데이터 공유가 핵심이슈로 등장하고 있다. 이에 따라 각 영역별로 시장원리에 따라 행동하는 민간 데이터 액세스 플랫폼인 데이터 허브의 조기 신설이 필요하다(해외에는 스타트업 형태의 다양한 자동차 데이터 허브가 이미 활발하게 활동 중인 데 비해, 국내에서는 아직 OEM 자체적으로 운영하는 데이터 플랫폼을 제외하고는 존재하지 않음).

둘째, B2B 데이터 공유의 필요성에 대한 자동차 생태계의 인식 제고 및 지원기구 설립이 요구된다. 모범적 사용 사례를 파악 및 공유하고 자동차 데이터 공유를 위해 필요한 기술 지침 제공 등을 지원하는 자동차 데이터 공유 지원센터 설립이 필요하다. 지원센터의 주요 임무는 데이터 공유의 혜택에 대한 인식을 제고시키고 데이터 공유를 통한 새로운

표 3 자동차 데이터 공유의 성공 요인

성공 요인	주요 내용
데이터 공급자와 데이터 사용자 간의 신뢰	<ul style="list-style-type: none"> • 데이터 사용자 및 공급업체와의 신뢰 구축은 가장 중요한 성공 요인이며, 신뢰를 성공적으로 구축하기 위한 주요 방안임 - 데이터 공유에 사용되는 API 등의 기술 메커니즘에 높은 보안 수준 보장 - 데이터 공급자에게 권한을 부여하고 데이터 세트 및 데이터를 공유하려는 회사에 대한 완전한 통제권 부여 - 데이터 공급자와 사용자 간의 소통을 가능하게 하고 신뢰할 수 있는 회사만 데이터 교환에 참여 허용: 데이터 공급자는 사용자가 데이터로 궁극적으로 수행할 작업을 이해한 이후에 데이터 세트 사용을 허용하며, 이를 위해서는 데이터 사용자와의 긴밀한 협력 필요 - 데이터 사용자가 데이터의 이점을 이해하고 데이터 세트를 사용하도록 지원 - 데이터 사용 조건 및 제한 사항을 설명하는 포괄적인 라이선스 계약서 작성 - 회사가 보유하고 있는 데이터가 합법적으로 공유될 수 있음을 데이터 사용자에게 설명(개인으로부터 제3자와 데이터를 공유할 수 있는 권한을 얻었음을 명확히 하는 것이 중요) - 정기적으로 회의 및 네트워킹 이벤트를 조직하여 데이터 공급자와 사용자 간의 직접적인 접촉 촉진
데이터 수요에 대한 명확한 이해	<ul style="list-style-type: none"> • 데이터에 대한 잠재적 관심이 있는 부문을 잘 이해하는 것이 주요 성공 요인 중의 하나임 - 기업이 보유하고 있는 데이터에 대한 구체적인 사용 사례를 제시하는 것이 유용(사용 사례는 데이터 공급자가 보유하고 있는 데이터에서 만들 수 있는 다양한 사용의 예를 말함) - 사용 사례는 잠재적인 데이터 사용자가 특정 회사의 데이터에 액세스하면 자신의 비즈니스에 어떤 이점이 있는 지 이해하는 데 도움이 되며, 이를 통해 데이터 공급자는 데이터 세트의 가치를 정의할 수 있음
기술적 보완성을 제공하는 파트너십	<ul style="list-style-type: none"> • 다른 회사 또는 다른 제3자(예: 연구기관 또는 협회)와 상호 보완적인 기술을 갖춘 파트너십을 구축하는 것도 데이터 공유 성공의 핵심 요소로 꼽힘 • 파트너십은 공식적인 계약 관계를 포함하거나 서비스 또는 편의의 교환으로 성립될 수 있으며, 다음과 같은 효과가 기대됨 - 데이터를 효율적이고 안전한 방식으로 제공하기 위한 올바른 기술 메커니즘을 구축함으로써 기술 투자 측면에서 상당한 절감 효과 가능 - 데이터 사용자의 요구 사항을 더 잘 이해 - 공유된 데이터를 사용하여 제품 또는 서비스를 공동 개발 - 기업 내에서 필요한 기술 구축 - 데이터 공유 활동을 위한 자금 마련
단순성과 사용자 친화성	<ul style="list-style-type: none"> • 데이터를 적절한 형식으로 교환하기 위한 간단하고 사용자 친화적인 메커니즘을 도입하는 것이 데이터 공유를 가능하게 하는 데 중요 • 데이터 사용자가 데이터를 이해하고 효율적인 방식으로 재사용할 수 있도록 하는 API(또는 SDK) 및 시각화 도구를 제공하여 데이터 공유 경험을 향상함 • 데이터 사용자가 데이터 플랫폼에서 얻을 수 있는 것을 쉽게 이해할 수 있도록 단순한 데이터 라이선스 요구사항 사용
명확한 법률 및 정책 프레임워크	<ul style="list-style-type: none"> • 개인 데이터 보호 및 지적 재산과 관련하여 명확한 법적 프레임워크의 제정이 중요 - 이러한 법적 프레임워크는 보유하고 있는 데이터로 합법적으로 수행할 수 있는 작업, 다른 사람에게 데이터를 제공할 수 있는 조건, 데이터 오용 시 의무를 집행하는 방법에 대한 확실성과 예측 가능성을 제공함으로써 참여 기업을 지원 - 궁극적으로 디지털 단일 시장을 통한 상품 및 서비스의 자유로운 이동과 관련된 내부 경계의 제거는 데이터를 수집하여 처리하고 다른 기업과 데이터를 공유할 수 있도록 하는 데 필수적
기타 성공 요인	<ul style="list-style-type: none"> • 모든 이해관계자에게 명확한 윈-윈 혜택을 창출할 수 있는 능력 - 데이터 사용자는 데이터를 활용한 성과 개선을 통해 데이터에 대한 대가 지불 유인을 갖게 되는데, 이는 OEM에게 데이터 공급자가 되고자 하는 인센티브를 제공 • 데이터 공유 경험을 시작할 때 적절한 자금 조달 기능을 갖추거나 제3자로부터 재정 지원을 받는 것도 하나의 성공 요인 • 사용 경험을 교환하고 지식을 공유하는 것은 데이터 공유 활동의 기반이 됨

출처 Reproduced from [4].

비즈니스 모델 개발 및 기업 간 데이터 공유를 촉진하는 툴킷 또는 지침을 제공하는 것이므로, 이를 통해 기업 간 데이터를 성공적으로 전송하고 액세스

하기 위한 권장 사항 제공, 표준 제정 및 공통 장애 요인 극복 등을 데이터 허브 등과의 긴밀한 협력하에 추진함으로써 국내에서 데이터 공유 및 재사용

표 4 데이터 공유를 위한 기술개발 과제

범주	기술 솔루션	주요 내용
데이터 공유의 위험 완화	데이터 침해 관련 기술 솔루션	<ul style="list-style-type: none"> • 데이터 수정 기술(Data Modification Technique): 데이터 무결성을 유지하면서 개별 데이터 요소 또는 전체 데이터 집합 변경 • 저장 및 전송 중인 데이터를 안전하게 유지하는 기술 • 블록체인과 같은 분산 원장 기술: 데이터를 분산된 방식으로 저장하고 공유 • 보안 계산(Secure Computation): 기본 데이터의 세부 사항을 공개하지 않고 분석을 가능하게 함 • TEE(Trusted Execution Environment): 칩의 암호화된 개인 영역 내에서 안전한 데이터 처리를 허용하는 운영 체제 및 별도의 하드웨어 모듈
	데이터 오공유 및 오용 관련 기술 솔루션	<ul style="list-style-type: none"> • API: 가장 널리 채택된 액세스 제어 형식으로 데이터에 대한 단일 정보 소스 역할을 하는 중앙 집중식 데이터 및 분석 플랫폼 제공 • 연합 학습(Federated Learning): AI 알고리즘은 이 기술을 통해 데이터 기여자가 보유하고 있는 분산 데이터를 이동하고 훈련할 수 있음(예를 들어, 민감한 환자 데이터를 공개하지 않고 다양한 병원 시스템의 데이터베이스에 보관된 이미지에서 암을 감지하는 기계 학습 알고리즘을 훈련하는 데 사용됨) • 합성 데이터(Synthetic Data): 개인정보를 공개하지 않고 원본 데이터 집합의 속성을 미러링한 다음 데이터를 파트너와 공유하여 알고리즘을 훈련하거나 소프트웨어를 테스트할 수 있음
	데이터 품질 위험 관련 기술 솔루션	<ul style="list-style-type: none"> • 데이터를 사용하기 전에 데이터의 출처, 신뢰성 등을 파악(예를 들어, 데이터 워터마킹은 제3자로부터 제공되는 데이터의 정확한 출처를 문서화할 수 있음)
데이터 공유 가치 제고	가치 식별 관련 기술 솔루션	<ul style="list-style-type: none"> • SI를 사용하여 데이터를 잠재적 사용 사례와 일치시키고 있음 • 데이터 판매자와 구매자를 연결하는 AI 솔루션을 구축함으로써 사용자가 자신의 데이터 관련 사용사례를 찾는 데 도움을 줄 수 있음
	가치 평가 관련 기술 솔루션	<ul style="list-style-type: none"> • SI는 전통적으로 데이터에서 통찰력을 도출하기 위해 사용되지만 데이터 자체의 가치 평가에도 사용될 수 있음 • SI를 활용하여 사용 가능한 품질을 구성하고 평가하기 위해 사용 사례 집합을 식별하면 기업은 데이터 공유의 대가로 받아야 하는 가치를 결정할 수 있음
	가치 획득 관련 기술 솔루션	<ul style="list-style-type: none"> • 소액 결제, 스마트 계약 및 구독 시스템과 같은 기능이 가치를 획득하기 위해 데이터 플랫폼에 내장되고 있음
데이터 공유의 마찰 감소	다크 데이터(Dark Data) 관련 기술 솔루션	<ul style="list-style-type: none"> • 데이터 공유 생태계에서는 기업의 다크 데이터를 발견하고 서로 다른 소스의 데이터를 집계해야 함 • 개별 기여자가 제공하는 다양한 데이터 소스(데이터 유형, 속성 및 시계열과 같은 차원에 걸쳐)를 통해 검색할 수 있는 데이터 시장 기술을 구축
	집계 복잡성(Aggregation Complexity) 관련 기술 솔루션	<ul style="list-style-type: none"> • 다른 소스의 데이터는 다른 정의, 형식 및 의미를 가지므로 이러한 문제를 극복하여 쉽게 이해하고 결합할 수 있도록 데이터를 식별, 정리 및 해석하기 위한 “데이터 랭글링(Data Wrangling)” 및 데이터 큐레이션(Data Curation) 도구 개발
	상호운용성 장벽 관련 기술 솔루션	<ul style="list-style-type: none"> • 이기종 데이터 유형과 데이터 소스의 보급으로 인해 데이터 표준 및 조화된 데이터 모델에 대한 투자 필요 • 데이터 세트에 컨텍스트와 의미를 부여하는 의미론적 메타데이터(Semantic Metadata)를 정의하고 데이터의 해석, 집계 및 분석을 용이하게 하는 레지스트리 구축
	전송 관련 기술 솔루션	<ul style="list-style-type: none"> • 공유의 기본은 데이터의 가치가 손상될 정도로 많은 오버헤드 없이 충분한 속도로 데이터를 원래 지점에서 의도한 용도로 전송할 수 있는 능력임 • API, 스트리밍 플랫폼, 연결 솔루션(5G) 및 Azure Data Share와 같은 클라우드 공급자 기능과 유사한 도구의 적절한 활용 필요

출처 Reproduced from [5].

을 위한 생태계 조기 확립이 가능할 것이다.

셋째, 데이터 공유의 핵심주체라고 할 수 있는 데이터 허브는 각 섹터별로 별도의 중립적 민간기업

이 운영을 주도하는 방식으로 추진되어야 한다. 아

직 OEM 등 자동차 관련 제조업체에는 데이터 중심문화가 미미하다고 볼 수 있으므로 데이터 거래

를 위한 중립적인 민간 전문기업이 운영주체가 되는 것이 효율적이다. 또한, 데이터는 일반적으로 동일 섹터 내에서 공유되는 특성이 있으므로 해외의 경우와 마찬가지로 자동차 생태계 전문 데이터 허브 구축이 필요하다(EC의 경우 유럽 오픈 사이언스 클라우드와 함께 모빌리티 데이터 공간 등 9개 분야에 대한 개별적인 데이터 공간 구축). 다만 궁극적으로는 섹터별 데이터 교환도 고려되어야 할 것이므로 장기적인 관점에서는 상호운용을 위한 데이터 공유 프레임워크의 사전적 검토도 필요할 것이다.

또한, 데이터 공유와 관련하여 개발이 필요한 기술 중 특히 정부 차원에서 기술개발 지원이 필요한 핵심 기술(솔루션)의 개발 지원이 필요하다. 앞에서 살펴본 바와 같이 데이터 공유의 위험 완화, 가치 제고, 거래상의 마찰 감소를 위해서는 다양한 기술 솔루션의 뒷받침이 필요하므로 국가 차원에서 확보가 필요한 핵심기술을 선정하여 개발을 지원할 필요가 있다. 현실적으로 표 4에서 제시된 데이터 공유를 위해 개발이 필요한 기술들은 대부분 기술수요기관인 개별 데이터 허브 차원에서 개발하기는 어려운 기술들이므로 국가 차원에서의 개발 추진 또는 지원이 필수적이다.

용어해설

Vehicle Data Hub(VDH) 다양한 데이터 제공업체로부터 차량 데이터를 수집하여 다수의 데이터 소비자가 사용할 수 있도록 하기 위해 자동차 데이터를 다수의 제3자와 교환하기 위한 플랫폼으로써 데이터 애그리게이터(Data Aggregator), 데이터 플랫폼, 데이터 포털, 데이터 거래소, 데이터 스토어 등 다양한 명칭이 사용됨

약어 정리

ADAS	Advanced Driver Assistance Systems
OEM	Original Equipment Manufacture

참고문헌

- [1] KPMG, Automotive Data Sharing, 2020. 10.
- [2] Car Data Facts EU, <https://www.cardatafacts.eu/risk-direct-access-car-data/>
- [3] everis Benelux, Study on Data Sharing between Companies in Europe, Final Report Prepared for the European Commission DG Communications Networks, Content & Technology, 2018.
- [4] fka & Strategy Engineer, Monetizing Automotive Data: A Perspective on How Data Drives Revenues and Growth in Future Automotive Business Models, 2016. 4.
- [5] Massimo Russo & Tian Feng, The New TECH Tools in Data Sharing, BCG. 2021. 3.
- [6] 관계부처 합동, "데이터 산업 활성화 전략," 2018. 6.
- [7] 관계부처 합동, "민·관 협력 기반 데이터 플랫폼 발전전략," 2021. 6.