MOBI

Building the
New Economy of Movement

March 2021

CONNECTED MOBILITY DATA MARKETPLACE

BUSINESS WHITE PAPER

MOBI CMDM0001/WP/2021
Version 1.0

INTRODUCTION

The Mobility Open Blockchain Initiative Connected Mobility Data Marketplace Working Group is a global, multi-stakeholder project working to co-design standards based on blockchain and distributed ledger technologies for connected mobility ecosystems. The project engages stakeholders across the transportation value chain, including vehicle manufacturers, technology solution providers, governmental, and non-governmental entities. This report is based on numerous discussions, workshops, and research. Opinions expressed herein do not necessarily reflect the views of individual members.

Sincere thanks are extended to those who contributed their unique insights to this report.

Authors

Chris Ballinger, MOBI
Roger Berg, DENSO
Benjamin Diggles, Constellation Network
Griffin Haskins, MOBI
Derek Pierre, NuCypher
Robin Pilling, MOBI
Tram Vo, MOBI

CMDM Working Group Co-Chairs

Roger Berg, DENSO
Michal Filipowski, General Motors

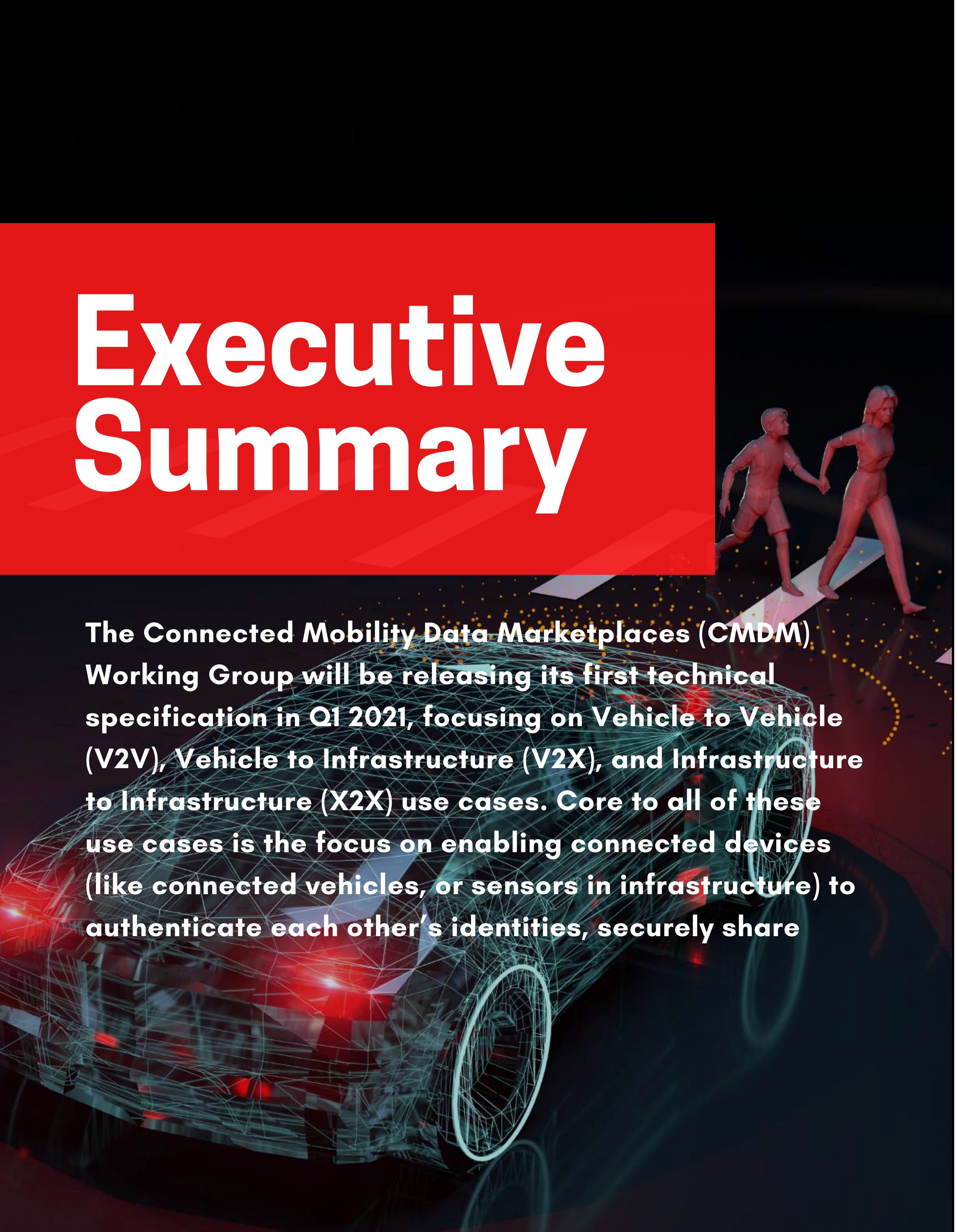
CMDM Working Group Team Members

Loren Adams, Accenture
George Ayres, IBM
Tim Bos, ShareRing
Joshua Cartellone, Accenture
Benjamin Diggles, Constellation Network
Arjun Hassard, NuCypher
Thomas Hulse, Fifth9
Jerry Kim, AMO Labs
Suresh Kumar Sundararaj, Ford
Duy Linh Nguyen, Cognizant
Chalid Mannaa, Ocean Protocol
Jim Mason, DMX
Pramita Mitra, Ford Motor Company
David Noack, Continental
Evripidis Paraskevas, GM
Manan Patel, Ocean Protocol
Derek Pierre, NuCypher
Anthony Salamone, RouteOne
Liu Shaowei, CPChain
Sudha Sriram, Ford
Sukesh Tedla, Swedish Blockchain Association
Wanda Wang, Toyota Insurance Mgmt Solutions (TIMS)
Chris Wood, Filament

MOBI Team

Chris Ballinger, Founder + CEO
Tram Vo, Founder + COO
Rajat Rajbhandari, Working Groups Lead
Robin Piling, Technical Lead
Griffin Haskins, Fellow
Kelly Clark, Communications Manager

TABLE OF CONTENTS



Executive Summary

The Connected Mobility Data Marketplaces (CMDM) Working Group will be releasing its first technical specification in Q1 2021, focusing on Vehicle to Vehicle (V2V), Vehicle to Infrastructure (V2X), and Infrastructure to Infrastructure (X2X) use cases. Core to all of these use cases is the focus on enabling connected devices (like connected vehicles, or sensors in infrastructure) to authenticate each other's identities, securely share data, and securely record transactions. The Standard defines a robust framework for a system that provides the above functionalities, in support of the development of those use cases within the mobility ecosystem.

> The CMDM Standard provides a rich set of digital infrastructure, providing frameworks for identity, permissioning, and more, with the goal of achieving interoperability between siloed mobility services.

The CMDM Standard prescribes a set of logical schemas to support interoperable data sharing between connected devices, and provides a new certificate for a connected sensor identity. Moreover, it provides a rich set of digital infrastructure, providing a framework for identity, permissioning, and more. These frameworks are intended to achieve interoperability between the "walled garden" mobility ecosystems that traditionally have an entirely siloed operation. Integrating these systems allows for many new solutions that provide operational efficiencies, superior insights, new revenue streams, and much more. The CMDM Working Group was acutely aware of this, which is why the CMDM Standard does not focus on any one particular application, rather, it ensures that appropriate infrastructure is in place to support organizations in creating their own applications within those use case areas.

> This document will investigate the impact of the CMDM Standard by overviewing some use cases that are supported by the Standard, broken down into their expected time-to-market.

This business review document begins with an overview of the current state of affairs with regards to mobility ecosystems and data exchange within the automotive industry. Then, the CMDM Standard itself will be described in greater detail. The document will also investigate how the Standard enables some example use cases on short, medium, and long term time horizons. Particularly, the Working Group defined short term use cases as being implementable within five years, while medium term use cases become feasible after five years, and long term use cases are likely to take more than ten years to reach scale. Ultimately, this document evaluates the use cases considered by the CMDM Working Group and discusses the value of the Standard supporting such use cases in real world settings.

Data Economy

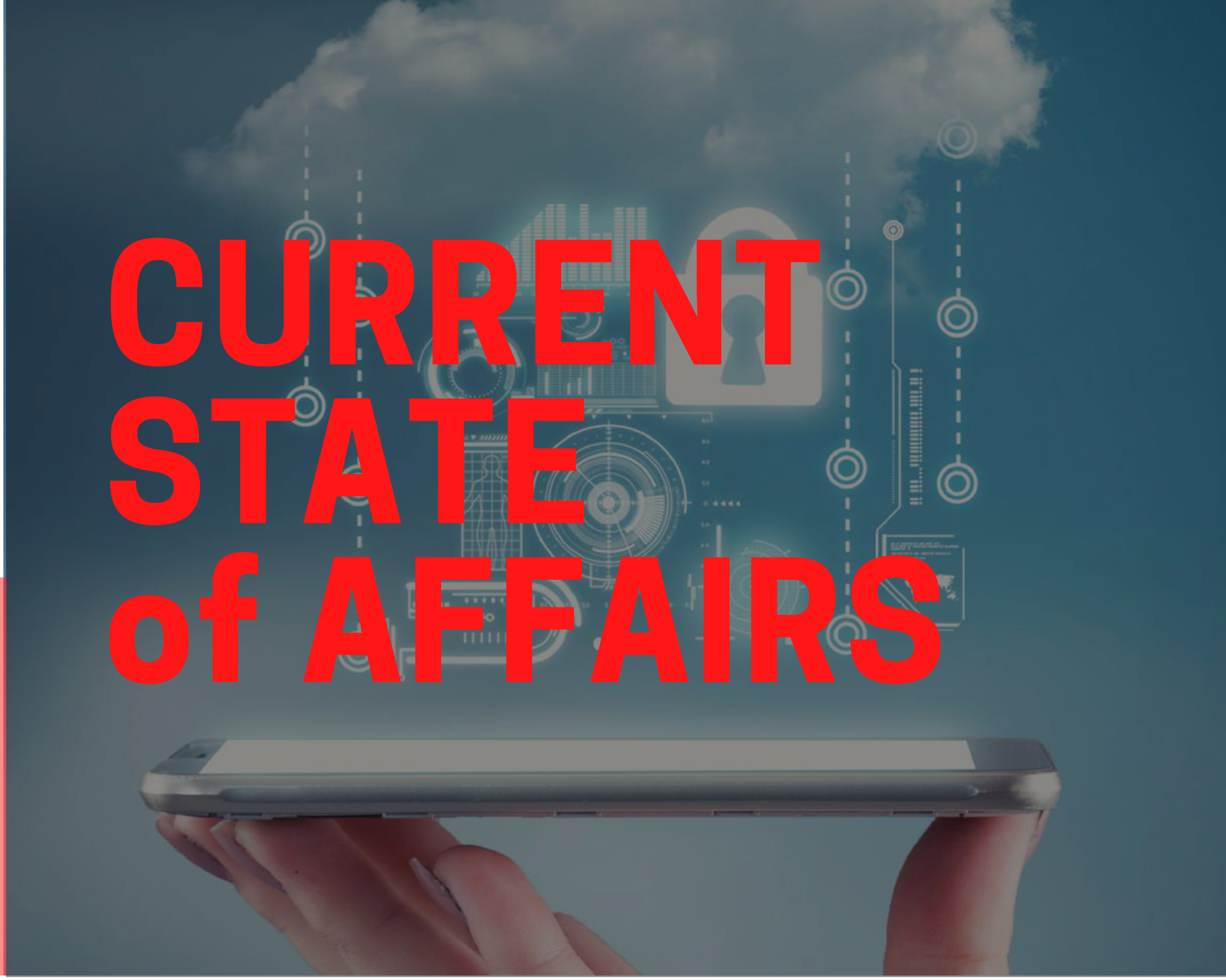
Data is digital, machine readable information that represents a foundational layer of wisdom in the digital age; a fundamental driver of the economy, almost akin to oil.

"Data is the new oil"[1] – arguably one of the most cited quotes in conjunction with digitization that portrays the ongoing shift into the wisdom age[2] and the corresponding economic mechanisms. The scope of this paper cannot stray too much into the influencing factors. Yet, it will test certain assumptions. There are three key assumptions that are important to understand the data economy and its importance:[3]

- The importance of physical products and related services for unique value propositions and revenue streams is declining, as they increasingly turn into commodities.
- Data is the foundational layer of wisdom in a digital environment, which will be detailed further in the subsequent passages.
- Data has an underlying monetary value based on derived knowledge – in this case, increased knowledge of users results in a better service and product experience.

Increasing complexity, the pace of change, and inherent determinants cause business leaders to actively seek opportunities that allow them to dominate their markets in the wisdom era. The wisdom economy has strong interdependencies with growing digitalization in the 21st century. Data represents digital, machine-readable information, from which knowledge, wisdom, and insights may be extracted.

1. "The World's Most Valuable Resource Is No Longer Oil, But Data.", Economist, May 2017 <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>
2. John Findlay, "The Wisdom Economy.", thewisdomeconomy.blogspot.com, August 2011, <http://thewisdomeconomy.blogspot.com/> [Accessed 1 January 2021].
3. Kiran Bhageshpur, "Data Is The New Oil -- And That's A Good Thing", Forbes, November 2019, https://www.forbes.com/sites/forbestechcouncil/2019/11/15/data-is-the-new-oil-and-thats-a-good-thing/#4d1da4dd7304



Economic power is derived from data, but that process is complex. Turning big data into actionable insights, which is how value is actualized for big data, requires ample computing power and supporting infrastructure.

Hence, they form a foundational layer of wisdom in the digital age. In this environment, the ever-growing amount of information is increasingly digested by machines. Humans must rely on computing power to be able to consume the vast volume of information and turn it into actionable insights. The big data movement implies that more consumable data translates into more knowledge and hence more economic power.[4]

4. "Big Data, What it is and why it matters", SAS, accessed January 2021, https://www.sas.com/en_us/insights/big-data/what-is-big-data.html; "Big Data Analytics", IBM, accessed January 6, 2021, https://www.ibm.com/analytics/hadoop/big-data-analytics

Pricing data is highly complex and difficult for two core reasons - firstly, there is no universal marketplace for data, and secondly, most datasets are not comparable.

This is where the data economy comes into play. Economic power directly derives from monetization of a product, which is data. In that context, the basic components of any market theory and corresponding mechanism apply. A marketplace as the basis for any economy requires a buyer, seller and price for the product, which is normally determined by supply and demand. Unfortunately, finding a fair market price for data is highly complex, as many influencing factors are involved. At its core, this is based on two facts: Firstly, there is no universal marketplace for data. Secondly, most datasets are not even remotely comparable.

Oil and data do have different market mechanisms, and pricing data is quite complex and difficult, which is fundamentally due to the fact that data is a non-excludable and nonrival good.

While data may be seen as "the new oil," it's clear that they act quite differently within an economic context. Oil as a raw material is traded on a global marketplace, which is not the case for data (as of now). While oil can be clearly quantified and priced (e.g. per barrel), this is not true for data: 1kb of raw data does not always have the same value. There is no generally applicable measuring unit.[5] Unlike oil, data has two characteristics which inhibit the smooth functioning of private markets. Data is "non-rival" in the sense that one user's consumption does not preclude another's consumption of the same data. Data is also often "non-excludable," meaning that once data is public, it isn't feasible to charge for use. These sources of market failure are both a challenge and an opportunity when considering new blockchain-based data markets.[6]

5. "Data Economy: Radical transformation or dystopia?", United Nations, January 2019, https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/publication/FTQ_1_Jan_2019.pdf; Jennifer Belissent, PhD, "Help Wanted: Data Innovation For The Data Economy", Forrester, October 2017, https://go.forrester.com/blogs/help-wanted-data-innovation-for-the-data-economy/; "Keynesian Economics", Investopedia, April 2020, https://www.investopedia.com/terms/k/keynesianeconomics.asp; Jason Fernando, "Law of Supply and Demand", Investopedia, November 2020, https://www.investopedia.com/terms/l/law-of-supply-demand.asp

6. D. Bergemann and A. Bonatti. "Markets for information: An introduction", Annual Review of Economics, November 2019.

Data Value Chain

To get a better understanding of the data economy, a closer look at the data value chain is key. At its core, raw data provides little value. Given that a lot of literature suggests a great demand for data, this could be considered a bold statement: market mechanisms imply that a great demand makes something valuable and drives the market price up.

Raw data does not have much value, but much of the literature suggests a high demand for data. Accessing and storing raw data can be difficult and expensive. Limited accessibility implies high demand, and high prices.

Looking at the demand side, it is clear that not all market participants have the same prerequisites and access to raw data might be limited. Obtaining and storing raw data can be hard, often requiring dedicated efforts and resources – think of consumer data, production data or sensor data, among others. With limited accessibility, high demand normally constitutes high prices.

Yet, this only holds true if the actual supply is limited, resulting in a scarcity of the raw material. In reality, quite the opposite is true. Compared to oil, which is a limited resource and can only be consumed once, there is a plethora of data out there, creating an abundance of almost unlimited supply. The principles of supply and demand indicate that an oversupply lowers the market price.

So why is data still considered valuable and how is a price established? The analogy to oil can be used again: raw data resemble crude oil, referring to a resource that, without refining, is useless. Data needs to be refined to become valuable. This refinement process includes structuring, analyzing and combining different datasets to turn them into actionable information and services. Hence, data is most valuable when it's

Current State of Affairs

Like oil, raw data must be refined and prepared in order to create value by being used in applications and solutions. The process that raw data takes, from generation, to data cleaning and preparation, and finally ending with usage in an application that actualizes the data's value is called the data value chain.

an integral part of a customer facing solution – the end of the value chain.[7]

If a company wants to find out if raw data is needed and where/how it should buy it, the value chain has to be applied backwards. First, the final offering must be clearly defined. Based on this definition, the chain can be worked backwards to determine what raw data is required to generate the final product. As a positive side effect, by starting at the service and actual application layer, there is a clear definition of the user and the value that should be generated.

With a concept of the final product as the starting point, the company can determine what is needed to create it. In most cases, the core problem is not the availability of raw data but the clear definition of a marketable service. It is vital to understand that in most cases buying data does not resemble the purchase of an end product, but raw materials or components that are needed to manufacture a service.

In a metaphorical way, the data value chain follows the principle of WHY – HOW – WHAT. The definition of concrete sets of raw data (WHAT – the data input) comes last. Simply generating and hoarding data as the holy grail of economic

7. Rado Kotorov, "The Data Value Chain: Steps for Monetizing your Data", Integration, accessed January 2021, https://www.idevnews.com/stories/6998/The-Data-Value-Chain-Steps-for-Monetizing-Your-Data; "The Data Value Chain Executive Summary", Open Data Watch, accessed January 2021, https://opendatawatch.com/reference/the-data-value-chain-executive-summary/; Rainer Sternfeld, "Designing a Better Planet with Big Data and Sensor Networks", Intelligent Sensor Networks Conference, November 2014,https://www.slideshare.net/rainer_sternfeld/designing-a-better-planet-with-big-data-and-sensor-network-for-isn-conference-philips-high-tech-campus; Sylvain Santamarta, Rash Gandhi, and Michael Bechauf, "Big Oil, Big Data, Big Value", Boston Consulting Group, August 2019, https://www.bcg.com/publications/2019/big-oil-data-value.aspx; Edd Dumbill, "Understanding the Data Value Chain, November 2014, https://www.ibmbigdatahub.com/blog/understanding-data-value-chain

Successfully extracting value from data means having a clear picture of how that data will be used in an application. Defining first WHY data is needed, and HOW that data will be obtained and processed is necessary before ever moving onto WHAT is needed, or the raw data itself.

success is a dangerous fallacy. The first step is knowing which service should create value for whom (WHY – the vision) and success is a dangerous fallacy. The first step is knowing which service should create value for whom (WHY – the vision) and how the input should be transformed into this service (HOW – the refinement process). Otherwise data collection will only translate into having a warehouse full of raw materials with no production process or final product to use them.[8]

Bridging the gap back to the wisdom economy, data only represents the foundational layer to wisdom and decisions – or in economic terms – the final product/service.[9]

Data Marketplace and Exchange in the Automotive Industry

Vehicles are increasingly Connected, Autonomous, Shared, and Electric, which together provide the core paradigm shifts driving mobility's shift into the digital age.

The automotive industry is undergoing a fundamental change that redefines market mechanisms and the roles of players in the value chain. While a magnitude of macroeconomic and microeconomic factors influences the ongoing shift, the most relevant are summarized under the CASE acronym. CASE stands for Connected, Autonomous, Shared and Electric and was mainly framed by Daimler.

8. Dave Chaffey, "Golden Circle model: Simon Sinek's theory of value proposition, start with why", Smart Insights, October 2020, https://www.smartinsights.com/digital-marketing-strategy/online-value-proposition/start-with-why-creating-a-value-proposition-with-the-golden-circle-model/; Peter Götz, "Why – How – What: From Product Vision to Task", Scrum.org, March 2018, https://www.scrum.org/resources/blog/why-how-what-product-vision-task; Chuck Frey, "Customer Centric Innovation", Innovation Management, April 2004, https://innovationmanagement.se/imtool-articles/customer-centric-innovation/

9. Rob Lokers, "Analysis of Big Data technologies for use in agro-environmental science", ResearchGate, October 2016, https://www.researchgate.net/figure/DIKW-hierarchy-from-Big-Data-to-decision-making-for-societal-challenges_fig1_307874029; Ayan Brahmachary, "DIKW Model: Explaining the DIKW Pyramid or DIKW Hierarchy", CertGuidance, May 2019, https://www.certguidance.com/explaining-dikw-hierarchy/

Current State of Affairs

This acronym gives a great overview of the most influential determinants of the shift, which are interconnected with each other as well as surrounding trends.[10]

Connected

Vehicles are increasingly connected with everything (V2X communications). This includes, but is not limited to: drivers, infrastructure, other vehicles, and third-party service providers. Communication is based on connectivity (e.g. 5G networks) and a constant data exchange between vehicles and the mobility ecosystem.[11]

Autonomous

The vision of self-driving vehicles is gradually becoming a reality. Based on the SAE automation levels ranging from zero (no automation) to five (full automation), current vehicle generations are between level three and four. Automation is based on different sensors, processing power, algorithms, machine learning/AI and the ability of vehicles to exchange data with other actors in traffic (machine or human).[12]

Shared

The deeper meaning of shared is a change in the concept of ownership. It subsumes car-sharing, ride-hailing, public transport and other fleet services that do not constitute private ownership and sole use of a vehicle by a single person.

10. "CASE – Intuitive Mobility", Daimler, accessed January 2021, https://www.daimler.com/case/en/

11. Charles McLellan, "What is V2X communication? Creating connectivity for the autonomous car era", ZDNet, November 2019, https://www.zdnet.com/article/what-is-v2x-communication-creating-connectivity-for-the-autonomous-car-era/

12. "Automated Vehicles for Safety", NHTSA, accessed January 2021, https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety

Vehicles can still be privately owned by one consumer but shared with other riders for a fee (see: Tesla robotaxi, in conjunction with autonomous driving). Arguably, shared mobility represents not only a technological shift but also a change in mindset of how the overall concept of mobility is perceived – which has the potential for far-reaching implications on value chains and the role of established actors. From a technological viewpoint, shared concepts demand data exchanges between involved parties inside and outside the ecosystem.[13]

There are four key factors driving mobility's shift into the digital age - Connected (connectivity in vehicles), Autonomous (autonomous vehicles and agents), Shared (sharing economy), and Electric (electric vehicles and eMobility applications).

Electric

Electric stands as a synonym for engines that are not based on fossil fuels and includes hybrid drives, purely battery-powered vehicles, fuel cells, and alternative synthetic fuels. For all systems, new infrastructural solutions are mandatory and require data to be exchanged between vehicles, passengers, and the infrastructure (e.g. smart grid and charging points).[14]

All of these major drivers have one thing in common: they require data as the fuel for future mobility ecosystems. However, the complexity of a mobility ecosystem is immense, and a functioning ecosystem requires prospering data exchange and collaboration between all actors.

13. Kirsten Korosec, "Tesla plans to launch a robotaxi network in 2020" TechCrunch, April 2019, https://techcrunch.com/2019/04/22/tesla-plans-to-launch-a-robotaxi-network-in-2020/; Anne Grosse-Ophoff, Saskia Hausler, Kersten Heineke, and Timo Möller, "How shared mobility will change the automotive industry", McKinsey & Company, April 2017, https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/how-shared-mobility-will-change-the-automotive-industry; Tobias Schönberg, "Shared Mobility", Roland Berger, July 2014, https://www.rolandberger.com/de/Publications/Shared-mobility.html

14. "Alternative Fuels and Advanced Vehicles", US Department of Energy, accessed January 2021, https://afdc.energy.gov/fuels/; "Alternative Fuels", US Department of Energy, accessed January 2021, https://www.fueleconomy.gov/feg/current.shtml; "The Smart Grid", US Department of Energy, accessed January 2021, https://www.smartgrid.gov/the_smart_grid/smart_grid.html

Current State of Affairs

Mobility Ecosystems

By definition, a mobility ecosystem encompasses all modes of transportation, a multitude of brands, single entities, subsystems that provide transportation, the infrastructure itself (roads, traffic signs, electric grid), other supporting actors, and ultimately, the consumers that require mobility. These moving parts are all interconnected and interdependent.[15]

Mobility ecosystems encompass all modes of transportation, entities, groups, companies, subsystems that provide transportation, infrastructure itself (roads, highways), and everything else that is used or participates in mobility.

Normally, a mobility ecosystem includes all aspects and functions that provide transportation, be it short-, mid- or long-distance. Yet, a holistic scope on mobility that connects all dots becomes extremely complicated. Various modes and providers of transport must be taken into account, different localities and infrastructures in one country have to be aligned. Internationally, even multiple languages, different time zones, and currencies must be considered.[16]

Imagine a city as a manifold: different components play together like gears in a well-oiled machine to form the overall concept of the city. This metaphor is also true for an urban mobility ecosystem. The foundational layer is formed by the urban population that needs transportation for daily routines and leisure. It can choose between different means of transport,

15. "Mobility Ecosystem: A wave of transformation", Moveco, July 2018, https://medium.com/@moveco.io/mobility-ecosystem-a-wave-of-transformation-c8f2eb7af658; Scott Corwin, Philipp Willigmann, "The evolving mobility ecosystem", Deloitte, accessed January 2021, https://www2.deloitte.com/us/en/pages/manufacturing/articles/the-evolving-mobility-ecosystem.html

16. Ralf Baron, Michael Zintel, Niklas Schemken, Christoph Uferer, "Integrated Mobility Platforms", Arthur D Little, February 2018, https://www.adlittle.de/en/insights/viewpoints/integrated-mobility-platforms; Shannon Bouton, Stefan M. Knupfer, Ivan Mihov, Steven Swartz, "Urban mobility at a tipping point", McKinsey & Company, September 2015, https://www.mckinsey.com/business-functions/sustainability/our-insights/urban-mobility-at-a-tipping-point

depending on the situation and reason for traveling: A privately owned car, public transport, bicycle or shared mobility offers with several types of vehicles. Most journeys even combine at least two or more of these means. They can be best summarized as the vehicle layer.

In a mobility ecosystem, the car, bicycles, taxis, trains, micro mobility options, motorcycles, etc. are all referred to, collectively, as the vehicle layer. The bridges, roads, electric grid, and other capital is referred to as the infrastructure layer. The drivers and travelers (users) are referred to as the consumer layer.

Aside from the vehicle layer, there is also the infrastructural layer, including bridges, roads, the electricity grid, and traffic management, among others.[17] This infrastructure layer may communicate with entities within its layer or with the vehicle layer.

From a data viewpoint, there are data silos on every aggregation point. Each consumer has their own set of devices, applications, and services they use. We call this the personal network, involving the individual datasets every inhabitant owns and shares. On the vehicle layer, every entity has its own siloed data network that incorporates routing, booking or payment services, among others. Examples include the data networks of buses, train and taxi providers, car-sharing fleets, or the data-driven services of every vehicle manufacturer. These data silos can intertwine within the vehicle layer. For example, if a car-sharing fleet uses cars of multiple makes, every car and car maker has their own branded network and portfolio of services, which are then combined in symbiosis within the network of the fleet operator (e.g. for telematics data). Aggregation of silos happens on various levels and in multiple combinations.

By exchanging data between the consumer and vehicle layer networks, mobility services are created. These offers are connected to the infrastructural layer that has its own set of

17. "How many times have you heard - "Automakers are sitting on a data goldmine"?" SBD Automotive, accessed January 2021, https://www.sbdautomotive.com/en/news-automakers-goldmine

Current State of Affairs

Within the vehicle and infrastructure layers exist data silos - consider the data networks of buses, car-sharing fleets, vehicle telematics, roadway sensor systems, and more. To expand the quantity and quality of mobility services, data must move between silos and aggregated on a superseding layer, called the ecosystem layer.

secluded data networks required for traffic management, telecommunications, or energy supply within the city. A practical example can be given with Tesla's (and other OEMs) charging services that allow the consumer to search for available charging points based on their location and incorporate charging dynamically into routing services. Especially regarding electric and autonomous vehicles, an open data exchange between the infrastructural, vehicle, and even consumer layer is mandatory.[18]

To allow for a functioning ecosystem, data must transcend between clusters and be aggregated on a superseding layer, described as the ecosystem layer.

Approaches and Problems of Data Marketplaces in the Mobility Industry

Collaboration and coordination between numerous players is necessary to create that ecosystem layer that enables a transition into a "Mobility as a Service(MAAS)" world.

Numerous players need to collaborate in the interwoven mobility ecosystem to enable the transition into a "Mobility as a Service (MaaS)" world. A term that is often used in this context is coopetition – cooperation on the network level and competition on the service level. Information and data has to transcend organizational boundaries and must be shared between multiple organizations and even industries.[19] Since most players pursue

18. Ricardo Bacellar, "What will the future mobility ecosystem look like?", KPMG, November 2017, https://assets.kpmg/content/dam/kpmg/br/pdf/2017/11/br-kpmg-congresso-sae.pdf; "Charging Your Car", Tesla, accessed January 2021, https://www.tesla.com/support/charging

19. Warwick Goodall, Tiffany Dovey Fishman, Justine Bornstein, and Brett Bonthron, "The Rise of Mobility as a Service", Deloitte, January 2017, https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/consumer-business/deloitte-nl-cb-ths-rise-of-mobility-as-a-service.pdf; Will Kenton, "Understanding Coopetition", Investopedia, May 2020, https://www.investopedia.com/terms/c/coopetition.asp

divergent goals or compete on the service layer, data exchange should be orchestrated in the form of a mediating marketplace. The competition, then, is rooted in the insights derived from the data, such as algorithms, that lead to optimized and high-quality products and services.

Although it's a massive challenge to build a universal marketplace for mobility data, there are a number of initiatives and platforms focused on it. Moreover, most of the important large players are pushing for the creation of mobility data exchanges. Understanding the pain points that hinder their creation is key.

Even though it is a massive coordinative challenge to implement such a universal marketplace, multiple platforms and initiatives have emerged over the recent years to tackle the problem. And within certain secluded segments, there have been successful examples, mainly driven by OEMs. Yet, all of these efforts still lack traction and there has been no global breakthrough. Considering that all actors can logically conclude that an open exchange of mobility data is the foundational layer for meaningful services and future offers, the question is: what stops them from just doing it?

To get a deeper understanding of pain points and why mobility data exchanges are struggling to gain wider traction, existing platforms have been inspected and related market studies have been analyzed. Overlaying the findings, five major pain points that hinder open mobility data exchanges were identified.

Current State of Affairs

A consensus on data access rights, roles, tasks, and overall framework is critical to the development of an open data exchange.

1. Data Sharing Standards

A universal shared data framework for the mobility sector still needs to be defined. It is less about technical hurdles – a common rule is that structured data in machine-readable formats are the preferred choice. The bigger issues are access rights and format as well as the infrastructure of datasets. Players closely guard the data for their own use and treat their inventory as proprietary. A consensus for data access rights and the degree of openness is required to establish a functioning data exchange.

A totally open exchange would support innovation and excel transitional speed of the ecosystem by allowing data-based services to prosper. What becomes apparent is that all initiatives suffer from an unclear distribution of roles and tasks: who decides what is public and what remains private, who manages access rights, who supplies, and who consumes.

Data quality is commonly determined by many factors, including its accuracy, completeness, consistency, validity, integrity, and currency.

2. Quality of Data

Mobility services depend on up to date, reliable, and adequately complete data within the ecosystem. Especially for applications needing real-time data, the half-life is extremely short. A lot of data must be updated daily, hourly, or, in some cases, every second. If a service provider wants to acquire datasets, they need to know that he receives quality – reliable sets without corrupt data. In that manner, quality does not only involve refresh rates, but also quality of the supplied datasets, referring to associated data points, format, and structure.

Usually, data quality is assessed along the following dimensions: accuracy, completeness, consistency, validity, integrity, and timeliness/currency. However, without a central governing body and given the magnitude of potential use cases for data in a mobility ecosystem, a universally applicable consensus on the level of quality can be hard to define.

Without a shared understanding and framework for valuing data and measuring its quality, determining an agreed value for data is extremely difficult.

The perception of quality in data is thus largely built on trust. Without knowing if a source can be trusted, a mediating trusted third party or product samples that give the consumer the chance for a product trial are required, assuming the data supplier wants to hand out freebies since it reinforces a freeloader syndrome: taking the data, without sharing their own.[20]

3. Monetization of Data

Considering the data value chain, data suppliers have no interest in sharing their information for free. Companies with access to raw data hoard it and have no natural interest in sharing, thereby creating data silos. Data owners are afraid to openly share or sell their data because they risk missing out on more valuable applications or services someone else could build with them.

Even if the provider is willing to sell its data, there is no universal measuring unit or definition for the value of data, as it is mostly based on perception and individual motives. Every dataset is different and for maximum use and interoperability, raw data need a level of normalization. Establishing a fair price for datasets still typically require negotiations in individual deal

20. Christian Bussmann, Magdalene Piec, Dr. Andreas Gutweniger, "MaaS and Data Sharing: Five Issues Closely Related to Practice", Detecon Consulting, April 2016, https://www.detecon.com/en/knowledge/maas-and-data-sharing-five-issues-closely-related-practice; "Data Done Right: 6 Dimensions of Data Quality", SmartBridge, accessed January 2021, https://smartbridge.com/data-done-right-6-dimensions-of-data-quality/

Current State of Affairs

flows between buyer and seller. Yet, data buyers are hesitant to buy, because accessible data might already have been tampered with and there is no comparative market price.

Data sellers are hesitant to sell their data, because that data could be used to power a new service or product whose profits they want to capture. Data buyers are hesitant because there is no comparable market price. Consequently, there is a lack of monetary incentives for data sharing.

It lays within the nature of private companies to be profit-oriented. Hence, sacrificing potential business models for the greater good is against their nature. Enforcing this through public organizations and governmental regulations could be a distortion of competition. In addition, processing and making data available in good time simply costs time and money, so free sharing is a calculatory loss for the provider.

Consequently, there is a lack of monetary incentives in sharing data. The only non-monetary incentive is the exclusive access to other mobility data through sharing, similar to peer-to-peer file sharing. However, this only works if there is a consensus on open sharing and everyone plays by the rules. This is also where the network effect comes into play, as the value of the exchange increases with active participants – so who will make the first move?

4. Ownership Models

Ownership and profit models around data sharing are unclear, resulting in friction between data sharing entities.

Control and ownership of aggregated data result in economic power. Hence, most actors pursue assuming the role of integrators and aggregators – in other words: mobility service providers. As a result, other actors would be degraded to data suppliers missing out on the monetization opportunities in the service layers. Naturally, this creates conflicts of interest. On another note, legal transfer of data ownership can be an issue if personal data based on previous opt-ins are involved.

5. Security

There are two major concerns that are closely linked to each other: protection against malicious attacks (cyber security) and data privacy. The data exchange has to be secured at the actor and network level to provide the necessary level of protection and prevention of a single point of failure. Especially in combination with autonomous driving, the security of data transfers is a top priority.

Liability from potential data breaches or exposure of PII makes data generators reluctant to share data, and so without robust security, the potential liability of data sharing makes the endeavor a calculatory loss.

Another factor is the protection of personally identifiable data: a unified approach to protect consumer data is required, which involves legal, technical, and ethical considerations. Pseudo-anonymous data can be reengineered to reveal identities, e.g. by cross-referencing location and other datasets. Reverse engineering of personalized data must be prevented in the design of the system architecture to comply with data protection laws. In addition, consent structures have to be set up for personal data or datasets have to be anonymized to a degree that does not require individual consent of end users.[21]

21. Regina Clewlow, "The Opportunity To Reshape Cities With Shared Mobility Data", Forbes, October 2018, https://www.forbes.com/sites/reginaclewlow/2018/10/10/the-opportunity-to-reshape-cities-with-shared-mobility-data/#5cb5ff3e617f; Nick Ismail, "The new front line of cyber security is mobility", Information Age, March 2018, https://www.information-age.com/front-line-cyber-security-mobility-123471096/; Leon Nash, Greg Boehmer, Mark Wireman. Allen Hillaker, "Securing the future of mobility" Deloitte, April 2017, https://www2.deloitte.com/us/en/insights/focus/future-of-mobility/cybersecurity-challenges-connected-car-security.html

THE CMDM STANDARD

Modern vehicles, trains, bikes, and scooters are increasingly internet-connected and intelligent. Even infrastructure like highways, bridges, and ferries are being endowed with sensors, telematics, and secure identities to generate and share data. Truly exciting services become achievable when these devices are able to seamlessly and easily share data with each other, with the necessary shared functionality guaranteed. These abilities are called data interoperability and functional interoperability; both are difficult to achieve at scale. In mobility ecosystems today, there is a dearth of such interoperability, across both the vehicle and infrastructure layers, resulting in the natural formation of silos at each aggregation point. Individual Mobility Service Providers operate entirely within their own silos, limiting potential value creation from new services by preventing the use of datasets across multiple silos. Collaboration between data owners necessitates costly integrations due to incompatible datasets, functionalities, etc.

Connected devices, and the success of markets that monetize those devices' data, are defined by the degree of interoperability that exists in the ecosystem. Indeed, every use case involving connected devices relies on secure, scalable data sharing. Blockchain provides a secure avenue for these stakeholders to authenticate each other's identity and to immutably record/securely expose their collected information to other stakeholders. The immutability and shared governance of distributed ledger systems provide a necessary, foundational layer for creating interoperable applications on top. However, breaking down these silos requires more than just technology – it requires community and coordination.

The CMDM Standard

Exciting new services and applications are made possible when different silos in the vehicle and infrastructure layers are able to seamlessly share data between themselves, with the necessary shared functionality guaranteed. These features are called data interoperability and functional interoperability.

Interoperability can only be achieved when there are common standards for core infrastructure – identity, data, and functions. The CMDM Standard prescribes core services, certificates, and logical schemas for the sharing of data between connected vehicles and the connected infrastructure they drive through. In other words, the CMDM Standard is a key step towards achieving data interoperability and functional interoperability in mobility ecosystems. This allows data to smoothly flow between these silos, which enables a rich variety of applications.

The logical schemas prescribed by the CMDM Standard define the set of attributes, data formats, and their associated standards for data-generating devices to register an "event." An "event" is anything that results in the generation of data – from sensors in an engine detecting a part failure, to a vehicle detecting debris on the roadway, or even the detection of high concentrations of particulate matter in the air. The CMDM Standard provides a universal data sharing framework for all of these devices (vehicles, air quality sensor systems, etc.) to communicate and share information. These schemas are how the CMDM Standard accomplishes data interoperability. By prescribing underlying core services (identity, settlement, discovery, and more), the CMDM Standard also addresses functional interoperability.

CMDM-compliant systems will generate clean, easily parsed, and similarly organized data. For data scientists and engineers seeking to utilize that data within their own applications, this means much less time cleaning and processing data, and a much more efficient data pipeline from source to endpoint. From the perspective of the data value chain, this is critical. As previously explained, raw data has little to no value. It is when that raw data is processed, organized, and utilized in the extraction of insights and execution of services that value is generated. By greatly reducing the time and resource cost of ETL for raw data, that data is effectively moved further down the value chain, minimizing the number of steps required for that data to ultimately provide value at the end of the chain.

Blockchain allows for entities to authenticate each other's identity, as well as immutably record and securely expose collected information to each other. But this can only be achieved when those entities have common standards for identity, data, and functions. This is a major focus of the CMDM Standard.

In developing the CMDM Standard's schemas, the Working Group focused on ensuring that the schemas could appropriately capture all possible events – uniformly representing that event's associated data regardless of what type of data is involved. For "simple" events, like a recording of an engine's oil temperature, or a weight sensor detecting that someone has sat down, this is not relatively difficult. This is the case for most of the types of sensors that the Working Group considered, which also include force, torque, inertial, optical, ratio, speed, etc. However, when considering more complex data, such as LiDAR, camera, or video, representing them uniformly becomes much less straightforward. In order for data exchange to provide as much value as possible, it must support all possible sorts of data. The CMDM Standard thus also prescribes methods for registering large, complex data, primarily by creating a smaller representation of that data for use as a pointer.

Perhaps most importantly, the Standard is also the key first step in enabling data monetization and efficient data marketplaces. As with any good, the price of data is determined by its value – what that good sells for in an efficient market. But no efficient market for data exists today, so there is no ability to prescribe a price or common metric of value. The first step to achieving an efficient data market is for there to be secure systems for exposing and sharing data. Without such systems, there would not exist a secure, scalable way of delivering data to buyers.

The CMDM Standard

Traditionally, data is shared bilaterally, or within one closed, siloed ecosystem. Such one-to-one transactional structures don't lend themselves to a scalable marketplace model. But more importantly, the question is: why did the market for data evolve this way, while markets for other goods, like oil, did not?

The CMDM Standard provides a structure for data to act as a private good, allowing for the creation of an efficient market for it.

This paper has already discussed some causal factors. As previously mentioned, data is not a private good, as it is not rivalrous; in other words, the same data can be copied and distributed without additional marginal cost. For such goods, the key is to be able to charge for the usage of the goods. For example, consider roads. If one driver is on the road, they are not preventing others from driving on that same road (excluding issues of traffic congestion, of course). There is also no way to prevent someone from getting in their car and driving on a road. Roads are not private goods, and as a result, they represent a sunk cost for local governments, as they are a significant investment without direct returns. Governments solve this problem by introducing road user fees, often through toll stations or automated charging systems. Drivers are charged for their use of the road and can be turned away or fined for not paying that charge. By creating a structure that allows for charging for use of non-private goods, they can act as private goods within that structure, and an efficient market can be built for it.

Applying this model to the issue of creating an efficient market for data, it becomes clear that structures must be put into place that allow for entities to be charged for the usage of data. Charging for usage of data requires two sets of functionalities: secure access management of data, and logic for the execution of payments. The structures prescribed by the CMDM Standard allow for actors to securely grant data access to other actors. That is, the CMDM Standard provides the structures for data to act as a private good, which is critical for its monetization in a data marketplace.

Data privacy, security, and compliance are core to the CMDM Standard. Unobfuscated data is never directly shared. Additionally, the Standard is flexible enough to allow for various different governance models, reflecting that different sorts of data may require different marketplace structures.

The CMDM Standard was crafted with data privacy, security, and compliance as core considerations. The Standard's framework for data ensures that unobfuscated data never leaves any given actor's data lake, rather, attestations are made to a distributed ledger, with robust access management for any data reads. Moreover, data provenance can be guaranteed by comparing data to its recorded attestation. As that attestation exists on a distributed ledger, that record has no single point of failure. Security and data privacy exist on a protocol level and is therefore guaranteed in all CMDM-compliant applications, and (within reason) in all possible governance models. Notably, the Standard does not prescribe governance models for data markets, as these are better left to the implementers. Rather, it prescribes the necessary functionalities for the decentralized operation of such a marketplace. In all likelihood, there will exist many data marketplaces operating independently, each with differing features, datasets, etc. Depending on each market's context, they may prefer a holistic integrator to manage the top level of the mobility ecosystem, or a more distributed approach. In light of this fact, the Standard ensures applicability in a variety of governance models.

The CMDM Standard also introduces a brand new identity/ownership certificate for sensors. Currently, when a sensor generates a data point, there does not exist any protocol-level functionality for associating that data point with a given sensor. There are exceptions to this. Some manufacturers create more powerful devices that typically contain many sensors, and some of these devices will have some

The CMDM Standard

The CMDM Standard provides sensor identity certificates and sensor ownership certificates for use in large, complex sensor networks.

sort of device ID, but there is no standardization. As a result, complex systems with many sensors can struggle with diagnosing failures, software/firmware consistency across devices, and more. The certificate simultaneously gives a unique ID to each sensor, as well as associates that sensor with a given owner. Attestations of that certificate go onto the distributed ledger, and are therefore given the same security guarantees. This is key, as sensors generate raw data, which begins at the first step in the data value chain. Moving that data through the value chain, by processing, refining, extracting insights, etc. belies an often enormous, complex system that has to handle and organize data, which is streaming in real-time (or nearly real-time) from sensors. Sensor IDs and Ownership allow for much, much more precise control and understanding of a sensor suite.

The CMDM Standard's three core components - core services, logical schemas, and sensor identity/ownership certificates provide the base structure required for bridging the current silos in mobility ecosystems.

Ultimately, the CMDM Standard's three main components – ***core services***, ***logical schemas***, and ***sensor identity/ownership certificates*** – provide the structures required for smoother data sharing, allowing communication and collaboration between the siloed ecosystems of each Mobility Service Provider. Such structures lay the foundation for robust, distributed data marketplaces operating at scale. The Standard carefully and effectively distinguishes between what sorts of data and functionality must be standard across all implementations, versus what considerations are best left up to the engineers. Of course, there are many hurdles to overcome in achieving an at-scale data marketplace, but the CMDM Standard represents a key step in addressing many of those challenges. Standards are much like mobility ecosystems, data marketplaces, and distributed ledgers, in that they all rely on network effects to provide value. The CMDM Standard enables an exciting array of use cases, and those use cases will grow in number, complexity, and value as adoption grows over time. Therefore, the remainder of this paper will be spent discussing some use case examples that are feasible on short, medium, and long term time horizons.

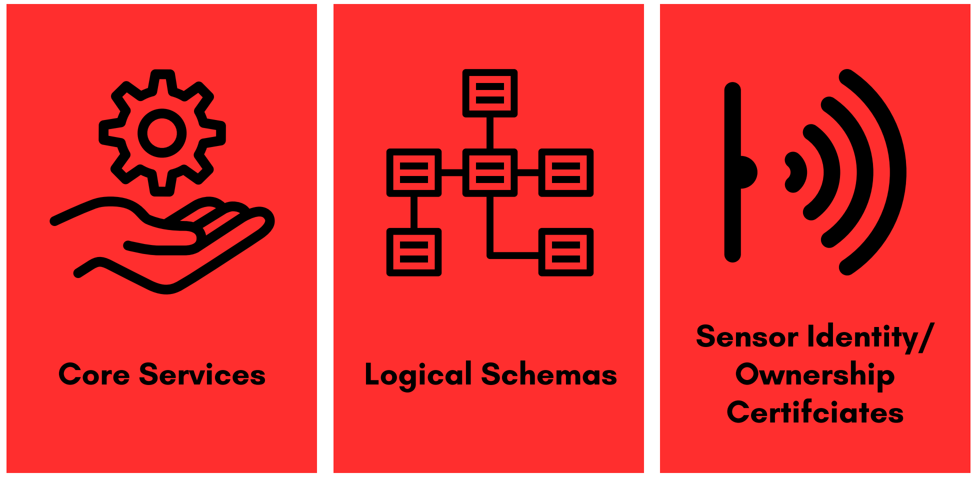


Figure 1 — CMDM Main Components

SHORT-TERM USE BENEFITS

The breadth and depth of use cases that are enabled by the CMDM Standard increase with the Standard's adoption. Even still, the CMDM Standard enables exciting applications in the immediate term. Modern vehicles are endowed with a suite of sensors, sufficient (and rapidly increasing) computing power, and communications capabilities. These are more than enough to support many v2x applications, and these capabilities will only grow more rapidly over time.

Vehicles can use their ECUs to collate sensor data to enable insights on safety factors. When a vehicle's optical sensors or cameras detect an obstruction on the roadway, its ECU can collate that information and send the key insights (the existence of the obstruction, its location, etc.) to other vehicles nearby, as well as to transportation authorities, and any other relevant parties. This capability - to detect events in the environment, measure and record that event, and share the resulting data with available parties (in the immediate vicinity or at distance) enables many safety-related V2X applications. For two simple examples, consider Intersection Movement Assistance (IMA) and Left Turn Assistance (LTA). In an IMA application, a vehicle that is about to run a red light can alert other vehicles before entering the intersection, allowing other vehicles to adjust course.[22] In an LTA application, a vehicle turning left can broadcast its intent to begin turning, and issue alerts to nearby cars attempting to drive straight through. Alternatively, a vehicle turning left can receive an alert from a vehicle that is about to come straight through, and automatically brake to avoid a crash.[23] These applications are not only buildable and achievable today, but are made quite simple with the tools and infrastructure provided by the CMDM Standard. The National Highway

22. "Safety - Intersection Movement Assist", US Department of Transportation, accessed January 2021, https://www.its.dot.gov/infographs/intersection_movement.htm

23. "Safety - Left Turn Across Path", US Department of Transportation, accessed January 2021, https://www.its.dot.gov/infographs/left_turn_across.htm

Short-Term Use Benefits

The CMDM Standard allows for vehicles and infrastructure to share safety-related information, like geolocation and movement intention, which can be used in many safety applications like Intersection Movement Assistance and Left Turn Assistance.

The usefulness of safety data can be augmented by cross-validating geolocation with other vehicles. Moreover, validation of geolocation data and availability of highly granular location information is directly monetizable.

Traffic Safety Administration estimates that implementing just these two applications could prevent up to nearly 600,000 crashes every year in the United States alone, which would save many lives, avoid many injuries, and additionally prevent up to 728,000 property-damage only crashes.[24] The returns on these two simple, but effective applications, are enormous. When additional safety applications are considered (Do Not Pass warnings, Forward Collision warnings, roadway debris detection, and more), the value in terms of reduced death and injury, as well as property damage, grows exponentially.

Key to many of these safety applications is the vehicle having robust geolocation functionalities. If the vehicle is not able to communicate its location, then the content of its data is made much less useful, For example, if a given vehicle is experiencing a catastrophic failure, but can't properly communicate where it is while the failure occurs (perhaps sending faulty data, or even no data), then other vehicles, emergency services, or other entities are not able to take action to help. This could occur in case of a fire, part failure, or other factors. However, there is a silver lining – other vehicles nearby could register that said vehicle is in danger, and broadcast its location on its behalf. What is occurring here is effectively a cross validation of geolocation data, and that application has value far beyond safety considerations. Location services are critical, and the ability to harness a vehicle's ability to sense its surroundings and geolocation for highly granular location data is highly useful and monetizable. This is true for safety applications as well as commercial applications. Vehicles that are communicating with other vehicles in overall service of accident prevention and driver safety have a major competitive advantage over vehicles that are not doing so. And the highly granular location data achievable through cross validation can be utilized directly in the creation of new applications, or access to that data can be sold to others – both of which are direct monetization avenues for that geolocation data.

24. John Harding, Gregory Powell, Rebecca Yoon, Joshua Fikentscher, CharleneDoyle, Dana Sade, Mike Lukuc, Jim Simons and Jing Wang, "Readiness of V2V Technology for Application", US Department of Transportation, August 2014, https://www.nhtsa.gov/staticfiles/rulemaking/pdf/V2V/Readiness-of-V2V-Technology-for-Application-812014.pdf

The CMDM Standard's underlying core services and software infrastructure that power these simple safety applications also power more exciting and complex applications, like coordinated autonomy.

Moreover, discussion of these applications here is focused on what is achievable in the short term. But, like the value of the Standard and of distributed systems in general, what is achievable in the realm of V2X safety technologies is entirely determined by the technological capabilities of the vehicles and the network effects and adoption of the Standard. This means that as these technologies continue to progress at a rapid pace, and adoption of the Standard grows with them, safety applications will, too, become increasingly robust, complex, and prevent more and more accidents. Consider a future scenario – fully autonomous vehicles coordinating to safely traverse an intersection. Those vehicles must be able to capture and communicate information about their surroundings. Even further, they must be able to process that information into actionable steps and communicate those steps as their movement intention to other vehicles nearby. This requires those vehicles to have a secure, distributed system for identifying themselves (otherwise, they would not be able to verify the authenticity, and therefore actionability, of information received from other sources), and the ability to quickly and seamlessly share data about their surroundings. These are some of the exact core services that the CMDM Standard focuses on – the same core services enabling more straightforward, immediate term use cases like communication of safety data, or geolocation cross-validation. More generally, the applications that the CMDM Standard can immediately enable will provide more value over time as technology improves and adoption grows.

MEDIUM-TERM USE BENEFITS

As time goes on, network effects and adoption will increase as the ecosystem matures. This means that there will be many more devices connected to the ecosystem. Consider the environment in which this ecosystem is functioning – an urban space, with diverse modes of transportation to traverse it. One fairly common option is personal vehicle ownership, driving that vehicle on the city roads to reach a destination. There are often subways, or overground trains, that move on fixed paths, and carry large amounts of commuters. There are complex bus systems that move many commuters across large geographic areas. There are on-demand rideshare services and taxis available, or on-demand motorized scooters or bikes available for rent or purchase. There are numerous modes of transportation operating in any urban environment.

All of these different modes of transportation have a few shared characteristics. They all involve using a device (vehicle, bus, scooter, etc.) that is capable of generating data about itself and its surroundings. They often intersect with each other – for example, people taking the subway, then using a rideshare service for last-mile transportation. Or someone driving their vehicle to a parking garage, then using an e-scooter to travel the rest of their commute to work. Finally, they are all services. All involve a relationship between a traveler and a Mobility Service Provider. Their services are both providing direct value (in the form of transportation) to a given user, as well as indirect value (to be realized further down the data value chain) as they generate useful data about the environment, traffic and commuter flows, and much more. That data is moved from these devices (broadly, the vehicle layer) to be used in applications. For example, rideshare vehicles expose their location and identifying information to the rideshare application in order to deliver that data to the rider. Municipal transit authorities utilize data on commuter flow in order to prepare more resources for peak travel times. These mobility services generate useful data, whose value is currently only realized by its generating entity, within their silo.

Medium-Term Use Benefits

Although different modes of transportation have many different definitional features, they share core characteristics that allow them to exist as equally-privileged peers in a mobility ecosystem.

The key is that, though transportation modes vary widely in their form, operation, and entire role, these shared characteristics enable them to all participate in one ecosystem. More and more transportation modes will connect to that ecosystem as it matures. And with them, their Mobility Service Providers, the users of those MSPs, and their data silos will also connect to the ecosystem. The CMDM Standard enables data to be securely shared within that ecosystem layer. With a diverse collection of data-generating devices, organizations, travelers, and other players, the ecosystem would represent a data source, for all different types of mobility data, that is far, far more complete and diverse than any other source of mobility data. Such a data source is able to support a wide range of new services.

The CMDM Standard provides the necessary infrastructure to support a connected multimodal transportation service, coordinating scheduling, booking, payments, and other necessities across different modes of transportation.

When the ecosystem is more mature, with ample mobility data being produced and made available, organizations are able to create many more, exciting services. One example of such a new service would be a connected multimodal travel planning platform. The platform could construct trip plans involving many potential transportation modes in order to optimize travel time or other commuter preferences. By combining multiple modes of transportation - for example, utilization of on-demand services like rideshare vehicles or e-scooters for first/last mile transportation, along with a subway or bus route for the bulk of the journey - would be able to create much more efficient, affordable travel for many people in urban environments.

The key to such an application is its ability to access the necessary data from each mode of transportation, or, in other words, each mode of transportation's ability to easily, interoperably, and securely expose its data to one application. The CMDM Standard is precisely the mechanism for this. The applications extend far beyond multimodal trip planning. Yet all ultimately require the same functionality - access to data, from diverse sources, that are nowadays generally inaccessible. In the medium term, the CMDM Standard will allow such access, enabling organizations to offer exciting new applications. In other words, the CMDM Standard, over time, will allow for the creation, purchase, and sale of new mobility services.

The CMDM Standard allows for entities to wed the monetization of their data to the success of the applications and services powered by their data, which is key to ensuring proper incentives around data sharing.

This paper has already discussed some of the difficult factors in reaching the production, scaled version of such a data exchange ecosystem. As previously discussed, one of the most salient of those factors is the fact that nobody wants to share their data for free. Companies extract value from that data, d do not want to lose out on monetization opportunities associated with the data they generate. In the real world, data providers seek to monetize their data by themselves and would require a comparable return from making that data available to others. Making that data available in a data marketplace, in which data providers can structure commercial transactions fo data access as they see fit, is key. For example, some data providers will insist on repeated costs for each call to their database, while others may accept one-time payments for access. The CMDM Standard provides the core services and interoperability required to construct a marketplace that allows for applications to easily make calls to data providers, fetch that data, and pay for that access.

While providing structures for the monetization of data is key for marketplaces to reach production on a meaningful scale, there are many examples of data sharing occurring without direct monetization avenues. Over the last 30 years, many P2P file sharing services have gained enormous popularity and usage. Many users offer their storage space and bandwidth, at no cost, to host files for other users to access. All of these services

Medium-Term Use Benefits

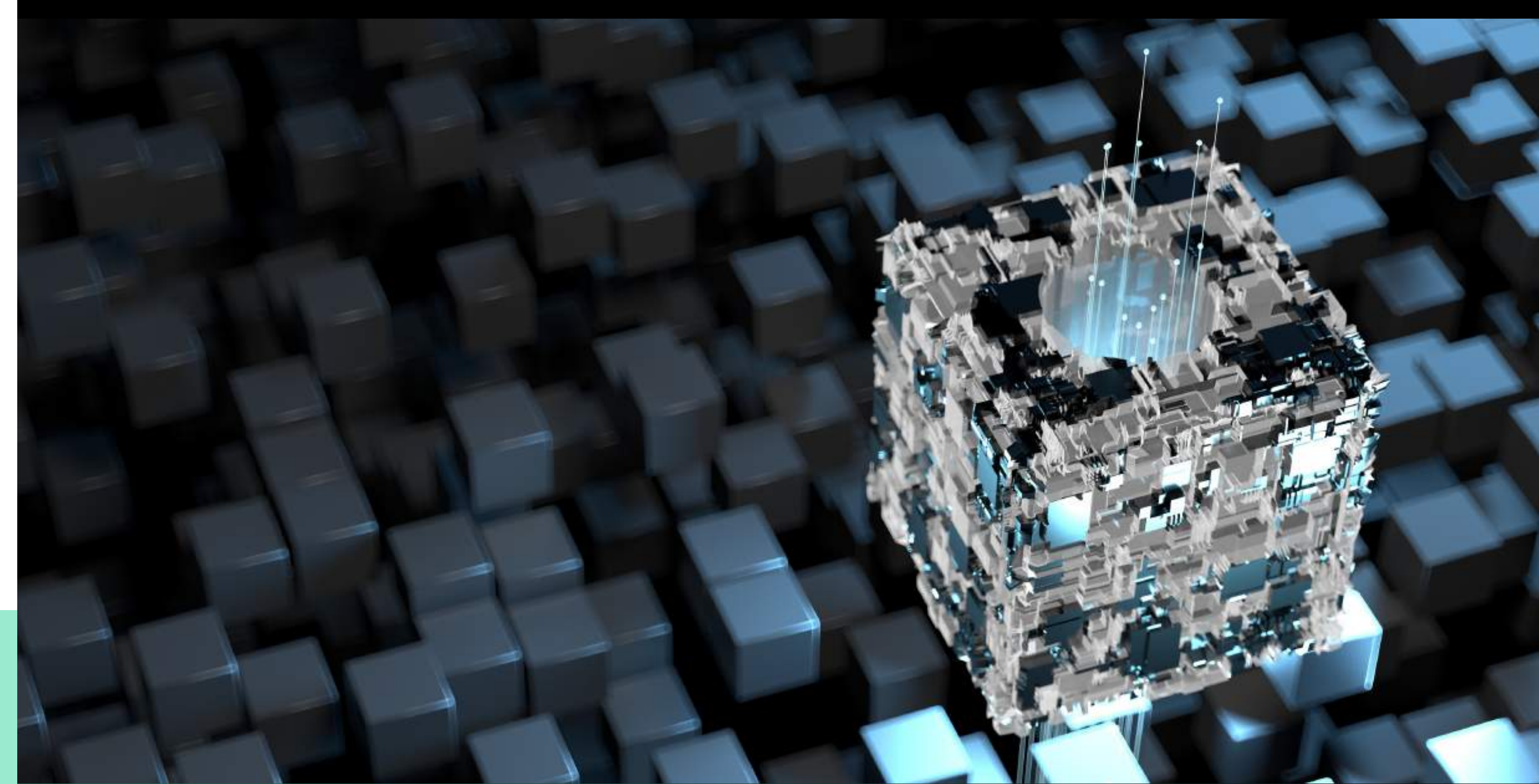
have relied on platforms, some of which were centralized, and others open sourced or otherwise operated in a decentralized way. This is fine in a B2C context, but when businesses are the users, it is much less straightforward.

There is enormous value in being the owner of an ecosystem layer, which disincentivizes joining any existing marketplace or ecosystem. Having that ecosystem layer be operated in a neutral way, and having network effects exist from the beginning are some of the driving forces that led to the formation of MOBI.

There is much value in being the owner of an ecosystem layer, so joining any already existing ecosystem is highly disincentivized. This is the problem of achieving network effects, that there's no incentive for anyone to be the first. This is precisely why MOBI was formed – to have a community of coordinating players to jointly deploy these systems. In the medium term, these network effects will exist, allowing for the creation of data marketplaces that provide the ability to monetize the sharing of data. The CMDM Standard addresses the previously discussed pain points that have prevented the success of previous data marketplaces. The success of these marketplaces are the singular most important piece of enabling the monetization of data sharing in the ecosystem layer.

When both secure data exchange and monetization exist, the ecosystem reaches an inflection point, as these key barriers separating data silos are finally addressed.

These two pieces – enabling secure data exchange and enabling the monetization of data shared externally – are two key milestones in the development of the new mobility ecosystem. It is these two factors that will truly catalyze the ecosystem in its creation of new solutions and value. In the short term, the CMDM Standard enables many exciting applications, with truly profound outcomes. But it is when these two pieces are achieved that the ecosystem reaches an inflection point, as data providers will finally have a proper incentive system to share their data externally. All of their combined data, within this ecosystem, will ultimately become critical infrastructure, supporting valuable and useful tools and applications for MSPs and users alike.



LONG-TERM USE BENEFITS

Over a longer time horizon, the share of vehicles on the road with robust V2X communication capabilities, powerful ECUs, vehicle wallets, and other key technologies will be very large. These vehicles will have ample computing power - powerful enough to execute large scale computations on bigger data. Equipped with the machinery and infrastructure provided by the CMDM Standard (and other MOBI Standards and implementations), these vehicles will be able to easily and securely share data and insights, allowing coordination. This coordination and communication will not only be happening between vehicles, but public transportation (trains, busses, etc.), stoplights, and other infrastructure. This fully realized interplay between traditionally siloed systems allows for exciting use cases centered around coordination, especially within the area of autonomous vehicles, big data, and edge computing.

Autonomous vehicle systems developed to date do not rely heavily on V2X connectivity. Instead, self-driving capabilities in most AV prototypes and early production systems rely primarily on in-vehicle sensors and on board computing to navigate safely through their environment. This “standalone” autonomy is inherently limited by the data available to that vehicle’s sensors. For example, on board sensors cannot see around corners, through traffic, or over obstacles. “Coordinated” autonomy has enormous potential to improve AV capabilities through greater data availability and consensus. Faster vehicle connections and edge computing will erase these limitations. Vehicles will be able to share information about road conditions, hazards, and traffic. They will be able to share their intentions and modify their behavior accordingly. Such coordination would present a wide array of benefits - large scale reductions in accidents and smoother, faster transit with less traffic are two key examples. Coordinated autonomy will far exceed the safety and efficiency of standalone autonomy.

Long-Term Use Benefits

In the long term, as these silos become fully bridged and integrated, autonomy will shift away from the current standalone model, towards coordinated autonomy for more robust autonomous decision making.

Standalone autonomy is limited by the data availability it implies – purely what one vehicle is able to ascertain about its surroundings with its sensors. The benefit of secure data sharing and coordination, in part, is due to the increase in data availability; rather than relying entirely on their own data, vehicles are able to access and make actionable decisions based on the data of other vehicles around them. In other words, the dataset that a vehicle can generate about its surroundings is much, much smaller than the dataset that an entire intersection's worth of vehicles can generate about their surroundings. This is true both within an autonomous driving context, as well as other contexts. Organizations seeking to carry out analytics and extract insights from vehicle data for any purpose gain a similar benefit – the smooth and frictionless creation of bigger, cleaner, and more diverse datasets than can be extracted from one vehicle, or even one fleet of vehicles. It's when a vast majority of vehicles have the ability to carry out this communication and data sharing where the resulting datasets can provide the most value, and enable the most use cases.

Coordinated autonomy allows for autonomous vehicles to make decisions that rely not only on the data that they can generate, but also the data from other vehicles and infrastructure around them, vastly increasing the available pool of information that informs driving decisions.

Much of the use cases discussed above rely on a vehicle's ability to generate data, guarantee that data's provenance, and share it with peers (other vehicles, infrastructure, servers, etc.). Modern, connected vehicles are already generating extremely large amounts of data, with some low estimates at 25 gigabytes of data per hour, or up to 300 terabytes of data per year. With hundreds of millions of vehicles on the roads, this results in truly enormous quantities of data being generated every hour, which presents a myriad of challenges for OEMs, telecoms, and other mobility players. Such large amounts of data aren't able to be affordably transferred via cellular or other network solutions, which prevents the majority of that data from reaching server farms for storage and statistical analysis.

Large computing power within vehicles and other infrastructure, along with their ability to securely authenticate, share data, and coordinate allows for larger workloads to be processed on the edge. Edge computing is key to scaling the Internet of Things.

This data is valuable and contains insights that can power monetization and value extraction further down the data value chain. However, much of this data is simply deleted, as there is just no mechanism to handle it at scale; in other words, this issue represents enormous lost economic opportunity. In the long term, however, these vehicles will accumulate greater and greater computing power, and with[13] the CMDM Standard, the ability to communicate and coordinate. So, instead of sending raw data to a datacenter to extract insights, these vehicles could pool their computational resources to carry out that statistical analysis in the vehicles themselves and send the resulting insights to a server for storage. Such a system would quickly result in gigantic savings from reductions in network costs alone. Edge computing has been a large focus for the distributed computing community, and the MOBI community believes that the widespread growth of connected and autonomous vehicles represents one of the most exciting and salient opportunities for edge computing to provide cost savings and new revenue opportunities.

CONCLUSION

This paper has discussed a variety of factors that have hampered the growth of data marketplaces, even as the meteoric rise of the data economy has clearly demonstrated that data itself has economic value. The key is that raw data, itself, doesn't provide any direct economic value. Rather, the economic value of data comes from the services that rely on the insights and analytics from data where economic value is realized. That process - the path that raw data will take from generation to providing direct economic return - is called the data value chain.

For organizations relying on data to power services, there is a direct cost of revenue that is implied here. Successful extraction of insights from data at scale implies the existence of a robust data pipeline, ample server power and space, skilled data scientists capable of executing such analytics, and more. Naturally, after making such an investment, organizations are reticent to make their data and insights open and available to others for free. Each organization would rather provide that end service and capture its full value, rather than sell it to a competitor. Additionally, since the value of data is subjective based on its use, there is no universal measurement of its value, necessitating individual negotiations between organizations for each transaction. Moreover, if an organization does wish to share their data with another entity, their systems may not handle and record data in the same way, necessitating costly integrations.

All of these factors have contributed to this odd reality, that in our modern, service-based, and data reliant economy, there are no scalable marketplace structures for data as a commodity. The result is a patchwork of siloed ecosystems, largely unable to communicate. The key to solving this problem lies in standardization efforts and industry-wide coordination and collaboration. Data sharing standards, monetization structures, ownership models, governance, and much more must be rigorously defined. But the CMDM Standard, and the MOBI community behind it, represents a critical step in this journey. By standardizing data and core services, the CMDM Standard provides a robust infrastructure to support secure data sharing between devices and entities. This is a foundational layer, on top of which marketplaces and many other applications can be built. With scalable structures that support the sharing, provenance, and value exchange of data, the CMDM Standard will provide the infrastructure for a multitude of use cases across the mobility ecosystem.

Bibliography

"Alternative Fuels and Advanced Vehicles", US Department of Energy, accessed January 2021, https://afdc.energy.gov/fuels/.

"Alternative Fuels", US Department of Energy, accessed January 2021, https://www.fueleconomy.gov/feg/current.shtml.

Anne Grosse-Ophoff, Saskia Hausler, Kersten Heineke, and Timo Möller, "How shared mobility will change the automotive industry," McKinsey & Company, April 2017, https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/how-shared-mobility-will-change-the-automotive-industry.

Automated Vehicles for Safety", NHTSA, accessed January 2021, https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety.

Ayan Brahmachary, "DIKW Model: Explaining the DIKW Pyramid or DIKW Hierarchy", CertGuidance, May 2019, https://www.certguidance.com/explaining-dikw-hierarchy/.

"Big Data Analytics", IBM, accessed January 6, 2021, https://www.ibm.com/analytics/hadoop/big-data-analytics.

"Big Data, What it is and why it matters", SAS, accessed January 2021, https://www.sas.com/en_us/insights/big-data/what-is-big-data.html.

"CASE - Intuitive Mobility", Daimler, accessed January 2021, https://www.daimler.com/case/en/.

"Charging Your Car", Tesla, accessed January 2021, https://www.tesla.com/support/charging.

Charles McLellan, "What is V2X communication? Creating connectivity for the autonomous car era", ZDNet, November 2019, https://www.zdnet.com/article/what-is-v2x-communication-creating-connectivity-for-the-autonomous-car-era/.

Christian Bussmann, Magdalene Piec, Dr. Andreas Gutweniger, "MaaS and Data Sharing: Five Issues Closely Related to Practice", Detecon Consulting, April 2016, https://www.detecon.com/en/knowledge/maas-and-data-sharing-five-issues-closely-related-practice.

Chuck Frey, "Customer Centric Innovation", Innovation Management, April 2004, https://innovationmanagement.se/imtool-articles/customer-centric-innovation/.

"Data Done Right: 6 Dimensions of Data Quality", SmartBridge, accessed January 2021, https://smartbridge.com/data-done-right-6-dimensions-of-data-quality/.

"Data Economy: Radical transformation or dystopia?", United Nations, January 2019, https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/publication/FTQ_1_Jan_2019.pdf.

D. Bergemann and A. Bonatti. "Markets for information: An introduction", Annual Review of Economics, November 2019.

Dave Chaffey, "Golden Circle model: Simon Sinek's theory of value proposition, start with why", Smart Insights, October 2020, https://www.smartinsights.com/digital-marketing-strategy/online-value-proposition/start-with-why-creating-a-value-proposition-with-the-golden-circle-model/.

Doyle, Dana Sade, Mike Lukuc, Jim Simons and Jing Wang, "Readiness of V2V Technology for Application", US Department of Transportation, August 2014, https://www.nhtsa.gov/staticfiles/rulemaking/pdf/V2V/Readiness-of-V2V-Technology-for-Application-812014.pdf.

Edd Dumbill, "Understanding the Data Value Chain, November 2014, https://www.ibmbigdatahub.com/blog/understanding-data-value-chain.

"How many times have you heard - "Automakers are sitting on a data goldmine"?" SBD Automotive, accessed January 2021, https://www.sbdautomotive.com/en/news-automakers-goldmine.

Jason Fernando, "Law of Supply and Demand", Investopedia, November 2020, https://www.investopedia.com/terms/l/law-of-supply-demand.asp.

Jennifer Belissent, PhD, "Help Wanted: Data Innovation For The Data Economy", Forrester, October 2017, https://go.forrester.com/blogs/help-wanted-data-innovation-for-the-data-economy/.

John Findlay, "The Wisdom Economy.", thewisdomeconomy.blogspot.com, August 2011, http://thewisdomeconomy.blogspot.com/.

John Harding, Gregory Powell, Rebecca Yoon, Joshua Fikentscher, CharleneKiran Bhageshpur, "Data Is The New Oil -- And That's A Good Thing", Forbes, November 2019, https://www.forbes.com/sites/forbestechcouncil/2019/11/15/data-is-the-new-oil-and-thats-a-good-thing/#4d1da4dd7304.

"Keynesian Economics", Investopedia, April 2020, https://www.investopedia.com/terms/k/keynesianeconomics.asp.

Kirsten Korosec, "Tesla plans to launch a robotaxi network in 2020" TechCrunch, April 2019, https://techcrunch.com/2019/04/22/tesla-plans-to-launch-a-robotaxi-network-in-2020/.

Leon Nash, Greg Boehmer, Mark Wireman. Allen Hillaker, "Securing the future of mobility" Deloitte, April 2017, https://www2.deloitte.com/us/en/insights/focus/future-of-mobility/cybersecurity-challenges-connected-car-security.html.

"Mobility Ecosystem: A wave of transformation", Moveco, July 2018, https://medium.com/@moveco.io/mobility-ecosystem-a-wave-of-transformation-c8f2eb7af658.

Nick Ismail, "The new front line of cyber security is mobility", Information Age, March 2018, https://www.information-age.com/front-line-cyber-security-mobility-123471096/.

Peter Götz, "Why - How - What: From Product Vision to Task", Scrum.org, March 2018, https://www.scrum.org/resources/blog/why-how-what-product-vision-task.

Bibliography

Rado Kotorov, "The Data Value Chain: Steps for Monetizing your Data", Integration, accessed January 2021, https://www.idevnews.com/stories/6998/The-Data-Value-Chain-Steps-for-Monetizing-Your-Data.

Rainer Sternfeld, "Designing a Better Planet with Big Data and Sensor Networks", Intelligent Sensor Networks Conference, November 2014, https://www.slideshare.net/rainer_sternfeld/designing-a-better-planet-with-big-data-and-sensor-network-for-isn-conference-philips-high-tech-campus.

Ralf Baron, Michael Zintel, Niklas Schemken, Christoph Uferer, "Integrated Mobility Platforms", Arthur D Little, February 2018, https://www.adlittle.de/en/insights/viewpoints/integrated-mobility-platforms.

Regina Clewlow, "The Opportunity To Reshape Cities With Shared Mobility Data", Forbes, October 2018, https://www.forbes.com/sites/reginaclewlow/2018/10/10/the-opportunity-to-reshape-cities-with-shared-mobility-data/#5cb5ff3e617f.

Ricardo Bacellar, "What will the future mobility ecosystem look like?", KPMG, November 2017, https://assets.kpmg/content/dam/kpmg/br/pdf/2017/11/br-kpmg-congresso-sae.pdf.

Rob Lokers, "Analysis of Big Data technologies for use in agro-environmental science", ResearchGate, October 2016, https://www.researchgate.net/figure/DIKW-hierarchy-from-Big-Data-to-decision-making-for-societal-challenges_fig1_307874029.

"Safety - Intersection Movement Assist", US Department of Transportation, accessed January 2021, https://www.its.dot.gov/infographs/intersection_movement.htm.

"Safety - Left Turn Across Path", US Department of Transportation, accessed January 2021, https://www.its.dot.gov/infographs/left_turn_across.htm.

Scott Corwin, Philipp Willigmann, "The evolving mobility ecosystem", Deloitte, accessed January 2021, https://www2.deloitte.com/us/en/pages/manufacturing/articles/the-evolving-mobility-ecosystem.html.

Shannon Bouton, Stefan M. Knupfer, Ivan Mihov, Steven Swartz, "Urban mobility at a tipping point", McKinsey & Company, September 2015, https://www.mckinsey.com/business-functions/sustainability/our-insights/urban-mobility-at-a-tipping-point.

Sylvain Santamarta, Rash Gandhi, and Michael Bechauf, "Big Oil, Big Data, Big Value", Boston Consulting Group, August 2019, https://www.bcg.com/publications/2019/big-oil-data-value.aspx.

"The Data Value Chain Executive Summary", Open Data Watch, accessed January 2021, https://opendatawatch.com/reference/the-data-value-chain-executive-summary/.

"The Smart Grid", US Department of Energy, accessed January 2021, https://www.smartgrid.gov/the_smart_grid/smart_grid.html.

"The World's Most Valuable Resource Is No Longer Oil, But Data.", Economist, May 2017, https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data.

Tobias Schönberg, "Shared Mobility", Roland Berger, July 2014, https://www.rolandberger.com/de/Publications/Shared-mobility.html.

Warwick Goodall, Tiffany Dovey Fishman, Justine Bornstein, and Brett Bonthron, "The Rise of Mobility as a Service", Deloitte, January 2017, https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/consumer-business/deloitte-nl-cb-ths-rise-of-mobility-as-a-service.pdf.

Will Kenton, "Understanding Coopetition", Investopedia, May 2020, https://www.investopedia.com/terms/c/coopetition.asp.

MOBI is a nonprofit industry consortium of many of the world's largest vehicle manufacturers, along with many startups, NGOs, transit agencies, insurers, toll road providers, smart city leaders, and technology companies working to accelerate adoption and promote standards in blockchain, distributed ledgers, and related technologies.

MOBI is creating simple blockchain-based standards to identify vehicles, people, and businesses in order to securely exchange and monetize data, and pay for mobility services, with the goal of making transportation more efficient, affordable, greener, safer, and less congested. MOBI itself is technology and ledger agnostic. For additional information about or joining MOBI, please visit our website.

dlt.MOBI   @dltMOBI   MOBI   @dltMOBI   cmdm@dlt.MOBI