A WHITE PAPER ON ARTIFICIAL INTELLIGENCE & BIG DATA IN TRANSPORTATION

CENTER FOR TRANSPORTATION RESEARCH

August 2018

PRELIMINARY REVIEW COPY

Technical Report Documentation Page

1. Report No. FHWA/TX-18/0-6806-CTR-4	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle A White Paper on Artificial Intelligence & Big Data in Transportation		5. Report Date August 2018
		6. Performing Organization Code
7. Author(s) Amy Fong, David Arredondo, Hali Hoyt, Andrea Gold, Kristie Chin, C. Michael Walton		8. Performing Organization Report No. 0-6806-CTR-4
9. Performing Organization Name and Address Center for Transportation Research The University of Texas at Austin 3925 W. Braker Lane, 4th Floor Austin, TX 78759		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. 0-6806-CTR
12. Sponsoring Agency Name and Address Texas Department of Transportation Research and Technology Implementation Office P.O. Box 5080 Austin, TX 78763-5080		13. Type of Report and Period Covered Technical Report May 2018- August 2018
		14. Sponsoring Agency Code
15. Supplementary Notes Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration.		

16. Abstract

Advances in computing techniques and processing capacity as well as increased data collection are beginning to enable artificial intelligence applications in a myriad real-world setting. Artificial intelligence (AI) algorithms at their most advanced can provide decision support, ease labor-intensive operations, perform predictive analysis, and inform targeted outreach. In the transportation sector such applications could reduce the administrative burden at public agencies such as TxDOT and the DMV, and collect higher resolution traffic data with less infrastructure, thus enabling detailed transportation planning models and predicting and identifying traffic incidents. Artificial intelligence is also being applied to traffic control devices, and preliminary deployments have been promising. However, with the advent of advanced models and the significantly higher quantity of data they typically consume and produce, key challenges will include managing complex data sources, ensuring their ethical application in decision-making, protecting the privacy of the public, and reducing cybersecurity risks. This white paper provides and overview of key technologies that are enabling AI, a menu of AI applications across five transportation application areas, and case-studies from deep-dive interviews with technology companies.

17. Key Words Artificial intelligence, big data, emerging technology, machine learning, transportation technology, emerging data sources, internet of things, service planning, predictive maintenance		18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161; www.ntis.gov.	
19. Security Classif. (of report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of pages TBD	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

THE UNIVERSITY OF TEXAS AT AUSTIN
CENTER FOR TRANSPORTATION RESEARCH

A White Paper on Artificial Intelligence & Big Data in Transportation

Amy Fong, David Arredondo, Hali Hoyt, Andrea Gold, Kristie Chin,
C. Michael Walton

CTR Technical Report: 0-6806-CTR-4
Report Date: August 2018
Project: 0-6806-CT
Project Title: TxDOT Administration Research
Sponsoring Agency: Texas Department of Transportation
Performing Agency: Center for Transportation Research at The University of Texas at Austin

Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration.

Center for Transportation Research
The University of Texas at Austin
3925 W. Braker Lane, 4th Floor
Austin, TX 78759

www.utexas.edu/research/ctr

Disclaimers

Author's Disclaimer: The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official view or policies of the Federal Highway Administration or the Texas Department of Transportation (TxDOT). This report does not constitute a standard, specification, or regulation.

Patent Disclaimer: There was no invention or discovery conceived or first actually reduced to practice in the course of or under this contract, including any art, method, process, machine manufacture, design or composition of matter, or any new useful improvement thereof, or any variety of plant, which is or may be patentable under the patent laws of the United States of America or any foreign country.

Notice: The United States Government and the State of Texas do not endorse products or manufacturers. If trade or manufacturers' names appear herein, it is solely because they are considered essential to the object of this report.

Engineering Disclaimer

NOT INTENDED FOR CONSTRUCTION, BIDDING, OR PERMIT PURPOSES.

Research Supervisor: Michael R. Murphy

Acknowledgements

The research team would like to express thanks to the Texas Department of Transportation Research and Technology Implementation Office for identifying this timely topic for research and for coordinating and support in gathering guidance from TxDOT staff. The research team also appreciates and values the insights and support from our invited subject matter experts who participate in meetings and key person interviews. The authors would also like to thank the following Texas Department of Transportation staff for their support:

Darran Anderson, Chief Strategy and Innovation Officer

The Office of Strategic Planning
Kent Marquardt, Director
Stacey Strittmatter, Deputy Director
Yvette Flores, Strategic Research Analyst

Research and Technology Implementation
Rocio Perez, Research and Technology Implementation Director
Kevin Pete, Research Project Manager
James Kuhr, Research Project Manager

Information Management Division
Dana Glover, Director
Cassie Jordan, Director of Strategy and Portfolio Management

TABLE OF CONTENTS

EXECUTIVE SUMMARY

A new wave of technological innovation has brought new opportunities and a new vocabulary into the world of transportation. Many useful and innovative solutions emerge every day that rely on cutting-edge technology, and for many of these promising solutions, the underlying technology is artificial intelligence, or AI. Innovation in this area will maintain its rapid pace for the foreseeable future, increasing both the difficulty and value of defining the impact of AI on the transportation space.

Recognizing that innovation will continue apace for the foreseeable future, this paper focuses on highlighting the various types of AI and AI-related transportation tools. Key findings from this perspective are the importance of data management infrastructure—both hardware and software—and the rising importance, bordering on need, of data scientist personnel. Many of these tools promise to improve efficiency, reduce costs, or increase capabilities. Some of these tools have delivered on this promise even in their infancy. Regardless of what problem these tools help solve, all are best utilized when an agency has resources and personnel dedicated to utilizing data insights.

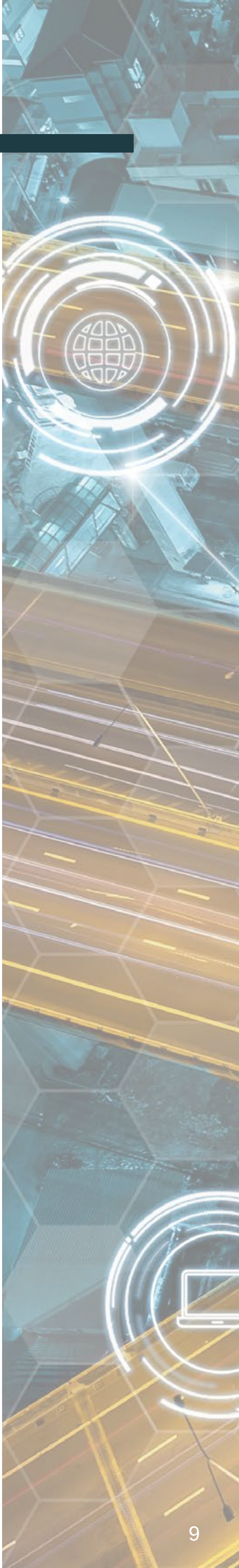
KEY TAKEAWAYS

Use flexible, open architecture databases.
In the era of big data, agencies will be well served by proactively adopting flexible and open architecture databases that will enable adaptation, scaling, and continuous availability to data stores.

Develop data governance across agencies.
Data governance provides public agencies with strategic guidance on the roles and responsibilities that are needed to ensure clear processes for collecting and reporting agency data and also helps to ensure accountability for data quality and security.

Extend data value via cross-sector applications.
With the increased availability of data there is increased opportunities to solve complex problems that span multiple domains. Transportation agencies can collaborate with agencies to look at transportation data crossed with energy, public health, and housing data, and more.

Build and integrate data science expertise.
Public sector organizations increasingly need to extract actionable insights from data in various formats from various sources. Although public agencies consistently face resource and budget constraints, it is important to hire and retain in-house data science expertise for data-driven decision making.



MOTIVATION

RESEARCH GOALS

- Inform and help make decisions
- Describe challenges and outline paths to overcome them
- Objective evaluation and presentation of options
- Introduce new concepts and distill technical content
- Describe business advantages to implementing AI solutions

Transportation planning has traditionally taken a long view—on the scale of decades—but the industry has recently been forced to turn its attention to technologies and service models that continue to emerge ever more rapidly. In the span of a few short years, private transportation network companies and other shared mobility service providers have commanded a not-insignificant mode share among the traveling public. Disruption from technology companies and startups has demanded that transportation agencies adapt the ways they plan for, regulate, and operate future infrastructure systems.

Private transportation service providers have been successful by leveraging the latest advances in information technology, such as AI, big data, and Internet of things (IoT). Although public agencies have vastly different goals from private companies, given the agencies are responsible for providing reliable and widely available infrastructure and services, they can borrow tactics from the private sector to deliver those public goods and services more efficiently or effectively. For a public agency, being able to adopt or emulate certain principles that led to private companies' success, successfully collaborating with those companies, or both, runs through understanding. Cultivating an awareness of emerging technologies and how they work best positions transportation agencies to adapt to the near constant changes in transportation technology occurring today.

AI and ancillary technologies like cloud computing and IoT provide new ways to use and generate information. For example, video data is now more available and more valuable. The reduced cost of quality cameras and viability of signal broadcasting equipment accounts for the greater availability; the greater value is a result of machine learning's ability to increasingly automate the task of continuously counting and evaluating everything a camera sees. Another example of leveraging traditional data sources through AI is in the area of surveys, which are rendered more powerful through cloud computing-powered algorithms and the ability to digitally merge many sources of data. These and other examples of data gathering and processing promise to either increase efficiency or lower the cost of transportation tasks.

From this point on, private-sector innovators in the transportation sector will only multiply: we can anticipate increasing numbers of technology demonstrations and deployments that promise their machine learning algorithms can deliver improved data collection, processing, and prediction paradigms or that their AI implementation will power novel transportation modes or services. In the meantime, public agencies will need to train their focus on continuing to provide access to transportation, essential services, and economic opportunity for all members of their communities. In some instances, not all products marketed towards to the agencies will necessarily be practical or useful. Instead, it is imperative that public agencies develop flexible standards and technologies, leveraging technology solutions that will allow public agencies to meet their responsibilities to the public more effectively or efficiently. Agencies must maintain an awareness of the ever-evolving technology sphere so that that they can partner meaningfully with private-sector innovators.

PROCESS

OVERVIEW

This paper aims to filter through some of the vast capabilities of emerging technologies that use artificial intelligence or machine learning, and provide a focused look at the applications relevant to transportation agencies.

This paper aims to filter through some of the vast capabilities of emerging technologies that use AI or machine learning, and provide a focused look at the applications relevant to transportation agencies.

First, we define our vocabulary, and demystify technology buzzwords, such as big data, edge computing, cloud computing, and AI, in the Technical Primer. This contextual base will allow the reader to synthesize the more complex topic areas to follow. The Technical Primer also outlines the historical trends that caused AI and adjacent technologies to go from theory to best-in-class products in less than a decade.

Next, this paper considers five distinct application areas that are broadly relevant to transportation agencies:

- system and service planning
- real-time system performance
- public safety and enforcement
- construction and asset management
- public administration and information management

These five areas represent distinct domains that help link different technological tools to common, specific goals. This organizational choice highlights the domains that AI most influences, as well as those domains in which AI plays a lesser role. Some domains, such as public administration and information management, have not been heavily influenced by AI. Others, like system and service planning and real-time system performance, were so influenced by AI that many companies now work in both. The Menu of Applications section will also shed light on which companies and startups have already taken innovative products to market, what types of data and analytics tools they used, and how some public agencies have started deploying or incorporating them into existing systems.

A series of company case studies follows the Menu of Applications. These case studies include information learned from deep dives on distinct technology solutions, and offer a glimpse into a real-world implementation of each technology. Regardless of whether the companies in this section solve the same or similar problems, the underlying technology differs widely from case study to case study. Graphics are provided at the end of this section in order to visually represent the differences, and similarities, between companies and technologies.

During each of the case study interviews, we asked each company what advice they would give to transportation agencies. Every single company recommended investing in data management as a whole, or some specific facet thereof. Our conclusion focuses heavily on this finding.

TECHNICAL PRIMER

OVERVIEW

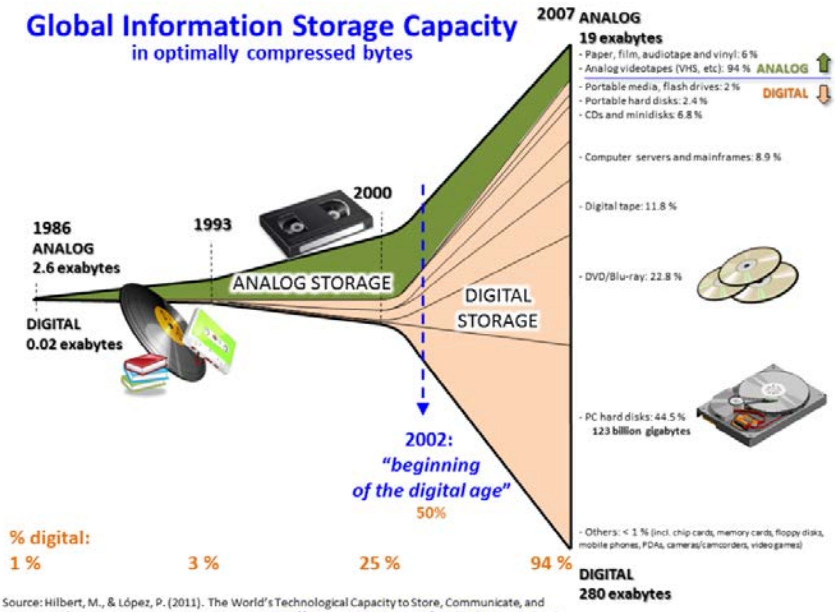
The objective of the Technical Primer is to demystify commonly used artificial intelligence buzzwords to give the reader a more contextualized view of the research being presented.

Artificial intelligence (AI) and big data along with cloud and edge computing have developed so rapidly that their names signify many things, including hype, hope, skepticism, and the promise of a more automated, efficient, and informed future. AI consists of complex mathematical underpinnings that can initially be difficult to digest for newcomers to the field of computing. Its capabilities and applications have been unlocked by hardware and software advances in the past two decades. This chapter explains how and why technology has developed so quickly in the last twenty years, and explains why these new technologies are so impactful.

Big Data

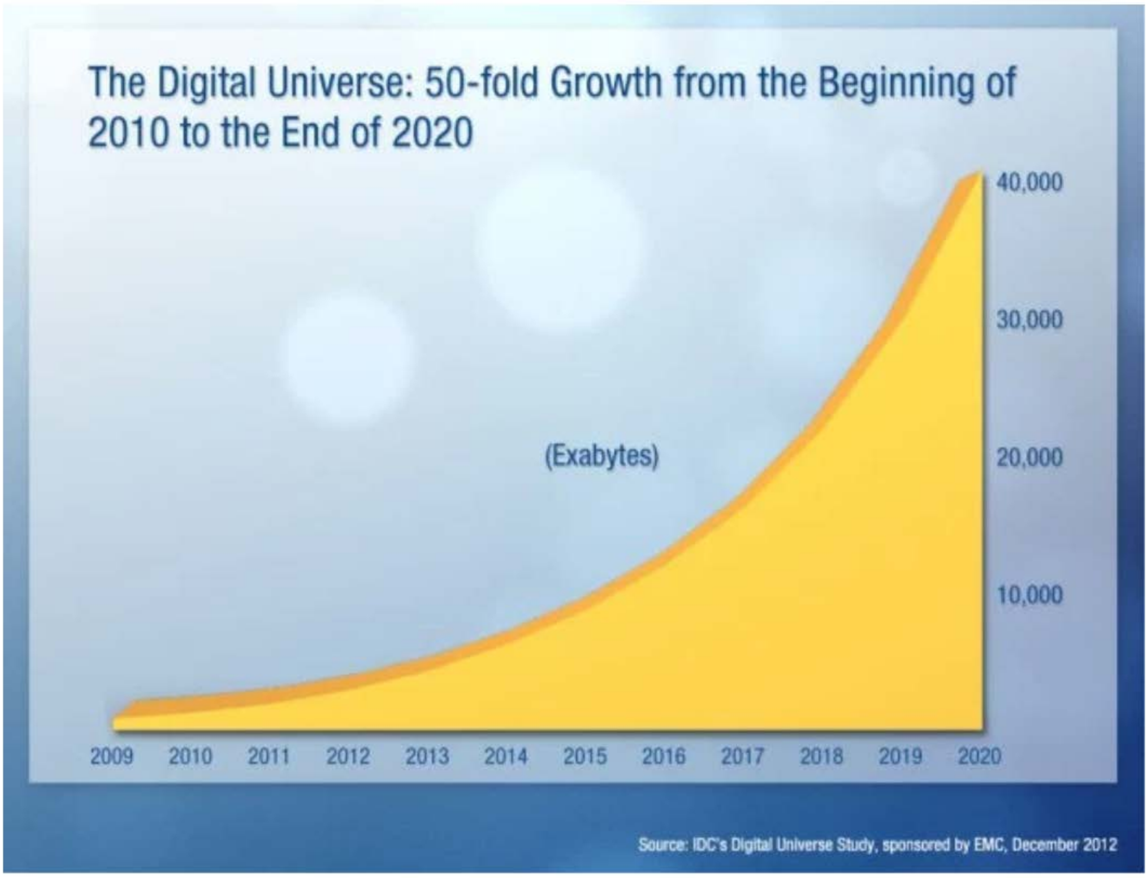
Big data's possibility stems from a historical trend known as Moore's Law (neither a physical law nor a legal one), which predicts that computing power and storage will double every two years. Moore's Law is at best a reasonably accurate boast from the computer processor industry. While the actual rate has fluctuated over time, this self-enforced innovation standard in the computer processor industry has transformed the way all parts of society store and interact with information.

The image below illustrates the exponential growth of data storage capacity.



Source: Hilbert, M., & López, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science*, 332(6025), 60 –65. http://www.martinhilbert.net/WorldInfoCapacity.html

Note that the above graphic is quite dated. Over the last ten years data storage has grown further—making even relatively new technologies like CDs inadequate for the current data environment. The difference between the graph above and the graph below illustrates this change. Ten years ago, it was possible to comprehend the amount of data in the world, and to interact with physical objects that stored data. Now, it's almost impossible to comprehend the amount of data that one business generates monthly—let alone the amount of data in the world.



From: https://www.emc.com/leadership/digital-universe/2014iview/index.htm

The image below illustrates the exponential growth of data storage capacity.

The sheer size of many organizations' data pools makes previously simple tasks like querying and summarization difficult. Almost paradoxically, it's difficult to identify and explore the insights that lie within. Even when given enough servers and computers to store data, finding the relevant information can still take a great deal of time and domain expertise.

The method of data collection and generation often produces unstructured data, further increasing computing and processing resources needed to process the data. Unstructured data is unorganized and, even with computing advances, it typically cannot easily be converted to more structured forms with human help. For most modern machine-learning algorithms, data must be organized into rows and columns for model ingestions. What's more, most conventional database tools that use Structured Query Language (SQL) also require data in row and column format in order to initialize and utilize. Another issue that arises with big data is that modern data often enters databases at a tremendous rate. Data from smartphones and IoT (Internet of things)-type sensors, and in the field (such as traffic cameras, or in-road sensors) comes in at a near-constant flood. This is one of the main reasons data often enters a system without any organization; at the time of collection, there simply isn't time to organize it.

The sheer size of many organizations' data pools makes previously simple tasks like querying and summarization difficult. Almost paradoxically, it's difficult to identify and explore the insights that lie within. Even when given enough servers and computers to store data, finding the relevant information can still take a great deal of time and domain expertise.

The method of data collection and generation often produces unstructured data, further increasing computing and processing resources needed to process the data. Unstructured data is unorganized and, even with computing advances, it typically cannot easily be converted to more structured forms with human help. For most modern machine-learning algorithms, data must be organized into rows and columns for model ingestions. What's more, most conventional database tools that use Structured Query Language (SQL) also require data in row and column format in order to initialize and utilize. Another issue that arises with big data is that modern data often enters databases at a tremendous rate. Data from smartphones and IoT (Internet of things)-type sensors, and in the field (such as traffic cameras, or in-road sensors) comes in at a near-constant flood. This is one of the main reasons data often enters a system without any organization; at the time of collection, there simply isn't time to organize it.

These issues in modern data storage and processing can be summarized with a pithy bit of alliteration: volume, variety (the types of data), velocity (how quickly data enter the system), veracity (how noisy the data are), and value.

To place this section in context, recall an old adage: knowledge is power. The more data, the more potential knowledge; with more knowledge, the potential to make better decisions, and so forth. Large information stores (big data) have the potential to increase efficiency and reduce cost—but there is no such thing as a free lunch.

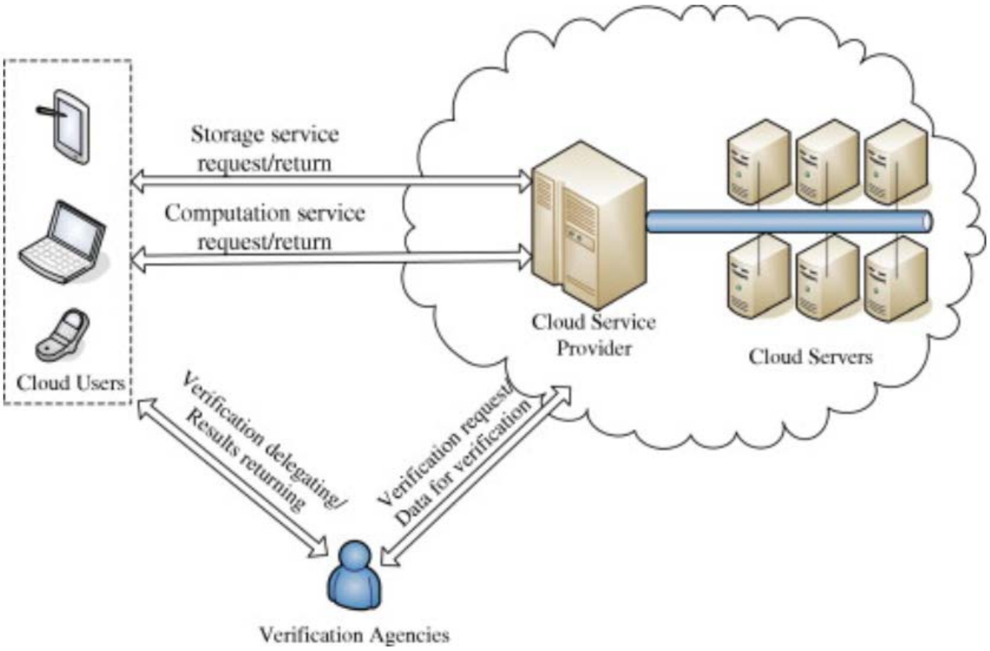
Key Takeaway: Big data is data that have any of volume, variety, velocity, veracity, and value in problematic amounts, requiring new and different tool to store, manage, clean, and analyze.

The Cloud
Data need a physical place to live. This place is often on computers, but it can also be on a six-foot-tall server, or a flash drive smaller than a pinkie finger. Advances in technology have made even the small devices, such as flash drives and smartphones, capable of housing dense data files, like videos. But when video data comes in 24 hours a day from something like a traffic camera, even a traditional desktop computer, with many times the storage capabilities of a smartphone, ends up overflowing with data.

In order to handle the data flow, more servers and technical personnel must be employed to manage the data—but these servers and personnel need not be employed on location. The Internet can provide a safe and efficient method for connecting clients to a large, off-site server farm capable of handling as much data as the modern world can generate. The same companies that provide the servers can also provide software and expertise to analyze and organize the big data the servers store.

Large software and hardware companies end up housing cloud computing infrastructure, because often large-scale infrastructure is needed to run internal operations. Some of the major players offering cloud services are Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform, and Oracle Cloud. Many other companies offer similar services, but the aforementioned providers often have more capacity, computing power, and a greater range of tools and services (analytics).

Although most cloud services are as secure (and sometimes more secure) than hosting data on local servers, using a cloud service means that another organization has access to the data stored within their servers. The graphic below outlines the relationship.



From https://www.sciencedirect.com/science/article/pii/S0020025513003320

Key Takeaway: Cloud computing is the act of employing remote computers—often managed by another company—to solve problems or store data via the Internet.

Note: Cloud computing often connotes big data, but the phrase "the cloud" can mean any size enterprise. For instance, Google Drive is a cloud service, but generally handles personal data, the type and content that would easily fit on a laptop.

Internet of Things and Edge Computing

The Internet of things (IoT) is a term that refers to putting small footprint computers in everything from street lights to the wheels of a bus. Edge computing means that those tiny computers are fast and powerful enough to clean data, i.e., they can help with the veracity and variety parts of big data. Edge computing used to mean cloud computing, but its meaning has changed, and it will most likely change again.

That said, much of the data that teaches machine-learning algorithms comes from IoT applications. Having the best computers in the right places deliver good results; a sensor in every bus wheel that constantly monitors tire pressure and general wear and tear provides actionable information that would otherwise not be accessible. It's even better if that sensor in the bus wheel has a tiny computer that emits only clean, easy-to-use data.

However, it cannot be overstated that the IoT, or edge computing, are upgrades in physical infrastructure. IoT means that actual tiny computers are installed and maintained inside all manner of objects, and that these tiny computers are beaming data at a near constant rate. Such data often requires immense data storage and data management capabilities, which is to say that IoT often generates big data. As previously mentioned, big data requires a bevy of servers—either an on-premise data center or cloud computing—in order to manage the sheer quantity.

Key Takeaway: IoT promises to supply copious amounts of useful data, so long as the underlying infrastructure is properly implemented and maintained.

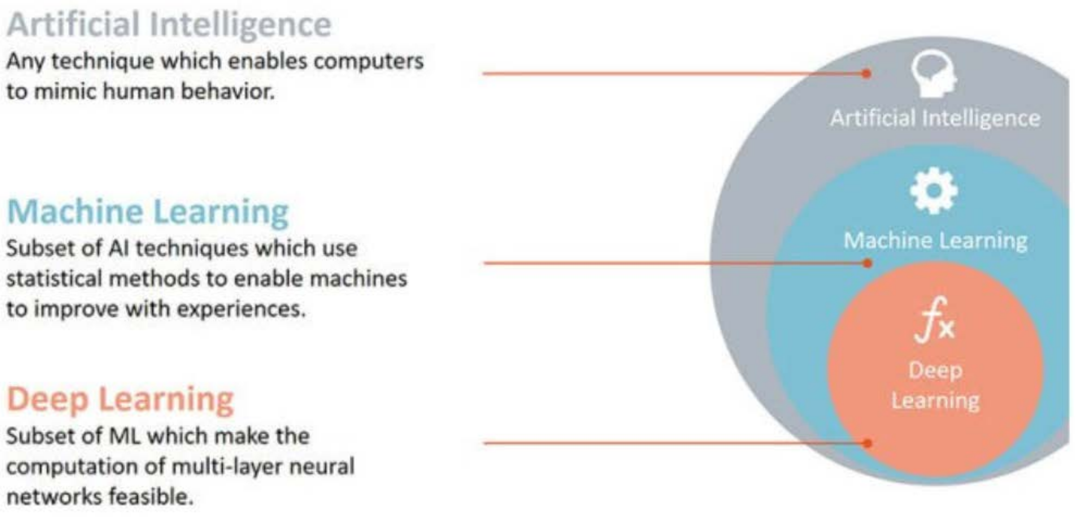
Cloud Computing and Big Data: Enabling AI

In the twentieth century, AI proved too rigid to be applicable to real-world problems and too reliant on data and computing power to be feasible. However, both big data and the cloud have removed these barriers and paved the way for AI to garner insights and increase performance across operations not only within the private sector, but also in the public space.

The best way to leverage big data is to utilize AI to glean constant, actionable insights from it. Combined with access, via the Internet, to a whole host of servers, AI can distill large amounts of information into meaningful facts. It can turn unending hours of video data into counts of every car, truck, and bicycle that went through an intersection.

Artificial Intelligence (AI)

The next section, and the rest of the primer, outlines how and why AI transforms data and computing power into useful knowledge about the world.



Artificial Intelligence
Any technique which enables computers to mimic human behavior.

Machine Learning
Subset of AI techniques which use statistical methods to enable machines to improve with experiences.

Deep Learning
Subset of ML which make the computation of multi-layer neural networks feasible.

From https://www.kdnuggets.com/2017/07/rapidminer-ai-machine-learning-deep-learning.html

Historically, there have been two major approaches to creating AI: expert knowledge systems and machine learning.

Expert knowledge systems are lists of rules written by an expert. The intuitive appeal of these systems lies in their exactness; the computer knows exactly what to do, because the rules are explicitly written. But their exactness freezes expert knowledge systems into a rigid position. Lists of rules leave no wiggle room for navigating the unknown. For example, consider the game of chess.

What then is machine learning and its relationship to modern AI? The magic comes from humility. We do not know everything, but we do have a good idea how likely something is, thanks to experience.

For example, consider darts, a game with a defined goal: hit the bullseye. On the first dart throw and without prior skill, all players are equally likely to hit any part of the dart board—if they hit the board at all. Given time to practice, players will become more accurate, which is to say, more likely to get closer to the bullseye. Now, the goal does not have to be hitting the bullseye. The game can be made easier (hit the board at all), or harder (professional dart throwers have to hit different parts of the dart board as the game goes on), and still, given time to practice, a person will get better at the game—the darts will get closer to the goal.

But where exactly is statistics, in all this dart throwing? Think about our game of darts in the following manner: experience makes us more likely to hit our goal. Each throw yields a result (success or failure), but it also provides data, information about what was tried. A person remembers how hard they threw the dart, where they were looking, how they were holding the dart, etc. These different factors, or variables, can be altered to improve performance. Statistical machine learning, or experience, tells us how much to alter these variables. After throwing a dart 1000 times at a dart board, a person remembers how to throw a dart in order for it to have the greatest chance of hitting a bullseye.

Without practicing, that player does not know how hard to throw the dart. An expert was not born an expert; they learned all their skills via experience. Likewise, without data, a statistical model does not know how to solve a problem; given enough data, a model can achieve expert level accuracy. Whether human or machine, expertise comes from having enough humility to learn from failure.

Returning to the chess example, let's again consider the humility necessary to learn. We will never know the outcome of every (or even many) moves beforehand, but we know some things:

- we have a goal: put a piece on the same square as the opposing king.
- we have a penalty: if an opposing piece lands on the same square as our king, then the game is over, i.e., it becomes impossible to achieve the goal.
- actions, such as changes in piece position or elimination of pieces, alter the likelihood of achieving the goal.

These rules form a model, which needs only the ability to execute actions and remember how close those actions brought it to its goal. In the beginning, this model isn't very good, because it doesn't know that a queen is often more important than a pawn. But as this model plays more and more games, it remembers what actions brought it closer to its goal, and begins to value a queen more than a pawn, a rook more than a bishop, etc. After playing a million games, the model would have more chess playing experience than any expert.

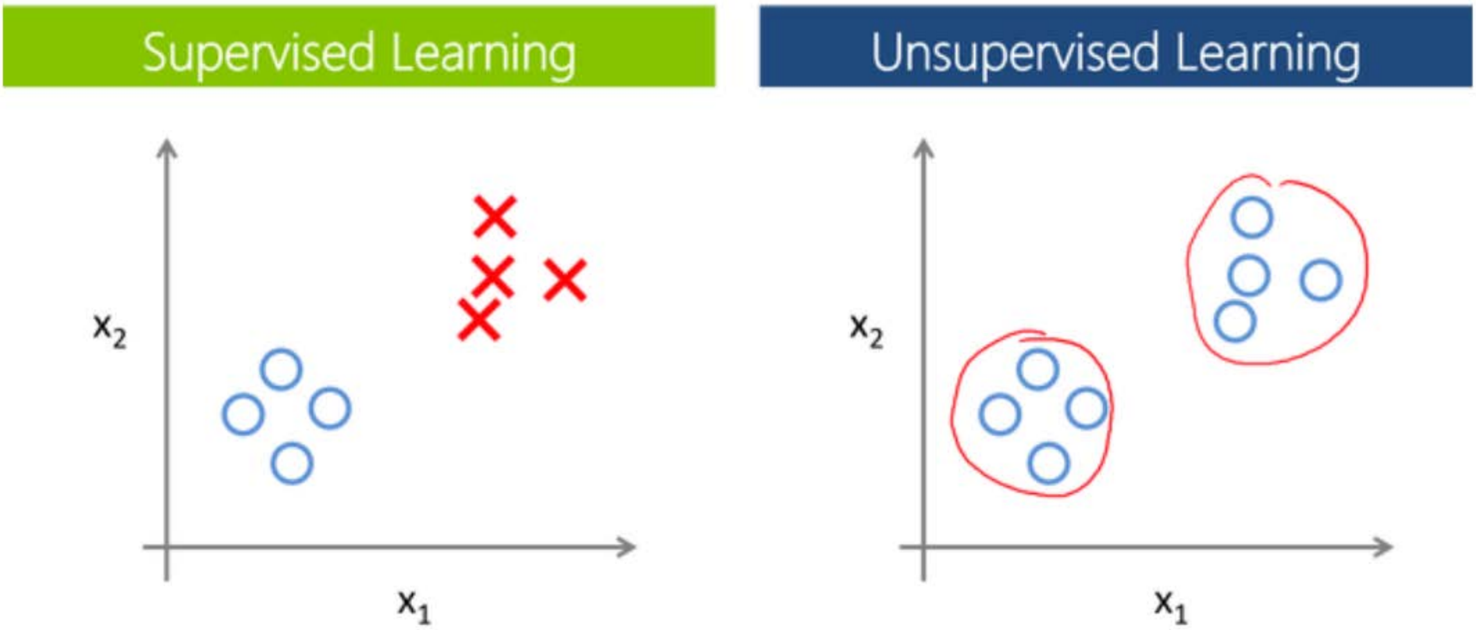
Remember that humans determine what any machine-learning model's goal is. Goals can range from predicting the stock market, to generating funny recipes. Recall that people invented chess; machine-learning algorithms only ever do what you ask them to do. In machine learning, we relinquish control over how the model achieves its goal, giving it the ability to learn from experience.

Supervised versus Unsupervised Learning

Extend the humility learned in the chess example to an almost entirely different problem: labeling an image as a cat or a dog. Instead of trying to list each possible way a cat is not dog, one could instead give a computer a bunch of images labeled either dog or cat. Like in the chess game, it is only needed to give the computer examples of what success looks like. The computer then tries to label an image as cat or dog, and learns after it's done whether it succeeded or not. As in in the chess game, it remembers what it did, and whether what worked or not.

Using human knowledge to define, or label, the solutions to a problem for an algorithm is called supervised learning, since a person supervised—or had input—in the learning process.

Another class of machine-learning algorithms, called unsupervised learning, does not need any human input on what is right or wrong. This comes with a trade-off: these algorithms do not solve defined problems. Instead, they summarize the data, often making it easier for the computer to run a supervised algorithm.



These rules form a model, which needs only the ability to execute actions and remember how close those actions brought it to its goal. In the beginning, this model isn't very good, because it doesn't know that a queen is often more important than a pawn. But as this model plays more and more games, it remembers what actions brought it closer to its goal, and begins to value a queen more than a pawn, a rook more than a bishop, etc. After playing a million games, the model would have more chess playing experience than any expert.

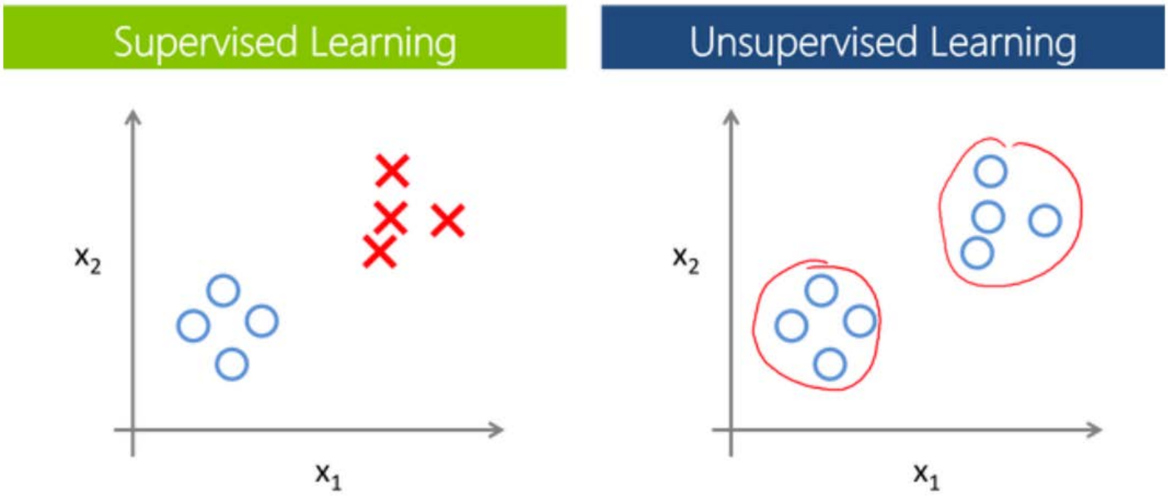
Remember that humans determine what any machine-learning model's goal is. Goals can range from predicting the stock market, to generating funny recipes. Recall that people invented chess; machine-learning algorithms only ever do what you ask them to do. In machine learning, we relinquish control over how the model achieves its goal, giving it the ability to learn from experience.

Supervised versus Unsupervised Learning

Extend the humility learned in the chess example to an almost entirely different problem: labeling an image as a cat or a dog. Instead of trying to list each possible way a cat is not dog, one could instead give a computer a bunch of images labeled either dog or cat. Like in the chess game, it is only needed to give the computer examples of what success looks like. The computer then tries to label an image as cat or dog, and learns after it's done whether it succeeded or not. As in in the chess game, it remembers what it did, and whether what worked or not.

Using human knowledge to define, or label, the solutions to a problem for an algorithm is called supervised learning, since a person supervised—or had input—in the learning process.

Another class of machine-learning algorithms, called unsupervised learning, does not need any human input on what is right or wrong. This comes with a trade-off: these algorithms do not solve defined problems. Instead, they summarize the data, often making it easier for the computer to run a supervised algorithm.



Consider the above picture. In the supervised setting, the points are labeled differently, as Xs and Os. An unsupervised setting considers a scenario where we don't know whether a point is an X or O, or maybe we do not care. Either way, there are two groups of points in the above picture. Differentiating between Xs and Os isn't necessary to find this information.

Think about unsupervised learning as a summary of data. For example, return to cats and dogs. A picture of a cat can be summarized as a series of facts: four legs, one tail, flat face, pointy ears. These facts require no prior knowledge about cats, but some of the facts, like pointy ears, can be useful for classifying cats in general once given labels. In fact, just knowing that there were pointy ears in an image makes it far more likely that that image contains a cat, rather than a dog.

Different contexts require different summaries. Image summaries might outline objects within that image, while a summary of text data might consist of word counts. Regardless, unsupervised learning seeks to highlight information relevant to the problem at hand.

Key Takeaway: Humans must label data in order for AI to learn how to solve a problem.

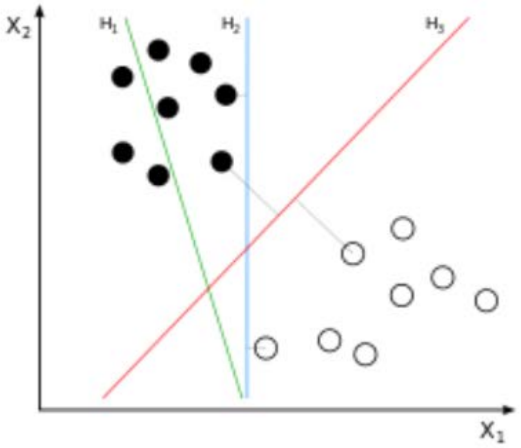
Deep Learning and Artificial Neural Networks: The Power of Complexity

A complicated model can do complex tasks, unlike binary setups that act like on and off buttons that can only label simple categories like “yes or no”, “light or dark”, and “cat or dog.” A big panel of buttons and mess of wires requires more maintenance and expertise to manage, but the buttons can now light up in a fun fancy pattern.

Until as recently as the last 20 years, complex algorithms like neural networks were, despite their promise, infeasible due to computational difficulties. If running one chess game takes a computer an hour, then it will take over one month for it to get even 1000 games' worth of experience.

Simpler options, like support vector machines (SVM) exist, but simple models underperform when asked to execute complex tasks. An SVM algorithm tries a destination between sets of differently labeled points, as in the picture below. The line marks the difference: one side of the line should have all black points, the other side of the line should be only white points. Think of the line as a decision; if your eyes cross the line, then your decision about what color the dots are changes from either black to white, or white to black. A human can look at the below graph and immediately draw a line between the two groups of points. How a person immediately knows this is not immediately apparent; people learn how to do this task in the first year of life, so the task has become so optimized that the need to try different lines (H1, then H2, then H3) seems silly and inane. Think about it this way: when first born, human infants aren't even sure if their arm is a part of them or not. Likewise, all algorithms, however complex, begin with the same amount of knowledge: none. The difference is the algorithm's learning potential, and how fast it reaches that potential. A dog can learn over 200 words, but a human can learn upwards of 60,000—even though the dog learns to walk before the human.

potential, and how fast it reaches that potential. A dog can learn over 200 words, but a human can learn upwards 60,000--even though the dog learns to walk before the human.

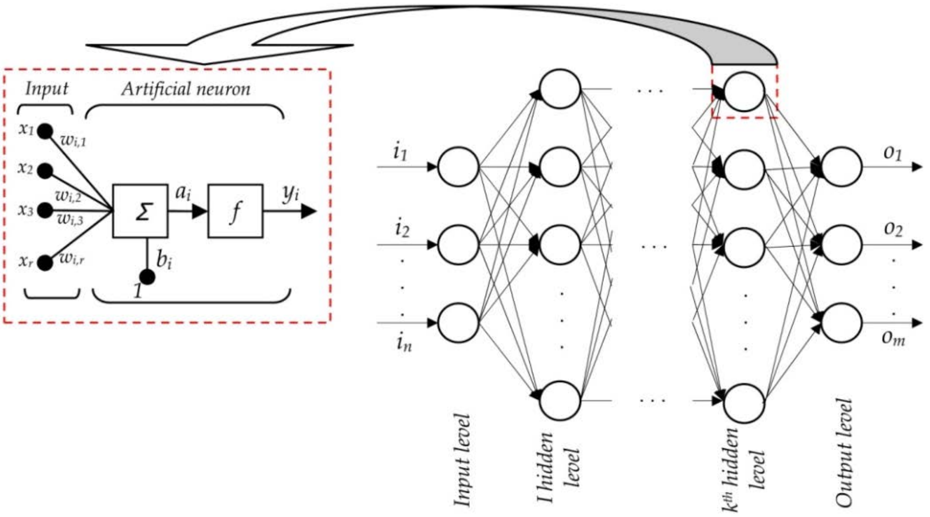


A visual example of a linear SVM from https://en.wikipedia.org/wiki/Support_vector_machine

Back to the math: SVMs, when modified, can eventually learn to complete harder tasks, such as classifying cats and dogs. However, their accuracy decreases with the complexity of the task. Even with the aid of unsupervised learning, SVMs struggle to correctly identify dogs and cats 80% of the time. And it is nearly impossible to even teach an SVM to play chess.

On the other hand, an artificial neural network can be trained to identify a cat or dog correctly over 97% of time, and its capabilities are diverse: they can beat grandmasters at chess and write text in the style of Shakespeare.

This power comes from neural networks' complexity. Each of the circles in the previous artificial neural network image contains an SVM-like model, and hundreds of these circles can be linked together into a powerful web. Networks of this size have the potential to safely drive cars, but that same complexity makes them hard to understand, and expensive to train (note: training is teaching/improving a model by giving it data; the "expense" is how long that process takes). Often PhD-level expertise and extensive infrastructure, such as access to a bevy of servers ("the cloud"), are required to design and train them.



Above image from:
https://www.analyticsvidhya.com/blog/2016/03/introduction-deep-learning-fundamentals-neural-networks/

Fortunately, neural networks require extensive time, computational resources, and mathematical expertise only while training; once finished, they can make decisions in as little as less than a second. This is as fast, or faster, than a human can process a problem and make a decision. For example, a neural network can drive a car and avoid objects as fast, or faster, than a human driver.

Use of larger neural networks is called deep learning (which refers to the "depth" in the number of layers of circles), and the phrase signifies both computational difficulty, and exceptional performance.

Key takeaway: Once fully trained, neural networks' speed, power, and performance—in certain tasks—rivals human intellect, but making them from scratch demands expertise, infrastructure, and time.

Note: There are many different types of neural networks, such as convolutional neural networks (CNN, often used in computer vision) and recurrent neural networks (RNN, often used in natural language processing). Regardless of their many differences and abbreviations, all neural networks share the same attributes outlined above.

Computer Vision

On a computer, an image/picture/video is just another form of data. If the data is labeled, i.e., you can grade performance towards completing a goal, then that data can be fed into a neural network or other algorithm. That algorithm, after processing enough data—seeing enough images—can identify objects from an image or video it hasn't seen yet.

What separates computer vision from lidar and radar is the ability to classify, or recognize, objects. Radar and lidar tell a computer where other objects are in relation to itself, but only computer vision can accurately identify a car from a pedestrian, or spot the flat lines that delineate roadway lanes.

Most algorithms for computer vision have two parts: feature selection (unsupervised learning), and prediction (training the algorithm with the data). Neural networks, with the aid of cloud computing, do both of these parts better than other any other algorithm.

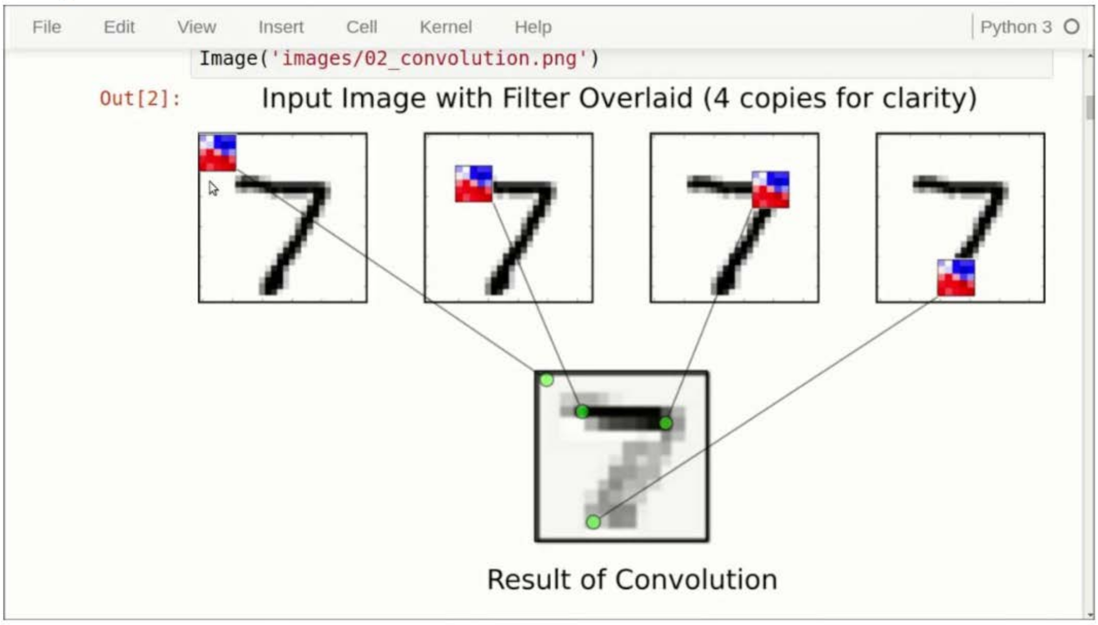
Process: one neural network makes a summary of the image or video—it finds all the relevant information. The second, with the aid of human classified data, learns to classify the images.

For a concrete example, consider a video of an intersection, and a goal: count people and count cars. The first neural network summarizes the video. For each image that makes up the video, it finds the outline of every object: the people, cars, stoplights, trees, etc. Armed with the outlines of each object, the second neural network tries to find people and cars in images that have already been classified by humans. This way, it can grade its success, and adjust its strategy based on how well its doing.

Not all computer vision applications use this two-step process. It is a matter of debate and research as to whether or not it is more powerful or more efficient to have a larger (deeper) neural network take in the raw video data and spit out counts of people and cars.

Regardless of which flavor of deep learning a data scientist chooses, the near consensus type of neural network used in computer vision is the convolutional neural network (CNN).

CNNs look at little square sections of an image at a time. Think about how humans see an image. When looking a picture of a number, the entire image is not taken in all at once. Rather, the squares with corners receive the most attention, followed by the sections with lines, etc. The background receives almost no attention at all. While this mode of vision is automatic to humans, remember that our computer is starting from scratch. Recall that our chess algorithm did not know whether a pawn or a queen was more important when it first tried to play chess. Check out the image below. Here, a CNN examines a random assortment of squares, and, through trial and error, learns where to expect a seven to appear, and where to expect background.



```
Image('images/02_convolution.png')
Out[2]:
```

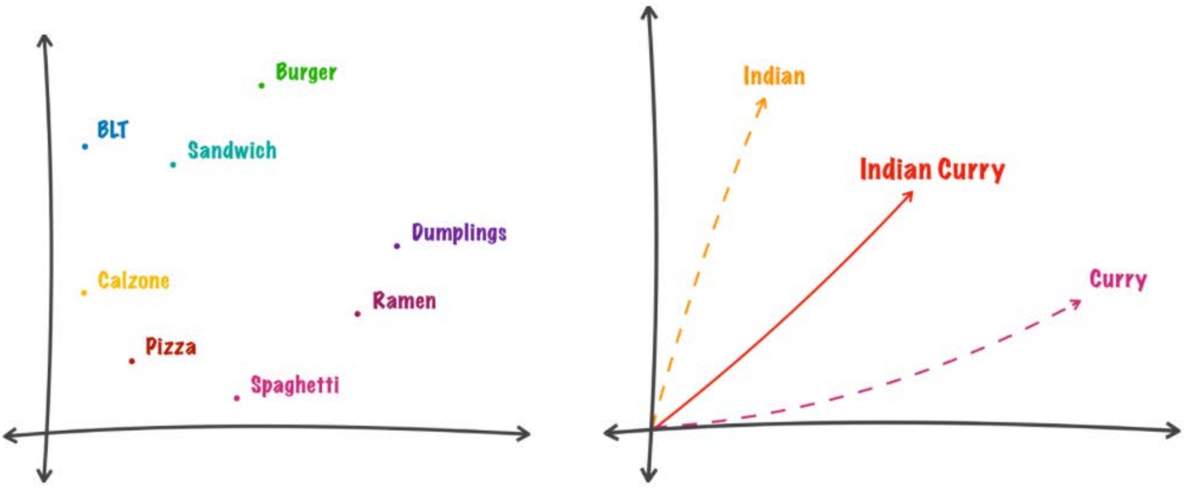
Key Takeaway: Using an algorithm (often a convolutional neural network) to identify objects in an image or video constitutes computer vision.

Natural Language Processing (NLP)
Just like images, text and audio recordings are just data to a computer. As with computer vision, if the data is labeled, then that data can be fed into a neural network or other algorithm.

For example, a neural network might be trained to identify tweets that call for help. If fed enough tweets—or bits of regular text—labeled by a human as calls for help, then our algorithm could filter through millions of tweets or other social media objects during a natural disaster to identify people in need.

Unsupervised learning often plays a key role in modern NLP. Text is often summarized, or converted, into counts (number of individual road occurrences), or more complex representations, such as vectors. Turning a word into a vector involves calculating the probability of its occurrence based on the surrounding words. Either way, unsupervised learning lets a computer know how similar words are best on how close their vector summaries are.

Examine the images below for a quick depicting of how the vector representation works.



From: https://medium.com/square-corner-blog/caviars-word2vec-tagging-for-menu-item-recommendations-13f63d7f09d8

Give these vector versions of word to a neural network, and you can expect it to write coherent sentences, or more easily understand written text.

Key Takeaway: NLP relies more heavily on unsupervised learning than computer vision or traditional analysis, but neural networks still feature in most NLP applications.

Conclusion

Computers have been getting steadily smaller, and thus steadily faster, since they were invented. Small computers, placed out in the field, can provide a constant supply of relevant data (IoT).This data teaches machines to react to the present, and predict the future, with human-like intelligence (AI)—so long as they are given access to a large amount of relevant data (big data), and they are powered with a suite of computers (cloud computing). Whenever the terms AI, big data, cloud computing, edge computing, or IoT appear, use this pipeline to sort out where their product falls in the digital landscape.

MENU OF APPLICATIONS

OVERVIEW

The following chapter summarizes existing AI technologies, with an emphasis on breadth and novelty of method. In order to tie technologies to their use cases, the following chapter contains five sections: Systems and Service Planning, Real-Time System Performance, Public Safety and Enforcement, Construction and Asset Management, and Public Administration and Information Management.

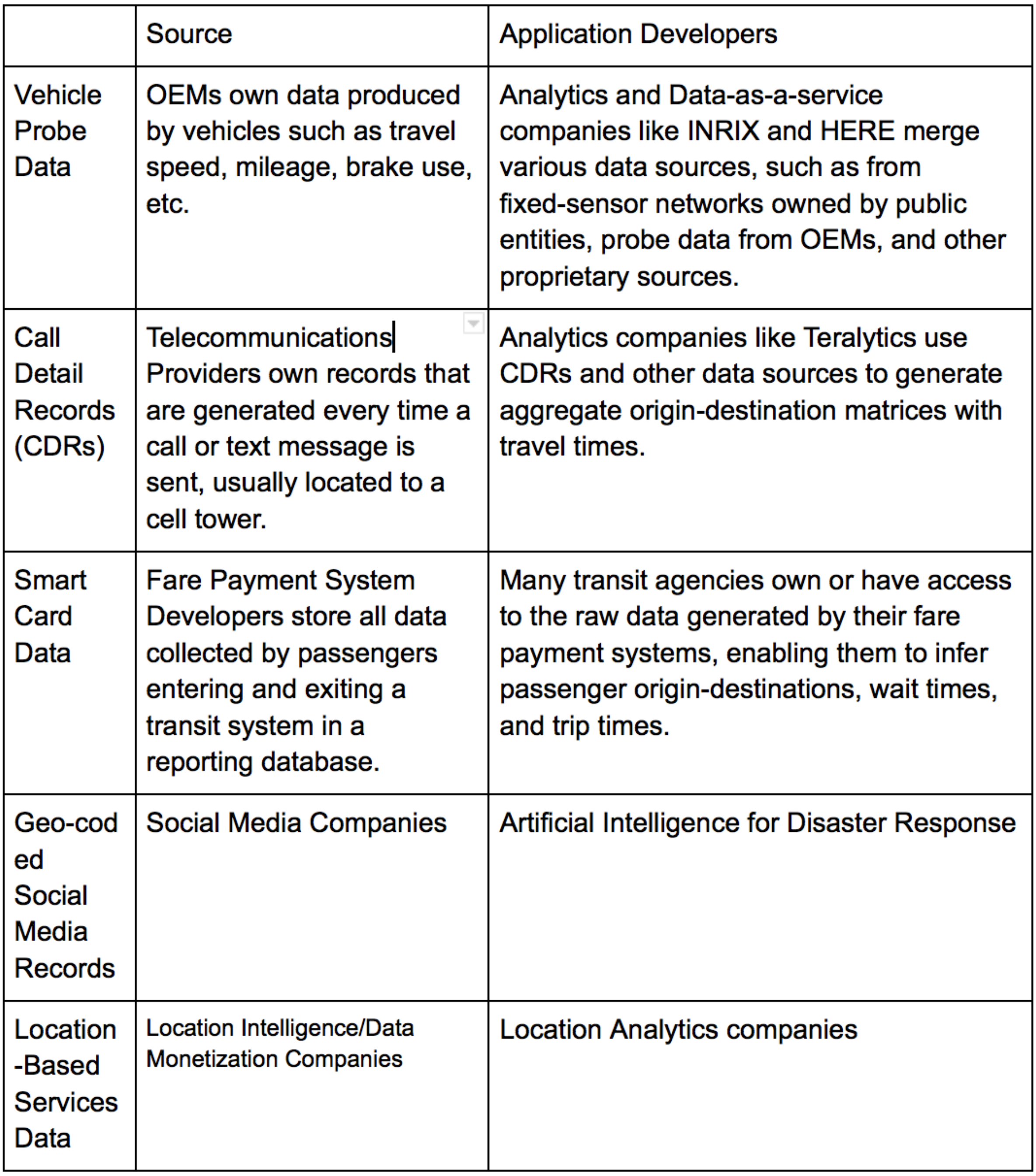
System and Service Planning

Big idea: Travel demand models and transit service planning rely on data about travelers and their travel patterns. Passive methods of collecting such data from cell phone records, Bluetooth and Wi-Fi traces, and vehicle probes. Machine-learning methods can process new and old forms of raw data into origin-destinations, route choices, mode choices, and even trip purposes. Not only will this source of big data supplement traditional household travel surveys, but the machine-learning analysis process will also expand the capability and precision of modeling predictions.

A variety of new data streams are being created as transportation infrastructure becomes increasingly digitized and smartphone adoption in the United States has grown from 35% in 2011 to 77% in 2018. These data streams include vehicle probe data, mobile phone call data records, and geo-coded social media records. Because these data streams are collected passively, they offer the opportunity to understand travel behavior at a much more granular level than traditional travel diary surveys and population censuses can achieve. These data streams could allow planners to observe, rather than infer, drivers' route choices, create more robust estimates of roadway link performance, and produce spatially and temporally disaggregate origin-destination matrices.

The sources of these big travel data vary. Some are owned by the private sector and are strategically monetized. Some researchers have been able to build partnerships with those private-sector companies, such as telecommunications providers, and have preliminarily demonstrated the viability of these datasets for application to transportation demand modeling and real-time operations. Statewide and regional planning models based on this data are beginning to be proven; for instance, the Illinois Department of Transportation has contracted with Sidewalk Labs, a subsidiary of Google, to develop a statewide travel demand model using GPS cellular data.

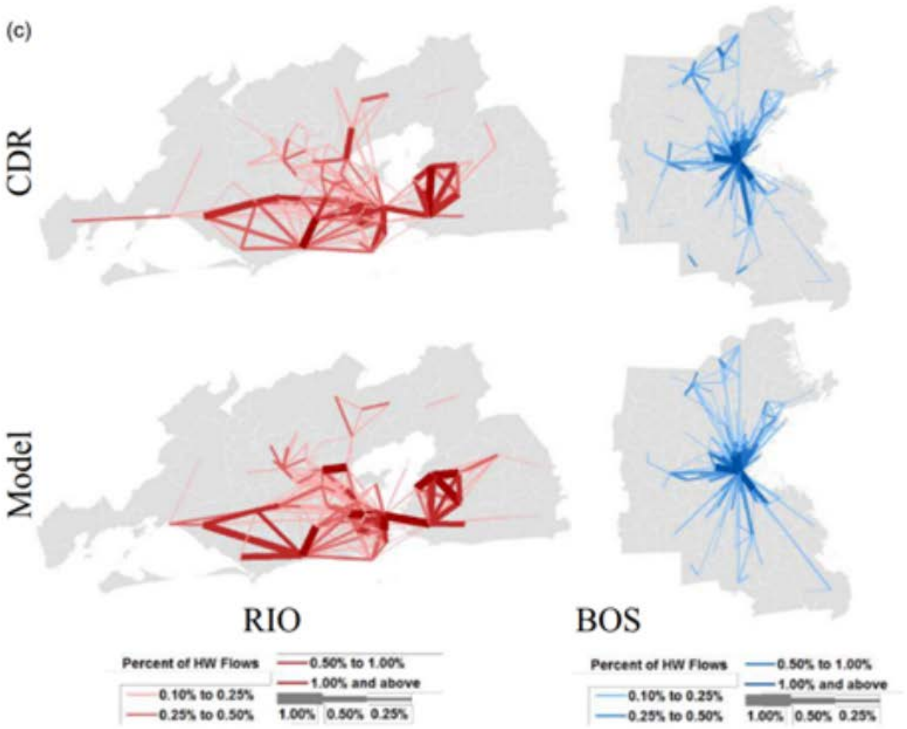
Historical geo-coded social media records are owned by the respective social media companies and historical record can be purchased from these companies, but some is publicly available or can be obtained for free by web scraping. For instance, open source tools such as AIDR utilize Twitter data but are free to access. Finally, smart card data may either be owned by the transportation authority that administers them or the private company that provides the fare payment infrastructure. This data source has useful applications for transit service planning but has not been substantially integrated with regional or statewide transportation demand models.



	Source	Application Developers
Vehicle Probe Data	OEMs own data produced by vehicles such as travel speed, mileage, brake use, etc.	Analytics and Data-as-a-service companies like INRIX and HERE merge various data sources, such as from fixed-sensor networks owned by public entities, probe data from OEMs, and other proprietary sources.
Call Detail Records (CDRs)	Telecommunications Providers own records that are generated every time a call or text message is sent, usually located to a cell tower.	Analytics companies like Teralytics use CDRs and other data sources to generate aggregate origin-destination matrices with travel times.
Smart Card Data	Fare Payment System Developers store all data collected by passengers entering and exiting a transit system in a reporting database.	Many transit agencies own or have access to the raw data generated by their fare payment systems, enabling them to infer passenger origin-destinations, wait times, and trip times.
Geo-coded Social Media Records	Social Media Companies	Artificial Intelligence for Disaster Response
Location-Based Services Data	Location Intelligence/Data Monetization Companies	Location Analytics companies

Chart: https://www.streetlightdata.com/location-based-services-data-transportation-rural-studies

Academics have been the first to demonstrate the powerful insights that can be derived from these emerging passive data sources. For instance, in 2014, MIT researchers demonstrated that data mining call detail records obtained from telecommunications providers can generate origin-destination trips of different purposes (include home-based work, non-home-based, and so on) and time of day. They sourced their data in Boston from AirSage. Each record in their datasets is a call detail record: it consists of an anonymous user ID, the latitude and the longitude, and the time at the instance of the phone activity, including calls and text messages. The coordinates of the records are estimated by the service provider (AirSage) based on a standard triangulation algorithm, accurate to an average of 200 to 300 meters. They validated their results against traditional travel diaries collected in Boston, and with similar results generated in Rio de Janeiro.



A variety of companies have brought products to market that utilize rich cellular or location-based data sources—or a synthesis of both—for transportation planning insight.

Teralytics is one of a few technology companies that have built on the framework for inferring travel flows from CDRs to bring a planning product to market. Their capabilities are featured in more depth in in the Case Study section of this report.

Another powerful data source that has been enabled by the digital revolution is location-based services (LBS) data. LBS data, like CDRs, can be interpreted to understand an individual's geolocation through time. LBS may offer better coverage of the population than data from telecommunications companies do because it is created by all smartphone applications that use LBS, which are typically available on all smartphones across all telecommunications providers. Furthermore, it uses a combination of in-phone GPS and Wi-Fi and Bluetooth sensors so that it is robust to gaps in cellular or Wi-Fi coverage, and thus may offer more geographic coverage than CDR data alone.

Sidewalk Labs utilized LBS data in tandem with a mix of other public and proprietary sources to produce an agent-based, activity-based travel demand model called Replica. Their current and future product offerings and methodology are examined in a case study later within this report.

Bluetooth and Wi-Fi sensing has proven to be a new source for big data in transit applications. Both have been applied to estimate transit riders' origins and destinations, estimate ridership, and evaluate service quality metrics like station wait times with often more accuracy than fare payment system data can provide. Researchers at the University of Washington "developed sensors which cost $60 per instrumented bus that can detect a unique identifier called a Media Access Control (MAC) address associated with a particular mobile device as it boards and leaves the bus to offer complete and real-time travel data. The system only collects MAC addresses and the time and location they are detected from Bluetooth or Wi-Fi signals, and each address is anonymized for privacy protection" [2]. They demonstrated that although the data collected by these sensors is very noisy due to the proximity of pedestrians and cyclists (who should not be counted in ridership estimates), it was possible to accurately sense and identify transit riders. Following academic proofs of concept, some cities have begun testing this technology and method of data collection at a larger scale in their transit networks. For instance, Transport for London conducted a 29-day pilot at the end of 2016 to "collect these device MAC address connections to better understand customer movements through and between stations, by seeing how long it took for a device to travel between stations, the routes the device takes and waiting times at busy periods" [3].

Machine-learning algorithms have been the primary tool for transforming these massive collections of disparate geolocated records into trip origins, destinations, and route choices. These models are also used to infer where an anonymized person may live or work, and thus the different trip purposes behind individual trips. Machine-learning approaches are best suited for these tasks because of the noisy, sparse, and intermittent nature of these passive datasets.

Real-Time System Performance

Big idea: The ability to near-instantaneously process mobility and transportation system data will enable a host of real-time alerts and real-time predictive responses to congestion and incidents. These upgrades are coming at the intersection and corridor-level thanks to computer vision, machine learning, and cloud computing. Benefits will be had system-wide as well: machine learning can diagnose causes of bottlenecks or high crash rates using a blend of real-time and historical data.

Creative approaches to collecting and integrating real-time data will enable the application of machine learning-based predictive analytics in infrastructure operations. Continuously improving machine-learning models will enable public agencies to predict negative externalities such as congestion, traffic incidents, or poor air quality with higher confidence or earlier. The insight could allow public agencies to organize in advance to mitigate the root cause of these externalities, reducing the impact they would have had on the public.

The early pilots in innovative real-time data collection and application have been based in advances in physical sensing and analytics. Networks of internet-connected sensors, often referred to as the Internet of Things (IoT), and computer vision in lieu of physical sensors are the two most common approaches to increasing the breadth and granularity of real-time data. The sheer quantity of raw data that these two systems collect demand cloud-based computing and machine-learning approaches to process them into useful information. Some companies have used novel combinations of many of these data streams in one platform and data mining to observe new patterns and associations in congestion, incidents, and pollution.

The IoT and AI do not depend on each other but are highly complementary. IoT devices generate the rich data that machine-learning algorithms are well-suited to turn into useful knowledge, such as monitoring indicators in the physical environment, detecting and removing anomalous data points, and predicting future outcomes.

In Chicago, Illinois, the Array of Things is an urban sensing or IoT project. It is composed of a network of interactive, modular sensor boxes that will be installed around Chicago to collect real-time data on the city's environment, infrastructure, and activity for research and public use. The project is the result of a joint initiative between Argonne National Laboratory and the University of Chicago, with funding from the National Science Foundation. The deployed sensor nodes can measure temperature, barometric pressure, light, vibration, carbon monoxide, nitrogen dioxide, sulfur dioxide, ozone, ambient sound intensity, and surface temperature. The team is also working to monitor other urban factors of interest, such as flooding and standing water, precipitation, wind, and pedestrian and vehicle counts, at intervals of several minutes, creating a measure of pedestrian and vehicle flows over time. The project's physical vibration sensors and magnetic field sensors will be used to detect heavy vehicle flow, and cameras will be used to detect traffic flow, including pedestrians and cyclists. In May 2018, the project team installed their 100th sensor, and continue to expand their network and the types of data their sensors will collect. The project team is currently still building out to a full network of 500 sensor-equipped sites and exploring how historical data collected can be analyzed. The long-term vision for the project is geared towards real-time alerts: they are looking to see how their system can provide notice of dangerous weather conditions, safe walking routes, urban flooding detection, and air quality alerts to citizens, which would likely require the use of cloud or edge computing and machine learning to rapidly process signal data [4].

Cameras are often a component in a larger IoT network. However, because computer vision is a relatively mature sub-field of AI, it has emerged as one of the most viable approaches to monitoring traffic without in-ground or other physical sensors. A number of traffic management solutions have come to market that rely solely on cameras to sense the traffic environment. Processing the massive quantity of video data that is produced by one or more cameras has become computationally possible only recently, often using cloud computing. Computer vision applications can be performed using traditional video data or infrared video data, and their technical underpinnings are described in the Technical Primer section of this report.

Miovision and Flir are two companies taking an innovative approach to traffic management using computer vision.

Miovision's traffic management solution, TrafficLink, uses a 360-degree camera to monitor traffic flow at an intersection. They have implemented algorithms that can detect and classify vehicles, pedestrians, and cyclists, enabling a variety of dynamic traffic signal control opportunities and new intersection performance measures. Miovision's technology is also featured in more detail in a case study in this report.

Flir was initially focused on offering high-performance, low-cost infrared (thermal) imaging systems for airborne applications. As such, they are best known for their infrared or thermal imaging cameras. However, they apply their camera technology to a host of markets and sectors, ranging from defense, industrial, and public safety and transportation, to security. In the transportation sector, they offer traffic management solutions based on video analytics. They use a combination of traditional and infrared cameras to monitor an intersection. The infrared cameras are particularly useful for identifying pedestrians and cyclists because patterns generated by humans' body heat can be isolated and identified. The applications they tout are therefore most focused towards V2X pedestrian and cyclist safety applications, ranging from dynamic pedestrian countdown, pedestrian count data collection, and automatic pedestrian green phases [5].

Computer vision can also serve as an unobtrusive solution for monitoring parking space availability, including truck parking at public rest areas and commercial truck stops. Computer science researchers at the University of Minnesota (UMN) recently demonstrated that this technology can detect real-time truck parking availability at above 95% accuracy. Their system uses multiple cameras to construct a three-dimensional (3D) representation of the parking lot, which is more reliable at filtering out false signals than a single camera. They deployed their system at three public rest stops along I-94 in Minnesota, a major freight corridor. The UMN team also disseminated real-time parking availability in real time through three mediums: a commercial operator accessible web parking information portal, an in-cab geolocation application that integrated within an existing onboard logistics device to support driver and carrier trip operations, and roadside electronic message signs. This camera-based system is continuing to be tested and deployed by the same research team in order to develop solutions to provide 24/7 parking information without disturbing the existing pavement structures or substructures [6].

IoT implementations and new approaches to intersection monitoring provide the opportunity to utilize new performance indicators at the intersection or corridor level. However, real-time data will also improve system-wide traffic management. Traffic management center (TMC) operations will be enhanced by the new integration of existing and emerging real-time data sources. Several platforms entering the market offer improved traffic congestion and incident prediction by machine learning using data sources such as navigation applications, traffic signal controllers, and weather monitoring.

Waycare is an Israel-based startup that uses proprietary deep-learning algorithms and diverse data sources to understand the causes of traffic congestion and incidents. They partner with more established data sources for traffic incidents such as Waze. Waycare also taps into non-traditional sources like TicketMaster to forecast the number of trips generated by sports events and concerts and estimate the resulting traffic impacts. Some of the use cases that Waycare anticipates include the dynamic re-timing of traffic signals in response to anticipated congestion, opening and closing roads, updating dynamic message signs, and coordination with public safety agencies for faster incident response. Waycare's products will eventually offer integrated and cross-cutting solutions in historical data collection and analysis, real-time traffic management, and public safety and enforcement [7].

INRIX is well-established in the transportation analytics market. Their variety of traffic data and analysis offerings are driven by their use of over 100 separate data sources, including anonymous, real-time GPS data from millions of connected vehicles and devices. Their data is ubiquitous, even collaborating with Waze to inform Waze's traffic estimations. Like Waycare, INRIX is analyzing real-time traffic data to discover new ways to anticipate congestion and incidents. For instance, a feature that INRIX recently launched within their traffic management platform detects unanticipated or unusual traffic slowdowns based on real-time vehicle data. Their algorithm can infer where traffic incidents have occurred by detecting an extreme change in travel speed between two adjacent segments of roadway. INRIX then uses this insight to provide location-based alerts to active drivers in the area and the TMC. Both the Iowa Department of Transportation and the Ohio Department of Transportation are early customers of this new feature. The use cases for these real-time traffic data platforms in incident management and public safety are also detailed in the Public Safety and Enforcement application area.

Public Safety and Enforcement

Big idea: Advances in AI and real-time data applications promise to reduce response times for almost all first responders by alerting authorities of any issues on the ground as they happen. Many of the benefits will be the result of increasingly integrated traffic data collection and cross-sector coordination.

Real-time data offers many applications for public safety and enforcement. All the cameras and sensors that IoT applications provide can be used to monitor public spaces, identify potential conflicts, and react to incidents faster and more efficiently than before. Computer vision can visually recognize dangers, while audio sensors use natural language processing (NLP) to identify threats—or calls for help—through sound.

Traffic light camera data, from a company like Miovision, can identify potential zones of conflict—areas where cars are more likely to hit cyclists or pedestrians, due to either poor intersection construction or non-compliance like jaywalking. Miovision's intersection monitoring system can also immediately identify a crash at an intersection and send an alert to the relevant public agencies. It can even analyze the traffic patterns that occur around previous crashes and use that data to predict times and areas where future incidents are likely. With that knowledge, the TMC can coordinate with police departments or emergency responders to strategically position enforcement and responders in these areas, to anticipate or mitigate incidents.

Likewise, system-wide real-time data integrators like Waycare can decrease response time to accidents on highways and other roadways. Also like Miovision, Waycare can predict times and areas where potential accidents can occur, and station the appropriate personnel near those areas [8]. As a result, the Nevada Highway Patrol is a prominent partner in a pilot project with Waycare in Las Vegas, along with the Regional Transportation Commission of Southern Nevada and the Nevada Center for Advanced Mobility.

Each of the above applications could form a part of a “smart” city—meaning that they provide real-time data about a city via IoT technology. These technologies will eventually be able to work in concert with one another. Waycare’s sensors can inform Miovision’s traffic lights about incoming traffic from a highway, allowing the traffic lights to adjust accordingly. The system could work the other way around as well, with the traffic lights letting highway patrol or an app like Waze know where to expect the most traffic congestion around a city.

Many other players are entering the smart-city niche, and many of these companies are combining many features and applications into a single technology. City IQ is a consortium including AT&T, Intel, and Current by GE. They have deployed 200 smart streetlights on the three roads with highest crash rates in Portland, Oregon, under the Traffic Sensor Safety Project. The system combines cameras, microphones, and environmental sensors, which can identify pedestrians, cars, and cyclists; help pinpoint the location of gunshots; and sniff for pollutants (the implementation is primarily in support of Portland’s Vision Zero plan to reduce traffic fatalities). This same consortium has also begun deploying over 3,000 sensors, plus 14,000 LED lights, in San Diego, California; some of the initial applications in this deployment will include real-time parking information and the ShotSpotter technology, which detects and locates gunshots [9].

Gunshot detection works in the same way a smart phone recognizes a voice; the underlying algorithm is trained on bits of audio labeled as gunshot or non-gunshot. The problem itself is simpler than speech recognition, even if real-world data on gunshot audio tends to be nosier (the street light is often farther away from a gunshot source than a phone is from a speech source). ShotSpotter, a fully operational gunshot detector, uses machine learning to not only detect gunshots, but also triangulate their location. As with the other technologies mentioned, the advantage gained is reduced response time for police, EMTs, and other relevant parties.

Other AI applications reduce response time as well. Some proof-of-concept work has involved combing through social media posts to identify calls for help during disasters. The Artificial Intelligence for Disaster Response (AIDR) platform has created several NLP implementations that can, among other things, filter out tweets that ask for assistance during a disaster, and then show the relevant tweets to the relevant response teams [10].

Construction and Asset Management

Big idea: AI can inform decision-making on a construction project site or by maintenance planners. The outcome will be improved efficiency and safety for construction workers and maintenance crews. Platforms that integrate various forms of data enable much of this new insight.

The oversight of all stages of a transportation project lifecycle is becoming data-driven to improve safety and find efficiencies, not just system planning and operations. Work zones and maintenance operations and planning are becoming increasingly digitized; with this comes new types of data collection and analytics.

Construction management software platforms consolidate a variety of project metadata, such as drawings, markups, issues, checklists, RFIs, submittals, clashes, and project and business profile project metadata. Even the use of drones or the ubiquity of smartphone digital cameras at work zones is adding to the explosion of data collected at a job site. Machine learning and AI can be used on this wealth of data to manage construction site risk, improve safety, and eventually optimize the scheduling of project tasks.

Autodesk’s project delivery and construction management software BIM360 provides solutions along multiple phases and aspects of the project delivery process, from design, contracting, and construction, to project management. Some of the functionalities of this software include using the cloud-based platform to share design and other files to streamline stakeholder coordination, building checklists and tracking issues, and tracking deliverables. Autodesk is extending the functionality of this platform with AI-based deep-learning techniques with the Project IQ suite of applications. Project IQ will be built into BIM 360 and will use project data to analyze past and current projects for safety and efficiency and provide targeted warning about delays and threats to workers’ safety. Using the data already stored in the BIM 360 platform, Project IQ will automatically scan all safety issues reported on a job site and attach a tag to them indicating whether they could lead to potential fatalities. The algorithm will identify linkages between hazards, dangerous behaviors, and job site injuries or fatalities to inform job site managers as to where they should target safety training efforts. Project IQ is still in the pilot phase and is being tested among current users of the BIM360 platform who have already collected multiple projects' worth of data.

Smartvid.io, a technology startup, is also applying AI to work zone safety. They extend the integration of work zone data through project management platforms such as Autodesk's BIM 360 by identifying unsafe construction behaviors. They apply computer vision to digital photographs taken on job sites in order to classify different worker behaviors captured in an image, and tag unsafe behaviors. As files from digital cameras, GoPros, and cell phone cameras are uploaded to the Smartvid.io media management platform, they are automatically tagged by their visual and audio content based on what the AI engine sees and hears in the content. Their technology is capable of identifying hazardous conditions based on similarities with previously identified hazards, such as misplaced hole coverings, improperly used ladders, and incorrectly installed barriers. It can also identify individual people in photos and videos, and analyze the presence or absence of safety protocol, including whether they are wearing appropriate PPE. Their engine processed 1,080 images in less than 10 minutes (while a human team required over 4.5 hours) and with greater accuracy than the human team. The engine also flagged 32 images containing personnel missing hard hats, and 106 images with workers missing safety-colored clothing. Smartvid.io's technology will make it possible for human supervisors to review flagged safety hazards and then target work zone safety education to those project- and staff-specific issues.

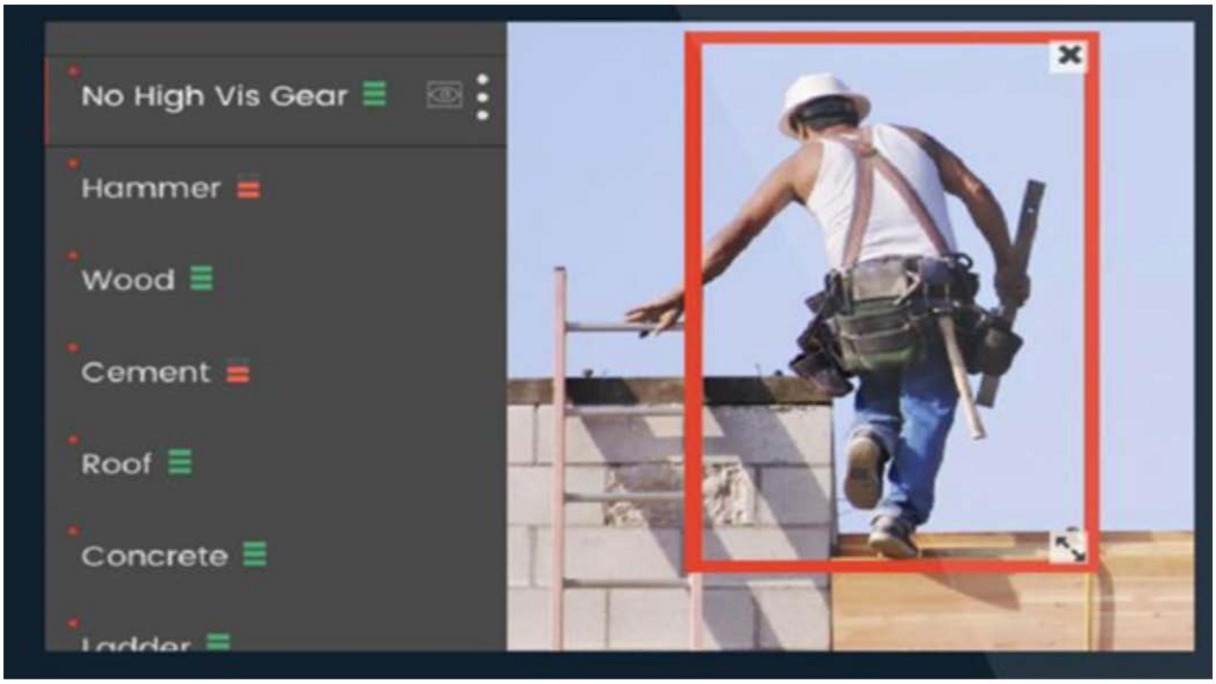


Figure: Smartvid.io (https://medium.com/autodesk-university/the-rise-of-ai-and-machine-learning-in-construction-219f95342f5c)

Maintenance and operations functions are also evolving with the introduction of technologies that use AI, automation, and analytics. Currently in development are automated machines that assess pavement conditions faster and with greater accuracy and coverage than human teams of surveyors can. Analysts envision using this newly comprehensive and more frequent data to enable predictive maintenance, optimizing preventive maintenance efforts on highways and transit systems.

Automated systems for pavement condition detection have been the subject of academic research since the 1990s. The most mature technology in this space uses scanning lasers to construct and analyze a 3D profile of a roadway. Pavemetrics' Laser Crack Measurement System (LCMS) has been implemented by many agencies worldwide. Outfitted on a moving van, LCMS records 5,600 4-meter-wide transverse profiles per second and combines them together to create a very high resolution 3D profile of the road. Then, the system uses algorithms to process the road profile and identify potential road defects. Typical outputs of the LCMS include crack detection, rut detection, macro-texture evaluation, ravelling evaluation, pothole detection, and more. The Texas Department of Transportation conducted a study with the Texas A&M Transportation Institute in 2016 to evaluate two automated visual distress data collection vendors compared to human-generated or manual Pavement Management Information System ratings in the Austin, Bryan, and Waco districts. The research team found mixed success using the automated data collection methods. The automated results were reasonably comparable to manual ones for asphalt distress surveys but were inconsistent across different distress types in jointed concrete pavement. Additionally, one vendor had more accurate rut depth measurements than the other based on the reference measurements obtained by the research team.

As these laser-based technologies continue to mature, their usefulness will increase and their implementation costs will inevitably lower. However, other solutions are entering the market that propose pavement condition monitoring methods at lower resolution but also substantially lower up-front cost. One such vendor is RoadBotics, which outfits fleet vehicles with smartphone cameras and processes video frames to identify various road defects; because their implementation costs far less than traditional laser monitoring systems, an agency can automatically gather pavement condition data with higher frequency and on more of their road network. RoadBotics is also the subject of a case study in this report.

The pavement data that will one day be collected by automated systems is integrated into state pavement management systems. Pavement management is becoming increasingly data-driven, generating outcomes such as assessing performance trends, calibrating design models, evaluating the cost-effectiveness of different treatment strategies, and recommending candidate projects for a preservation program. However, pavement preservation programs still face barriers to integrating preventive maintenance activities into their systems, primarily due to lack of useful data at a network level. As highway programs in the United States continue working towards properly implementing preventive maintenance, leading transit agencies or service providers have begun advancing from preventive to predictive maintenance models.

Predictive maintenance is made possible by predictive models that can forecast damage and deterioration in detail over long time frames. By being able to accurately predict infrastructure conditions over time, maintenance teams can intelligently coordinate maintenance activities to minimize the total cost of operations. This is accomplished by optimizing the scheduling of such activities so that no maintenance activities are repeated more often than necessary and so that preventive maintenance or replacement occurs far before catastrophic failure, both of which save agencies time and money. Machine-learning algorithms in forecasting and optimization can take the increasing quantity of collected data (such as from automated pavement condition monitoring systems) into consideration in intelligent maintenance planning. Academic researchers have begun developing neural network and kernel methods to provide more accurate forecasting results for infrastructure condition models, so predictive maintenance enabled by AI has become conceptually feasible for highway pavement programs.

On the other hand, predictive maintenance is not a new idea in other spheres. IBM's asset management and maintenance business line, called Maximo, has many established working partnerships with private and public customers. Furthermore, the private sector in particular has been familiar with the idea of "smart" asset management since before the turn of the century. Although it was first applied in industrial settings, public agencies like transit operators and water and energy utilities have also begun exploring predictive maintenance schemes. Looking ahead, there are ways in which AI and deep learning are poised to improve the efficacy of predictive maintenance, and they are discussed in an IBM Maximo case study. The case study will also highlight some of that team's work with Yarra Trams, a tram system operator in Melbourne, Australia.

Public Administration/Information Management

Big idea: Machine learning and AI could automate various administrative or time-consuming tasks. Early examples lean on NLP to support human resources processes, and it is expected that increasingly complex tasks will become subject to attempts at automation.

Given a cloud computing data management architecture, AI has the power to streamline and even automate certain administrative tasks. Many cloud computing solutions also offer repositories for the unstructured data streaming in from IoT devices. These database solutions often have built-in machine-learning implementations that can automatically organize data, draft documents, and suggest different types of analysis.

As IoT devices become more widely implemented, and other data sources and collection methods are digitized, many organizational tasks can be put under the purview of machine learning. Data collection and organization of road survey data can be left almost completely to machines: as the data from the IoT devices enters the cloud system, machine-learning algorithms will organize and classify the data. The algorithms can even suggest useful analyses, or helpful ways to reorganize the data.

These helpful suggestions extend to the realm of human capital management (HCM). Oracle's upcoming adaptive intelligence HCM product can suggest the best candidates for a position by automatically reading digitized resumes using NLP techniques, thus reducing the number of employees needed to process new hires. UtiliPro, a different HCM tool, can even read and summarize open ended questions from consumer or employee surveys. Other services use NLP to draft documents. Companies like Thought River and Luminance can write legal documents, or other pieces of writing with rigid structures.

It should be noted that cloud computing technologies all offer the same or similar services. Informatica, Oracle, Amazon Web Services, Microsoft Azure, etc., are all working on dynamic databases that self-organize, and offer suggestions about how to manage the data. Most of the machine-learning augmented services are currently in their beta phase, but most companies predict offering these services within 1 to 2 years.

____

[1] From phone conversation
[2] http://www.washington.edu/news/2016/01/20/bluetooth-and-wi-fi-sensing-from-mobile-devices-may-help-improve-bus-service/
[3] https://tfl.gov.uk/corporate/privacy-and-cookies/wifi-data-collection-pilot
[4] https://arrayofthings.github.io/
[5] https://www.flir.com/products/trafione/
[6] https://conservancy.umn.edu/handle/11299/185538
[7] https://www.haaretz.com/israel-news/business/waycare-an-israeli-startup-takes-charge-of-las-vegas-roads-1.5790998
[8] http://www.govtech.com/Las-Vegas-to-Pilot-WayCares-Accident-Prediction-Artificial-Intelligence-Software.html
[9] http://fortune.com/2018/06/18/portland-sensors-smart-cities-traffic-death-att-intel-ge/
[10] http://aidr.qcri.org/

CASE STUDIES

OVERVIEW

To complement the breath of the Menu of Applications chapter, this paper provides in depth case studies into a select group of companies. Few of the selected companies solve the same problems, and all use different methods and technologies from each other. Use to find concrete examples of transportation specific AI implementations.

IBM Maximo

Case Study

COMPANY OVERVIEW

The Maximo arm of IBM has traditionally been focused on maintenance and asset management. Maximo's original scope of work entailed ensuring that infrastructure systems remain operational in the short term by identifying parts and equipment that are most likely the fail in the next few days and then scheduling preventative maintenance or replacement. However, in their work with various public agencies and private companies, and with the proliferation of real-time data sources, the Maximo team has observed that their products have become useful for operations, maintenance, and longer-term capital planning and personnel alike.

CURRENT CAPBILITIES

Currently, IBM Maximo maps sensor outputs to the piece of equipment that sensor is installed on, such as a track rail or a train wheel. They consolidate these mapped sensor data in a platform, which allows a platform user to manage and observe the asset health of every outfitted part.

With the help of subject matter experts (such as a maintenance manager who can identify on sight a part that is about to fail) and data scientists, the Maximo team guides the development of a predictive maintenance framework customized to their customers' needs. They work with customers to first understand what their critical assets are, how they evaluate their health, and how they maintain them. Then, they work together to devise a way to measure an asset's health quantitatively, and determine how to use sensors and analytics to score those assets. This last step requires that they derive the patterns in sensor outputs that are associated with parts and equipment requiring attention. For instance, a certain part may display a unique vibration pattern or an extreme temperature when it is approaching the end of its lifecycle, relative to how it behaves when it is in good working condition. Then, the next time such a pattern is detected, Maximo can raise an alert to maintenance personnel.

IBM Maximo

Because these predictive maintenance implementations usually encompass an entire infrastructure system, once each unique pattern has been associated with a part and encoded into the platform, it can prioritize the order in which maintenance crews should schedule in-person inspections and decide whether to repair or replace a part.

FUTURE CAPBILITIES

For Maximo's future, the focus is on improving the predictive powers of the system, and even working the implementations into a real-time data analytics role. For example, IBM will be able to project the lifetime of different pieces of equipment—and adjust that lifetime as unforeseen bouts of wear and tear occur. IBM can then coordinate that lifetime with all the other lifetimes of the objects in the system, and generate a projected maintenance schedule over time. That schedule will be optimized such that the fewest equipment failures occur using the fewest parts with the least amount of labor over time. This will evoke a shift from shorter-term insight and maintenance action on the order of days to a monthly or yearly asset management perspective.

Another capability IBM Maximo is aiming towards is being able to not only identify the presence of a general maintenance need, but diagnose it. From there they can make longer-term recommendations on the next best step to take. For instance, they could determine whether it would be more efficient to either repair or replace a part.

HOW IT WORKS

Detection and tracking are made possible by computer vision algorithms. These algorithms are fed with over a decade's worth of human-labeled intersection data, and with human-labeled data from Miovision's current smart intersections. Miovision processes this data, and trains their algorithms, on the Amazon Web Services cloud computing platform.

While detection is done with a traditional computer vision pipeline outlined in the technical primer, tracking involves more human intervention upon hardware installation. Once Miovision's camera/computer outfit is placed at an intersection, a human operator views video feed of that intersection on a normal day, and highlights areas of interest. These areas of interest contain the common paths of cars, cyclists, and pedestrians. The below image illustrates what these areas of interest look like for cars.

USE CASE

Yarra Trams, the tram service operator in Melbourne, Australia, is a prominent example of the IBM Maximo asset management solution at work. Yarra Trams outfitted its tram network with 91,000 data sensor points on separate pieces of tram equipment.

These sensors range from automated wheel-measuring machines to track sensors that detect signs of track wear or breakage. The IoT approach to inform asset management has enabled Yarra Trams to implement predictive maintenance scheduling [11].

[11] https://www.igi-global.com/chapter/application-of-artificial-neural-networks-in-predicting-the-degradation-of-tram-tracks-using-maintenance-data/167562

ADVICE TO TXDOT

- IBM echoed other case study subjects' emphasis on the importance of starting any machine learning or predictive analytics application with high quality data. They highlighted the critical role that the data scientist will play in deriving as much insight as possible from large datasets. They also observed that public agencies can tend to be siloed across different departments, which can pose a challenge to addressing cross-functional needs efficiently with a single product or data source. Recognizing such existing organizational barriers is the first step to overcoming them and identifying shared opportunities.

Miovision

Case Study

COMPANY OVERVIEW

Miovision is a traffic solution company founded in 2005 and headquartered in Canada. The company collects data on multiple modes (using their proprietary traffic signal hardware), analyzes that data in near-real time and real time, and integrates information into a collection of smart intersections. Deep-dive interviews with the Miovision team on computer vision-based hardware for real-time traffic operations, intersection safety analysis, and data collection as well as their data platforms were conducted.

CURRENT CAPBILITIES

The Miovision TrafficLink solution suite combines a camera, and a compact, edge computer to process video data, and traffic signal cabinet hardware with a real-time portal for traffic management and data collection. Presently, the TrafficLink system can detect vehicles, pedestrians, and cyclists in an outfitted intersection. This enables the system to implement real-time signal extensions that can ensure that cyclists and pedestrians have adequate time to exit the intersection. Miovision's approach to the detection of pedestrians is novel because it doesn't require that agencies carry a beacon across the intersection in order to be detected in the way that previous V2X approaches to pedestrian intersection safety applications have proposed. TrafficLink can also detect vehicles in almost all weather conditions, further enabling real-time signal actuation; for instance, once stopped or approaching vehicles are detected, the system can make a call to the signal's server to request to hold or shorten a certain phase. Furthermore, being able to detect all agents that enter and exit the intersection provides them the ability to generate traffic counts on roads in all directions of the intersection.

Miovision

FUTURE CAPBILITIES

The next technical challenge that Miovision faces concerns the tracking of vehicles, cyclists, and pedestrians through an intersection. Differentiating between pedestrians, cyclists, and vehicles is particularly exciting, because computer vision software has, up until now, been unable to make these distinctions. This will enable Miovision to provide turning movement counts through an intersection, which has been particularly challenging for traffic engineers to collect beyond traditional manual counting efforts. Miovision promises to count how many vehicles turned right, left, went straight, etc. The ability to track vehicles and people through an intersection opens up a host of novel, data-driven intersection safety analyses. For instance, Miovision will be able to identify and classify “near-miss crashes” in the zones of an intersection. This has not been measured empirically before. With this information Miovision can bolster currently sparse crash data with more potential conflicts, and inform the selection and implementation of safety countermeasures. These new insights could potentially reduce the risk for cyclists and pedestrians.

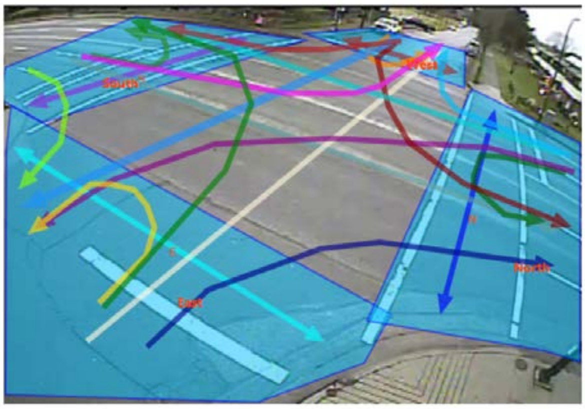
Furthermore, because the company is developing the capability for multimodal detection and tracking, these analyses can be tailored to the most vulnerable road users such as cyclists and pedestrians.

HOW IT WORKS

Detection and tracking are made possible by computer vision algorithms. These algorithms are fed with over a decade’s worth of human-labeled intersection data, and with human-labeled data from Miovision’s current smart intersections. Miovision processes this data, and trains their algorithms, on the Amazon Web Services cloud computing platform.

While detection is done with a traditional computer vision pipeline outlined in the technical primer, tracking involves more human intervention upon hardware installation. Once Miovision’s camera/computer outfit is placed at an intersection, a human operator views video feed of that intersection on a normal day, and highlights areas of interest. These areas of interest contain the common paths of cars, cyclists, and pedestrians. The below image illustrates what these areas of interest look like for cars.

Miovision



Given this information, Miovision's beta testing has proven to be adept at tracking which cars turn right, which go through an intersection, and which make left or U-turns, and other movements.

Gaps in their computer vision algorithms' capabilities exist for less common weather conditions, such as hurricanes, but Miovision's technology will improve with the collection of more data.

USE CASE

Miovision and the City of Detroit worked together to launch "The World's Smartest Intersection," where Miovision is currently piloting a version of their tracking software. This partnership provided marked advancements in the connectivity of Detroit's traffic infrastructure, while allowing Miovision to test and refine their upcoming capabilities in vehicle tracking. Together, five intersections in a corridor were upgraded with Miovision's complete TrafficLink platform. A notable aspect of this pilot was that because the City of Detroit had previously instrumented those intersections with Miovision's hardware, Miovision was able to remotely update the software to contain the new packages capable of collecting more intersection data and actuating real-time signal phase adjustments.

ADVICE TO TXDOT

- Prioritize open architecture data management so that the structure of the platform and databases can be easily adapted to meet evolving needs within a changing digital landscape. This will enable a public agency to integrate many data streams from a number of third-party providers.
- Consider how technology asset investments can be synergistic across multiple agencies, such as with operators in other sectors such as emergency response and public safety.
- Choose technology solutions that put the ownership of the platforms and the raw data generated in the hands of the public agency.

RoadBotics

Case Study

COMPANY OVERVIEW

RoadBotics originated from research conducted at Carnegie Mellon University's Robotics Institute. Although the company's use of computer vision to classify pavement defects began as a purely academic pursuit, the research team quickly realized the broad usefulness of its technology, and formed the company RoadBotics in late 2016. Since then, its technology has been deployed in 56 cities and counties in the US and Canada, primarily with public works departments.

CURRENT CAPBILITIES

RoadBotics uses cameras, currently from smartphones, to scan pavement from a moving vehicle, then feeds data into computer vision algorithms to identify various defects and provide a pavement condition map. Currently, their platform serves as an intermediary between manual pavement inspections (which are labor- and time-intensive, subjective, and dangerous to conduct from the roadside) and automated laser-based scanning systems (which are incredibly precise but generally too costly to use on an entire network). RoadBotics can diagnose road defects more objectively than a human inspector can, and also collect data passively at a low cost by outfitting fleet vehicles from industry partners of public agencies. The level of precision that RoadBotics can achieve using a smartphone camera is sufficient to inform maintenance decisions that can mitigate costly defects like cracking, rutting, and potholes before a road is irreparable. Therefore, their technology complements the existing but much more expensive technologies such as laser-based scanning and 3-D profiling by informing where such detailed monitoring may be most critical.

RoadBotics

FUTURE CAPBILITIES

Describing themselves as a software company, RoadBotics is currently working on identifying other infrastructure visible from the road, such as guardrails and signage. Another next push will be identifying damage in power lines from their existing video feed, and potentially sharing that information with the relevant utilities or operators. Additionally, they are looking into upgrading their recording technology, in order to improve their data and classification quality.

Beyond upgrades and diversification, RoadBotics is currently attracting interest from various automated vehicle technology companies and automobile manufacturers, some of which have already invested in RoadBotics, and Tier 1 automotive suppliers. RoadBotics anticipates the profound value that could be generated by the integration of their pavement monitoring software with automated vehicle providers, especially as more miles on roads are driven by fleets or transportation network companies.

HOW IT WORKS

RoadBotics relies on several years of research, done by one of its founding members, to identify different types of road damage via computer vision. Expert staff has trained a series of neural networks to identify four dozen features (different types of cracks, etc.) on labeled video of roads. Their recording equipment of choice is still currently low tech: a modified smart phone. Despite the relatively primitive data collection device, RoadBotics' algorithms provide an accurate and consistent classification of roads into five groupings. Combined with the GPS information collected by the modified phone, RoadBotics produces a heat map of a city's roads, highlighting areas that warrant more attention, and distinguishing between roads in good condition and roads with a moderate degree of wear.

This lowbudget approach means RoadBotics can quickly outfit a fleet of vehicles with their modified smart phones, and thus map out an entire city's roads in only a few days. The company does have plans to upgrade their hardware, but it's important to keep in mind that their primary wares are the software and the data it generates. The road-condition-identifying software can be applied to any video recording device; RoadBotics is currently leasing their technology to several autonomous vehicle manufacturers.

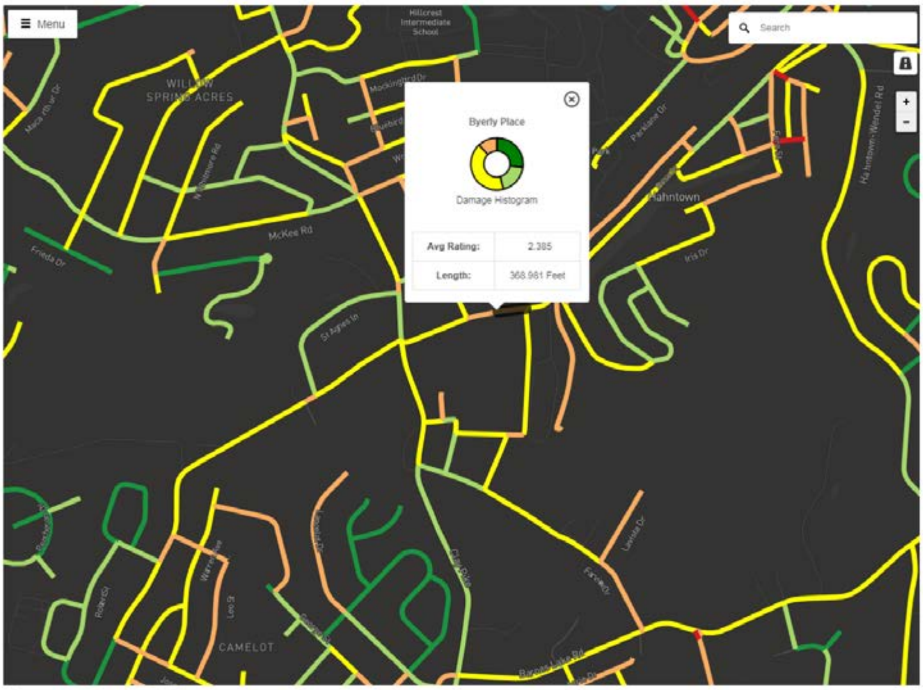
The product is almost entirely derived from computer vision and statistical analysis. The underlying math does not differ greatly from that outlined in the computer vision section of the technical primer; the neural networks and analysis are powered by cloud computing, and the video data is stored in the cloud as well.

RoadBotics

USE CASE

In South Bend, IN, RoadBotics conducted video of 100 miles of roads in partnership with the public works department. Their detection algorithms are calibrated to the Pavement Surface Evaluation and Rating (PASER) standard, which the South Bend Public Works Department uses. [1]

Below is a RoadBotics assessment map of a roadway section near Pittsburgh, PA, as an example of their roadway scans' output.



Avg Rating:	2.385
Length:	368.981 Feet

Source: RoadBotics

[1] https://www.abc57.com/news/south-bend-tests-new-technology-to-assess-road-conditions

ADVICE TO TXDOT

- Digitize data collection (such as across highway maintenance programs) so that data can be processed easily and that information can be used for many different applications.
- To enable data-driven decision making, start by ensuring that data collection processes are not only digitized but well-designed. Collecting good data through automated, planned, or routine processes consistently through time will allow analysts to derive far more insights from data.

Oracle

Case Study

COMPANY OVERVIEW

Compared to many modern AI achievements, database technology and theory has been around for a long time. Likewise, Oracle, a database technology company, is more mature and established than most of the other companies featured in this whitepaper. Advances in machine learning have changed and enhanced database systems, but not nearly as much as advances in database systems—i.e., cloud technology and the continued exponential increase in computing power—have enhanced machine learning. Thus, Oracle offers state-of-the-art data management systems that can self-organize data and make recommendations based on machine learning, but the bulk of their operations focuses on data management software and cloud computing services.

CURRENT CAPBILITIES

Oracle offers data management services (such as software tools, system integration, and training for those tools), cloud computing, and a limited number of smart databases. For maximum data security, a company like Oracle can offer consulting on how to upgrade hardware and software to handle large data streams and complex computing problems. Oracle, like all cloud service providers, can offer via cloud computing all the same state-of-the-art capabilities without requiring a client to upgrade any of their existing hardware. Either way, the extra storage space can house the data from IoT implementations, such as 24-hour traffic cameras, and power computationally expensive machine-learning operations, like processing 24-hour traffic camera footage into useful information, such as traffic counts or travel speeds.

FUTURE CAPBILITIES

In addition to general power and storage increases, future databases look to incorporate an AI component to provide users with virtual assistance. Many of Oracle's upcoming smart databases offer recommendations to the user about how to structure and use the data. In some use cases, the data will be able to self-structure, or self-organize, for easier future querying.

Oracle

HOW IT WORKS

Smart databases look at the current structure of the database, remember what each database manager did with past data, and then guess what to do with new data. The guesses are then communicated to the database manager as suggestions or recommendations.

Some smart databases can work with text data, as well as the more numerical data from sensors or cameras. For instance, a smart database can read resumes, and then, based on criteria outlined by the human resources department, identify the best potential candidates.

USE CASE

Marseille FR use case: “Safe” City using Oracle Big Data platform
https://www.forbes.com/sites/oracle/2017/12/12/marseille-turns-to-data-to-plan-a-safer-city/#7f6230211095

The Oracle Consulting solution for Marseille’s plan to radically improve public safety by using big data analytics and machine learning is their Big Data Appliance platform, which is geared towards employing social intelligence for public safety. The machine-learning algorithms will analyze disparate data sources—including data from sound sensors, social media streams, weather patterns, and automobile and pedestrian traffic flow—to predict potential instances of civil unrest and prevent terrorist attacks. The platform encompasses a data acquisition engine from social media platforms, a data pool that integrates variety of data, a semantics analytics toolkit to process social data, a “discovery lab” or analytical environment that aids data scientists in discovering patterns and relationships within data, and a reporting and visualizing tool for analysts to convey information to public safety decision-makers such as police officers and dispatch staff.

ADVICE TO TXDOT

- Oracle offered some focusing questions as their advice to TxDoT:
- What is the problem(s) to be solved or priority use cases and applications?
- What are current internal data science capabilities?
- What data science tools area available to helps solve the problems?
- What resources do those tools require? These can include IoT hardware, cloud computing, a slew of data scientists, etc.

Teralytics

Case Study

COMPANY OVERVIEW

Teralytics is a transportation analytics company that uses telecommunications companies' data to infer origin-destination matrices, travel times, and other useful planning metrics. It is an international start-up, founded by researchers at ETH Zurich in Switzerland based on a research collaboration with a Swiss telecommunications company. Since then Teralytics has expanded to work with various agencies in Germany, Hong Kong, and the United States, developing partnerships with more telecommunications companies along the way.

CURRENT CAPBILITIES

Oracle offers data management services (such as software tools, system integration, and training for those tools), cloud computing, and a limited number of smart databases. For maximum data security, a company like Oracle can offer consulting on how to upgrade hardware and software to handle large data streams and complex computing problems. Oracle, like all cloud service providers, can offer via cloud computing all the same state-of-the-art capabilities without requiring a client to upgrade any of their existing hardware. Either way, the extra storage space can house the data from IoT implementations, such as 24-hour traffic cameras, and power computationally expensive machine-learning operations, like processing 24-hour traffic camera footage into useful information, such as traffic counts or travel speeds.

FUTURE CAPBILITIES

Although Teralytics has demonstrated that call detail records in themselves provide a great deal of potential insight into mobility patterns, they are looking to integrate more data sources into their platform. One of those potential data sources is Wi-Fi network data which would provide even better coverage of the traveling public.

Teralytics

Although Teralytics has demonstrated that call detail records in themselves provide a great deal of potential insight into mobility patterns, they are looking to integrate more data sources into their platform. One of those potential data sources is Wi-Fi network data which would provide even better coverage of the traveling public.

The Teralytics data team is flexible with designing a data visualization or compiling metrics that addressing the specific needs of an agency. Because they have telecommunications data coverage over the entire United States, they are able to quickly put together metrics at appropriate granularity and scope to answer novel questions posed by agencies.

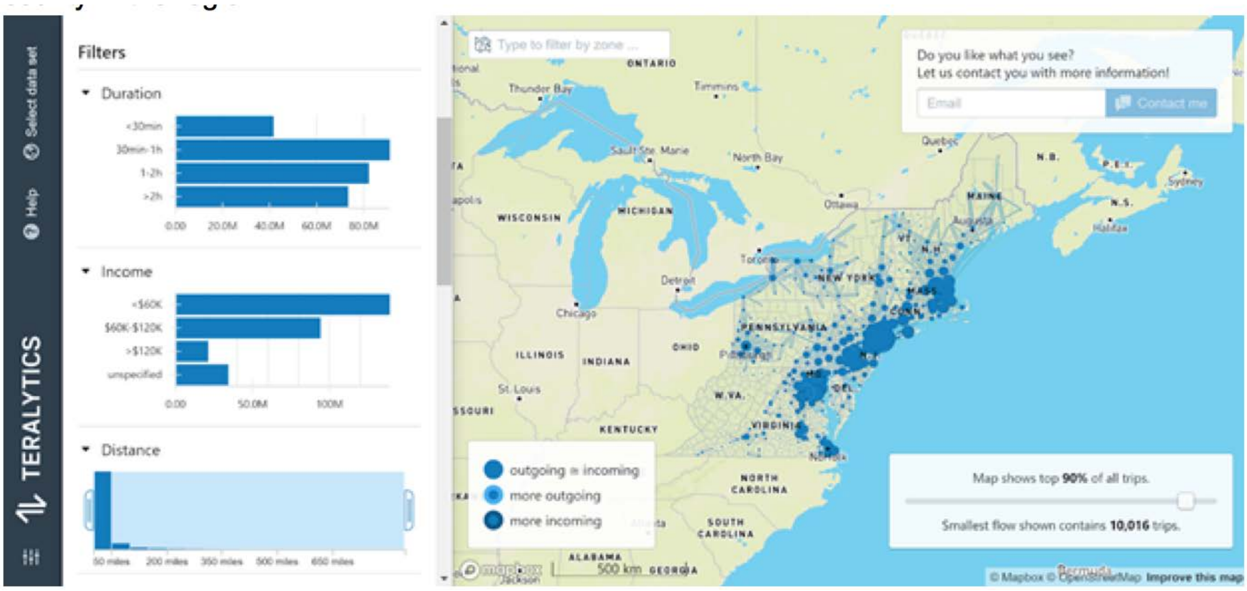
HOW IT WORKS

Teralytics primarily sources their data from the CDRs created by 60 million devices in the United States, which amounts to roughly 25% of the US cell-phone user population. They use machine-learning algorithms to analyze the CDRs. They are also able to access clickstream data from mobile devices so they can infer mode choice based on the phone applications in use and the user's travel speed. For instance, they might conclude that the most likely travel mode of a person who at one point was going 65 mph down a highway while using a navigation app is a personal vehicle, and then they can attribute that flow to a personal vehicle on all the roadway segments they were recorded on. Likewise, if a person is using a TNC's app while in motion, they can infer that it was a ride-hail trip. Because they can also infer a person's home and work location, they can combine that knowledge with public datasets such as the American Community Survey to estimate the distributions of travelers' socioeconomic characteristics, such as their income, age, and race.

USE CASE

Below is a screen capture of Teralytics' demo platform, based on CDRs in the northeastern U.S. A user can examine the distributions of travel duration, income, travel distance, ride hail and carshare mode split, travel mode, trip volumes, and incoming and outgoing flows from every county in the region.

Teralytics



Do you like what you see?
Let us contact you with more information!

Map shows top **90%** of all trips.

Smallest flow shown contains **10,016** trips.

Figure: Demonstration of the Teralytics Online Dashboard

Teralytics also discussed some of the other use cases they have addressed with their customers. For instance, in Los Angeles they were able to estimate the volume of traffic attributed to a baseball game at Dodger Stadium, helping the MPO understand the reach and distribution of trips that were attracted by the stadium at the zip code level. Teralytics was also able to pull out the time of day of travel and travel time.

ADVICE TO TXDOT

- Develop a data governance policy and have personnel who are responsible for actively managing and updating that policy.
- Prepare for common data formatting or even data standards and define workflows that would make it easier for a public agency to distribute their data for useful applications. This is especially salient for statewide agencies like TxDOT which have far-reaching jurisdiction and whose data could be useful for many agencies and applications.

Replica

Case Study

COMPANY OVERVIEW

Sidewalk Lab's primary project is the redevelopment of a Toronto, Ontario waterfront neighborhood. Their activity travel model, Replica, is the first of potentially several tools to be commercialized as a result of their Toronto project. There, the team will design and populate a mixed-use neighborhood that uses the latest in digital technology to demonstrate innovative approaches to providing an energy-efficient, livable community. Their research or lab-oriented approach to urban design means that as they develop planning tools and models oriented towards their Toronto project, they discover solutions to shared challenges that could be commercialized, such as Replica. Sidewalk Labs is a subsidiary of Alphabet, which also owns Google.

CURRENT CAPBILITIES

The agent-based Replica model uses person-level trip origin, destinations, route choices, and mode choices to construct a travel demand projection that can be fully segmented by mode, trip purpose, sociodemographic factors (such as income), and down to any minute of any weekday. Replica uses over 30 proprietary and public data sources to generate a synthetic population (synthetic proxies of each individual in the community of study). The proprietary data includes location-based services data from sources like Streetlight Data and vehicle probe data from sources like INRIX. They use local data such as transit, traffic, cyclist, and pedestrian counts to calibrate the ultimate model. Replica's model team works with the partner agency to calibrate the model within an acceptance criteria, such as being within a 15% margin of error at all intersection counts in a city.

So far, seven regions or states have begun work with Replica to build a regional or statewide model. The early use cases are diverse. Some agencies that don't already have robust real-time traffic data plan to use Replica as a source for operations data. Other agencies plan to use Replica for short-range planning (3 to 5 years) or for the evaluation of newly implemented transportation policies or projects. Finally, because many cities, counties, and MPOs do not have well-estimated origin-destination information for all the travelers in their

region, the Replica model serves as a much more detailed origin-destination matrix than ever before. They also make available interesting metrics that are not often shared from typically obscure travel models, such as an individual's probability of making each mode choice—this type of metric could be used to examine which travelers would be most likely to shift to an alternate mode given certain transportation policies or improvements.

Replica has an interesting pricing structure: the final cost of the model amounts to 20 cents per person in each jurisdiction per year. Sidewalk Labs offers the model on a subscription service, in which the year's subscription begins once the contracting agency and the model team reach an acceptable model calibration. However, in an effort to make detailed travel models more accessible, any jurisdiction within a regional or statewide model will be able to access the Replica model during the subscription period.

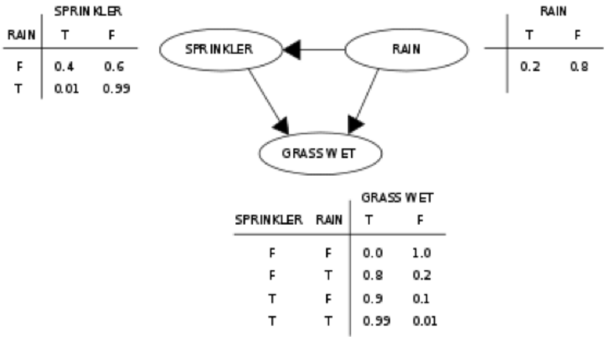
FUTURE CAPBILITIES

The model team is developing another product, called Scenario, that will be able to provide insight to policy analyses and scenario planning with high granularity. Because the Replica model is re-calibrated to account for short-term changes in travel behavior and seasonality every 3 months, it can theoretically already be used to estimate the effect on travel behavior that a change to the transportation network had by comparing to previous calibrations of the model. On the other hand, Scenario will be able to model questions at a micro-level, such as the effect that a new bicycle lane will have or what segment of the population would use a new transit stop. The team estimates Scenario will be available for integration with the Replica model within 6 months.

The Replica model team is also developing a common data standard for public agency data such as transit, traffic, cyclist, and pedestrian counts akin to Google's General Transit Feed Specification (GTFS). This is a result of the efforts their team has made to integrate multiple public datasets into the calibration of their model, but with its adoption could also ease many of the existing data silos amongst disparate transportation operators in a region.

HOW IT WORKS

Replica uses a statistical technique known as Bayes networks to create precise, yet anonymized, populations. Bayesian networks are directed probability graphs; to see what that means, consider the following graph.



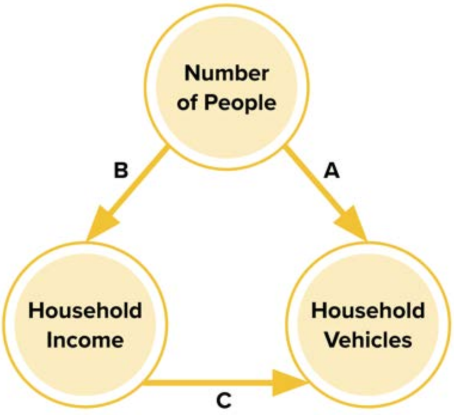
	SPRINKLER	
RAIN	T	F
F	0.4	0.6
T	0.01	0.99

RAIN	
T	F
0.2	0.8

		GRASS WET	
SPRINKLER	RAIN	T	F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

This toy example uses the probability of rain and the sprinkler turning on to calculate whether or not the grass is wet. Both rain and the sprinkler influence the likelihood of the grass's wetness, but the rain also influences whether the sprinkler turns on or not. The Bayes network provides both a visual interpretation for this series of interconnected probabilities, and a numeric-based definition. People can easily interpret the interconnected nodes, while computers can process the tables that represent those nodes.

The next image depicts how this idea generalizes to populations. The graph models a household. Each household has a different number of people, and that number of people influences the number of vehicles and the income of the household. However, the household income also influences the number of vehicles. A graph communicates the definition of a household much better than those last two sentences. More importantly, the graph provides a template for generating a synthetic household.



Adding more nodes to the graph, such as number of children/dependents, job types held, and maybe race/ethnicity, strengthens the similarities between a real household and the synthetic one. These synthetic households can then be placed in synthetic representation of a metropolitan area. That means adding more nodes, which denote how far the household is from their places of employment, as well as their access to different roadways, bus lines, and other means of transportation.

The final step requires survey data about the households in a metropolitan area, in order to guess at the probability in each node. For example, the only way to find out the distribution of people per household in a neighborhood is to look at a survey that asked households for that asked households for that information. With that information, Replica can say that each household in that area has a 20% chance of having 5 people, 40% chance of having 3, etc. This is how all the other nodes' base probabilities are found as well.

A final note: because the Bayes network uses probabilities to calculate the size and complexion of households, the synthetic population it generates is totally anonymized.

USE CASE

Although there is no full implementation of the Replica model in place yet, several regional and statewide models are in the development pipeline. They range in scale: from mid-sized regions such as the Portland, OR or Denver, CO; to mega-regions such as a Northern California swath that stretches from Sacramento to San Jose; to statewide applications in New York and Illinois. Kansas City will receive the first full implementation in late August 2018.

Because Sidewalk Labs provides open access to the model for all contained jurisdictions, this presents an opportunity for strategic investment at the state level. For instance, the state of Illinois has commissioned a statewide model from Replica, but they are using a planning grant to pay for a model of only the Chicago region initially. Ultimately when the model is available statewide, the coordination will provide substantial benefits to all transportation and planning agencies in the state.

ADVICE TO TXDOT

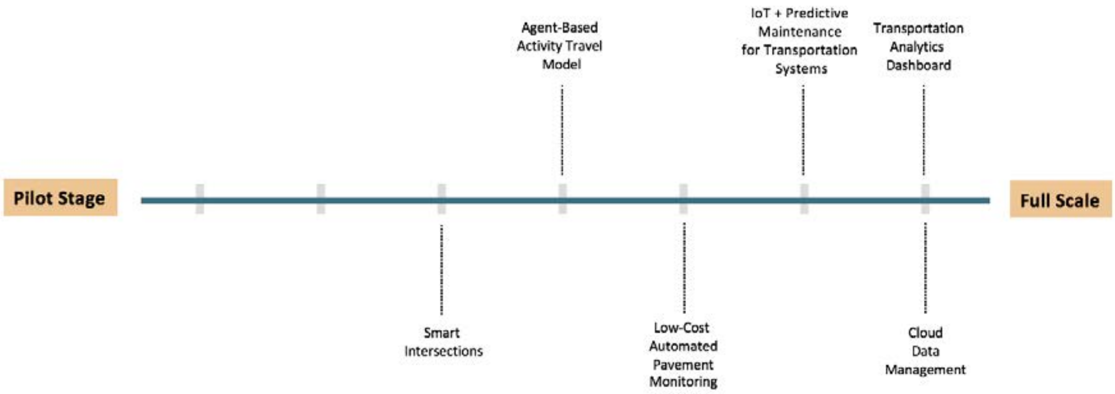
- Replica's model team has observed that inconsistent data practices among public agencies pose a significant barrier to building increasingly granular and comprehensive travel models. This obscures many of the analyses and insights that could be performed even without the use of a sophisticated model like Replica. Transportation agencies such as TxDOT should continue to emphasize the importance of adopting thoughtful, common data standards.
- Create consistent data practices.
- Recognize the power of clean data, and robust data pipelining.

CONCLUSION: TECHNOLOGY HORIZON

It is useful for public agencies to be able to assess the maturity of a particular technology, analytics technique, or data platform. This allows them to understand the level of risk that they may take on when implementing a new technology. A more risk-averse agency may only wish to invest in a product or service once it has been proven out by multiple other public sector customers and once a great need for it is identified. As a result, the risk of a lackluster solution and a perceived inefficient use of public dollars is minimized. On the other hand, an innovation-forward agency may be willing to assume more investment or resource risk by being an early adopter of a technology. These agencies may use creative or collaborative funding models to implement small-scale deployments of emerging technologies. At these agencies, even a small deployment has great value because it serves as a learning experience not just for the agency but for the industry at large.

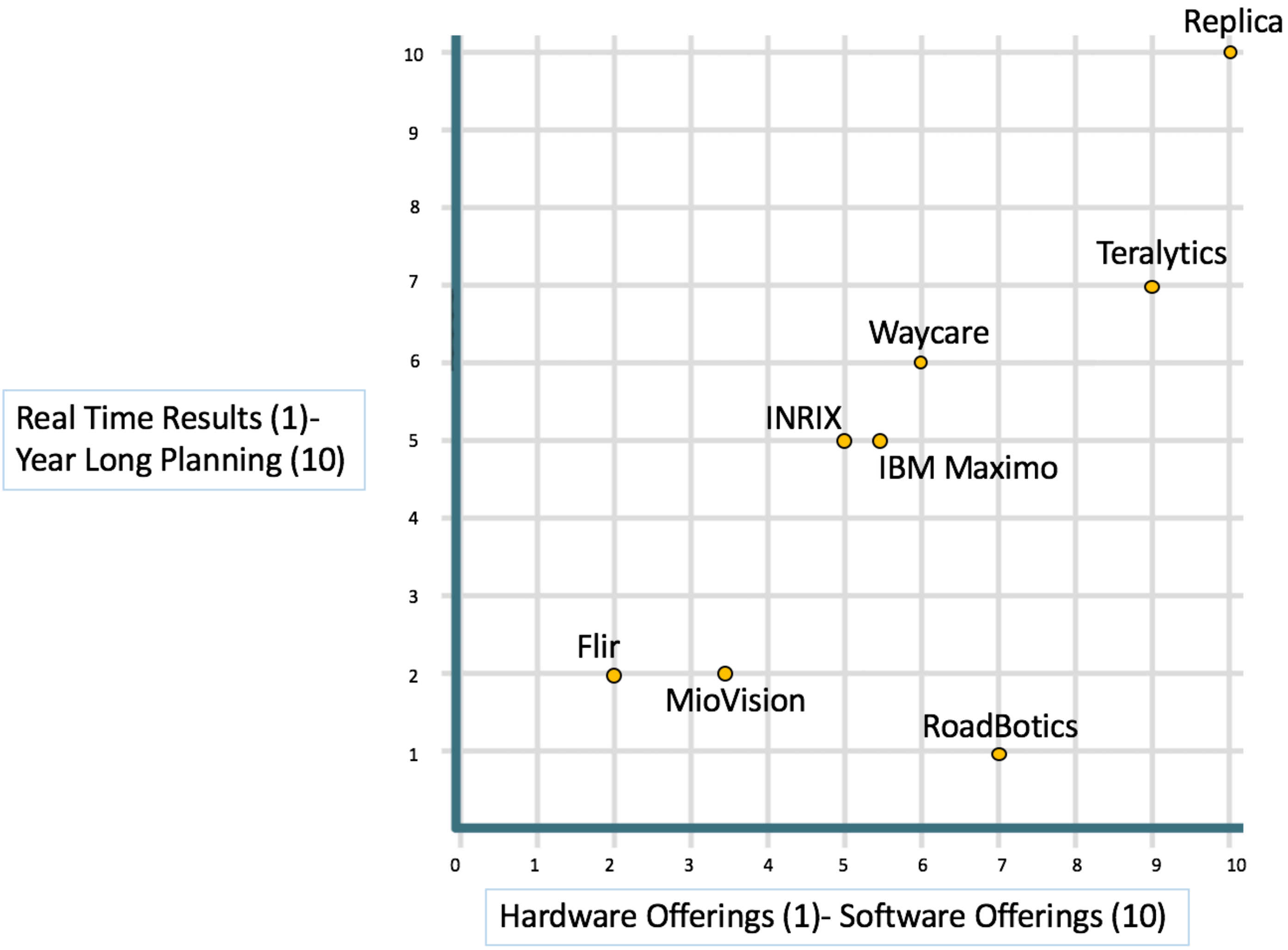
Here, a simplified scale for risk and maturity assessment is demonstrated using some of the featured technology products in this paper. There are three tiers: beta, pilot, and scale implementation.

Beta implementation represents the lowest maturity and highest risk products - in this stage, an idea may still be an advanced research product or only have a handful of test implementations in a lab setting. The concept is established, but there is no formal product on market. Pilot implementation represents a medium level of maturity and risk: a near or fully operational product is tested on private and eventually public right-of-way.
When partnering, a private sector partner may assume a greater share of risk in a pilot project with a public agency. Finally, scale implementation represents the highest maturity and lowest risk: multiple agencies have already procured and deployed a technology, and can share their lessons learned and best practices to guide following agencies.


Pilot Stage
Full Scale
Agent-Based Activity Travel Model
IoT + Predictive Maintenance for Transportation Systems
Transportation Analytics Dashboard
Smart Intersections
Low-Cost Automated Pavement Monitoring
Cloud Data Management

Stakeholder Map

As the digital revolution in transportation takes hold, it is sometimes unclear how different companies fit into the technology marketplace. Here they are characterized as either software, hardware, hybrid, or cloud computing (services?). The type of product that they sell or the pricing model they use may have implications for procurement and contracting.



When asked about what advice to give TxDoT in order to prepare for current and coming technology innovations, the interviewed companies chose focus on data. Some emphasized making data compatible across different platforms, others mentioned the importance of valid and appropriately sourced data, and others simply advised standardizing data practices. All the companies suggested having completely digital data stores, and highlighted the importance of making all of the data consistently formatted and easy to access.

AI and IoT work best with a clean and consistent digital ecosystem to function. That ecosystem is even more important in order for AI and IoT's results to be actionable. These emerging and emergent technologies may not require a thorough investment in data architecture, but a well maintained data environment magnifies their potential.

The best practices gathered include:

- digitize the collection of what data is not already digitized
- flexible, open architecture of databases
- extend value of data collected by sharing across agencies and considering cross-sector applications
- similarly, consider developing data standards across agencies
- take advantage of access/ownership to raw/disaggregate data collected within private sector solutions/platforms in addition to prepared dashboards/statistics
- consider efficiencies/opportunities to coordinate data collection at a higher authority level (like state or region-wide) to share down with all jurisdictional agencies
- Have personnel dedicated to managing data

RECOMMENDATIONS

OVERVIEW

When asked about what advice to give TxDoT in order to prepare for current and coming technology innovations, the interviewed companies chose focus on data. Some emphasized making data compatible across different platforms, others mentioned the importance of valid and appropriately sourced data, and others simply advised standardizing data practices. All the companies suggested having completely digital data stores, and highlighted the importance of making all of the data consistently formatted and easy to access.

AI and IoT work best with a clean and consistent digital ecosystem to function. That ecosystem is even more important in order for AI and IoT's results to be actionable. These emerging and emergent technologies may not require a thorough investment in data architecture, but a well maintained data environment magnifies their potential.

The best practices gathered include:

- digitize the collection of what data is not already digitized
- flexible, open architecture of databases
- extend value of data collected by sharing across agencies and considering cross-sector applications
- similarly, consider developing data standards across agencies
- take advantage of access/ownership to raw/disaggregate data collected within private sector solutions/platforms in addition to prepared dashboards/statistics
- consider efficiencies/opportunities to coordinate data collection at a higher authority level (like state or region-wide) to share down with all jurisdictional agencies
- Have personnel dedicated to managing data