AUTONOMOUS DRIVING

BIG DATA *Value* MYTH

MANU NAMBOODIRI & NITIN KUMAR

Autonomous Driving: The Big Data Value Myth

A lot has been said about data being the new oil and gold etc. Data value cannot be generalized, its intrinsic and monetizable value varies. In this article, we explore the value of data and what makes it valuable, through the autonomous driving lens.

Introduction

Today, there is a ton of data generated from people, devices and about many things e.g., weather, location, demographics, purchase patterns etc. Moreover, we are shifting from an era of user generated data/content to an era of machine generated data/content increasing the volumes exponentially.

> Unlocking value from autonomous driving value chains is not about having more data – but the right data, at the right time, for the right ecosystem participant

Most of these large sets of data are considered valuable because they can be packaged and sold to someone else. However, the common myth is that collecting lots of data about lots of things creates value. It is important to understand the right criteria and attributes of data to generate value effectively. For example, Google collects a lot of data, but value is accrued only when it is meaningfully processed thorough applications, APIs, or service to create value for customers.

The same holds true for an autonomous vehicle. There is a lot of optimism with companies looking to benefit from the massive amounts of data generated from modern automobiles. We believe that without deep understanding of automobile and the autonomous data value chain, and viable mobility business models, real value from data cannot be realized.

The AV Data Value Chain

Autonomous vehicles collect a wide variety of data from a multitude of sources and formats e.g., sensors, maps, traffic, vehicle health, operations - the usage of this data is also spread across a wide spectrum of use cases. Before we look at the value of this autonomous vehicle data, let us understand data better by looking at it through two lenses i.e., its useful life (fast vs slow) and its usage (internal vs ecosystem). The illustrative example in Figure 1 provides us a non-exhaustive list of data sources and usage from the autonomous vehicle.

Useful Life

Useful life can be described as how quickly the data is used after initial capture and how long will it be useful. Fast data is newly created data – and data that is acted upon quickly. Note that all data is born fast – as time passes, its value usually decreases as it becomes slow data.

Data Usage

Usage pertains to utilization of the data within the autonomous mobility ecosystem i.e., whether it is used, transported, aggregated directly as data or programmatically accessible and used by applications and services, either locally on the vehicle edge or outside in the cloud. The usage attribute defines the actionability of the data – actionable data is the most useful of all.

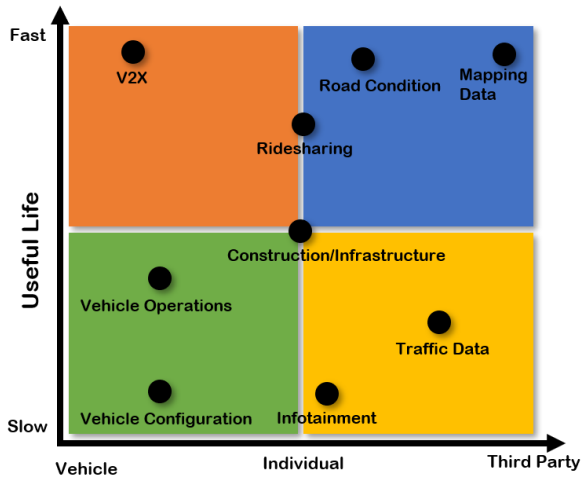

Figure 1: AV Data Value Context

The above attributes also drive inform the storage location of data, ease of accessing data (smart phone vs specialized sensor), level of modification/manipulation of data etc.

There is also an inherent cohesion and linearity between valuable data within the autonomous vehicle. For example, data from construction and traffic is utilized into Maps and Maps directly informs navigation of the vehicle in real time. The more cohesive and interdependent the data is, the faster it needs to be available to multiple components of the ecosystem and hence more valuable it becomes. In some cases, sensors are deployed for redundancy and the data arriving from these sensors are used as confirmatory and improves machine extracted intelligence. The data is also processed at the device or at the edge leveraging machine inference, whereas slow data is typically stored in the cloud and informs deep learning.

The AV Data Chain: Key Value Drivers

All things being equal, fast data creates more value per byte – since it is available for immediate action. As data gets older, while still useful, it has reduced value per byte. Value is driven from integration, combination for analysis, reporting, planning etc. Figure 2 illustrates linear and exponential value of data by its characteristics at each point of the value chain.

The next level that dictates further value creation, beyond fast or slow data, is how actionable the data is. If the data is critical for immediate decisions and action, this increases its value. Furthermore, if the action is not limited just to the company generating the data, but the data is available via APIs to ecosystem partners for them to act, this has the potential of increasing value tremendously.

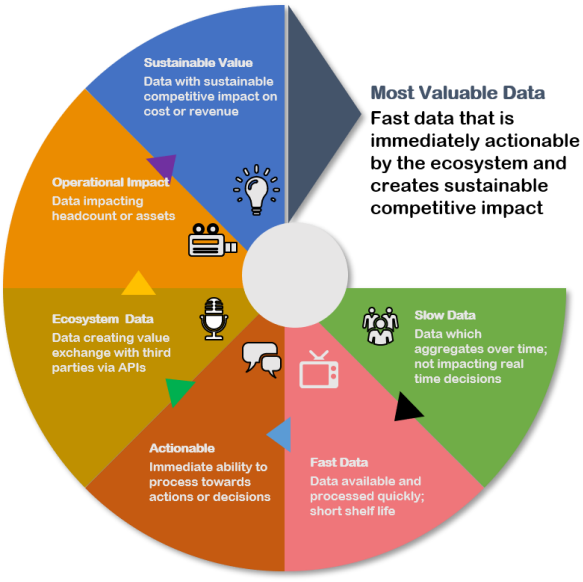

Figure 2: AV Data Value Analysis

The next levels in the value chain are the types of value created by the data – i.e. do they provide an operational advantage by streamlining operations or can the data provide sustained value from a competitive standpoint. The most valuable data has the characteristics of being fast, actionable by the ecosystem and providing sustained competitive impact.

AV Data: A Monetization Framework

As we just saw, the useful life, actionability and ecosystem impact of the data determines its value - and any data that creates sustained value obviously has significant monetization potential.

The monetization value can be simplified through a framework of the useful life of data and the breadth of usage of that data – i.e. is it used within the company on simpler use cases or available programmatically through APIs to the ecosystem for more complex use cases.

What this means is that monetization potential is enhanced if data is consumed as a service to the ecosystem via APIs – i.e., value of the data increases because it is now available and actionable to the AV mobility ecosystem. And consequently, this data now increases ecosystem value and vice versa, thus creating a sustainable competitive advantage for the AV vendor that controls and manages that mobility ecosystem.

The following framework can help the design of data driven business models for mobility and autonomous driving. As shown in Figure 3, Fast Ecosystem data is most valuable while Direct Slow data is generally the least valuable. Monetization of each requires different approaches and capabilities within an organization – below are some use cases and examples of how AV mobility organizations can leverage available data to create value.

Slow Internal

Use cases: An autonomous car fleet provider uses internal vehicle data, passenger pickup/drop off data to negotiate for better insurance rates or maintenance contracts. In this case, the fleet company can make a case that insurance rates should be lower when there are no passengers in the car. This data could be collected over a period and across vehicles types, geographies. While

some of this type of data could have been actionable at creation, much of the value from a cost or revenue perspective, is only realized in aggregation later.

Slow Ecosystem

Use cases: Automobile or AV data is collected, aggregated, and provided to 3rd parties and their applications for analysis. Imagine the same data collected by the above AV fleet provider now being programmatically available to its ecosystem e.g., battery charging partners, maintenance network, or another specialized sub-fleet servicing college students in a specific university town. Historic or seasonal demand and ridership data made available via APIs can allow the battery charging firm to provide dynamic pricing, the student sub-fleet to better position its vehicles throughout the town for maximum utilization and ridership.

Fast Internal

Use cases: Internal data about auto health, external data received from 3rd parties about temporary road construction are examples. While auto health data is not real-time sensitive, it has an immediacy that could impact the health and costs or revenue opportunities. Construction data is captured as it flows inwards and baked into the Maps for immediate navigation. The useful life lasts if construction continues. Useful life of data such as traffic jams, weather and congestion are much shorter and temporary.

One could also include the data collected by the sensors to enable the vehicle to drive autonomously in this category. The data is directly acted upon, to make decisions by the local AV stack. However, the ability to monetize this internal data, as safe motion, is limited to the AV stack owner.

Fast Ecosystem

Use cases: When immediacy and actionability of the data adds significant value. Imagine a company providing an autonomous vehicle service to take young children to school and back. The car will depend on the manufacturers' AV stack to navigate, while the company develops its mobility services on top of the data made available by the stack. When the car gets to school, the critical safety question that needs answering is "Who is allowed to open the car door?" A photo of the person opening the car door can be made available in real time – and is from real-time camera data accessible from the AV platform APIs. The mobility company applications, running both locally in the car and in the cloud, can confirm the identity of the person trying to open the door. If the validation is successful, the company application makes a secure request through the AV stack API to open the door – and the car executes the physical action and safety critical command of opening the car door.

In this use case, while the 'size in bytes' of a photo is small, the value of this interaction is very high from a value per byte of data perspective. This is an example of the high value of fast data that is actionable immediately by the ecosystem.

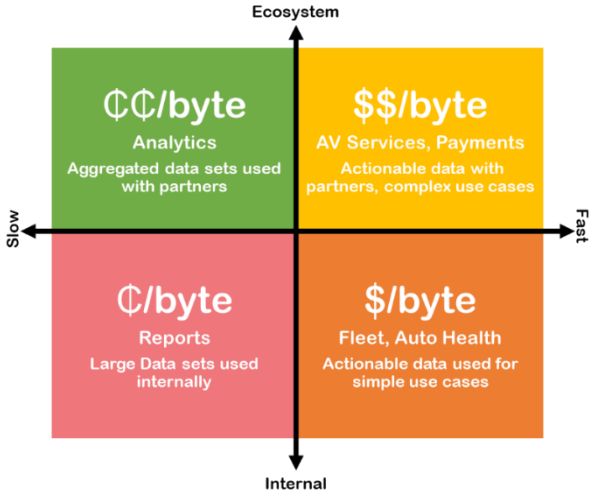

Figure 3: AV Data Monetization Framework

Broader Applicability:

While we have discussed data value in the AV context, the value of this data framework applies broadly to other industries and many IoT use cases. Examples include stock prices for financial industry, spot prices of tomatoes for the farmer in the agricultural sector, ticket pricing for travel, traffic and demand on an ecommerce website, weather data etc.

Conclusion

More data is not necessarily valuable – data is the new oil is a myth since all data is not equally valuable and there exists significant variability of the value per byte of data. The value of data depends on its direct impact to cost, asset efficiency, headcount productivity or revenue. In most industries, fast and immediately actionable data available over programmable access for the ecosystem is the most valuable. In summary, these frameworks are also applicable to broad use cases of big data value, outside of autonomous vehicles.

About the Authors:

Forbes | Technology Council 2020 OFFICIAL MEMBER

Nitin Kumar is a 20-year veteran in the Hi-Tech industry. He is currently the CEO of Appnomic but played a variety of hands on executive roles ranging from CEO, Chief Growth Officer, Chief Transformation Officer, M&A Integration/Separation Leader, BU Head and Management Consulting Partner (corporate and PE portfolio companies). Nitin Kumar is a member of the Forbes Technology Council and shares his ideas and thoughts on the forum regularly. In his role as a former Management Consulting Partner, Nitin has done multiple strategy and M&A engagements for Software, Hardware, Semiconductor and AutoTech sectors gaining invaluable insights in the value chain, technologies, and business models. He is also a Certified Autonomous Driving Professional.

Manu Namboodiri has for 20 years worked across industries such as autonomous vehicles, security, IoT, software and has broad experience ranging from strategy, product, marketing, and ecosystem development. He resides in the San Francisco Bay area and advises companies in various stages of market adoption.