Database Drift Detection

Artificial Intelligence: The Unforeseen Threats

An STS Research Paper submitted to the Department of Engineering and Society

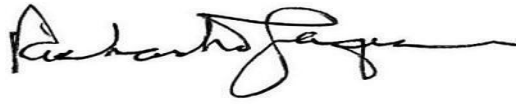
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By
Dev Kumar

Fall 2021

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

_____Date _11/7/2021_
Dev Kumar



_____ Date: 12/03/2021
Richard D. Jacques, Ph.D., Department of Engineering and Society

INTRODUCTION

The Information Age has made the industry more information-intensive and less labor and capital-intensive. With the required data, computers are in a state of continual improvement and can perform remarkable feats. In 1997, IBM's chess computer Deep Blue famously defeated the world chess champion Garry Kasparov (Nwankpa, 2018). Twenty years later, self-driving cars are being used as shuttles to transport the public to their offices. Artificial Intelligence (AI) is still in its very early stages. As of now, most AI systems have very narrow applications such as detecting faces or driving vehicles. Futurist Ray Kurzweil has predicted successful passing of the Turing test in 2019. But, until then, we remain in an era of Artificial Narrow Intelligence (ANI), which is weak AI, where special-purpose computer programs outperform humans in specific tasks such as games of skill and text analysis (Manheimer & Kaplan, 2019). However, “a long-held goal in the field has been the development of artificial intelligence that can learn and adapt to a very broad range of challenges” (“Risks from Artificial Intelligence,” n.d., p. 2). Eventually, it may be “superintelligent”, superior to humans in all domains. AI is a discipline that is able to be applied in practically all areas to aid us. Technology allows us to streamline tasks while also reducing the amount of error. I worked with the Database Engineering Services Team to produce an application that assists in standardizing databases' stored procedures that is significantly less susceptible to human error as the process is not only automated but also visualized in a table.

The rapid evolution of technology lets us automate decisions in order to increase efficiency. As a technical project, I developed an application that incorporates multiple

technologies in order to streamline a process usually done manually by developers. The technical project was started in the summer of 2021 and its scope and objectives changed as the summer progressed. While society's perception of AI is largely favorable, there are understated risks that come with its development. As a research project, a closer look to the uncertainties of AI will be investigated. An examination of current, future, and possible legislation will also be discussed in order to mitigate the likelihood of unanticipated consequences.

STREAMLINING DATABASE DRIFT DETECTION

While developing for the Database Engineering Services team for Salesforce, an American customer relationship management service cloud-based software company headquartered in San Francisco, I designed a program to discern the differences between special SQL Queries called stored procedures (sprocs). A database stored program is a "computer program that is stored within, and executes within, the database server" (Harrison, 2006). The store procedure is the most common type of stored program; it is a "generic program unit that is executed on request and that can accept multiple input and output parameters." These sprocs are saved SQL Queries to be rerun multiple times. Machines have different versions of each sproc, i.e. it is unstandardized. The project ensured that the sprocs were identical to each other or confirmed the differences. There are multiple teams in the Marketing Cloud organization that use and share sprocs and thus this update would ensure that code is standardized between teams.

The process starts with using a SQL Agent Job which calls a PowerShell script every day. Jobs calls scripts procedurally according to a set schedule. The scripts run once every night. The script gathers a list of instances in use by querying a database that stores active instances with "IsActive" set to "True".

Each instance is looped through and its sprocs are MD5 hashed. The code query itself is the only part hashed. The hashes are then extracted to a remote server. With the help of another team specializing in Splunk, the hashes are then forwarded to Splunk. Then Splunk Processing Language is used to form a query to count the number of unique MD5s for each of the sprocs. And so if a sproc has more than one unique hash, that means there are multiple definitions for it. Then the team manually checks the different instances and figures out what are the differences between the sprocs.

There are four total important pieces of data in the Splunk dashboard, the final product. The first is a report of procedures with multiple unique MD5 definitions. This is the main project feature. Another aspect is the ability to search the name of a sproc and calculate the number of servers that have each of its unique MD5s. If there are many servers with one MD5 and only one with another definition, it is highly recommended to check the latter for its inconsistencies, although ideally all servers should be checked why they have different definitions for their procedures. The third aspect lets the user search an MD5 and output which servers have the MD5. Team members usually know the names of different types of servers and can determine if those servers should have that particular MD5. The final feature is a line graph that illustrates the number of total unique sprocs per a user defined time interval. Optimistically, there is a downward trend, indicating that team members are finding inconsistencies and fixing them.

```
$instance=$args[0]
$query = "SELECT @@SERVERNAME AS ServerName,
SPECIFIC_CATALOG AS DatabaseName,
SPECIFIC_NAME AS ProcedureName,
SPECIFIC_SCHEMA AS SchemaName,
CONVERT(Varchar(32),HASHBYTES('MD5',Routine_Definition),2) AS MD5,
'ERROR' AS AlertLevel,
Convert(date, getdate()) as [TimeStamp]
FROM [Utility].[Information_Schema].[Routines]
WHERE ROUTINE_TYPE = 'PROCEDURE';"
```

Figure 1: **SQL Query used to extract an MD5 hash from each stored procedure**

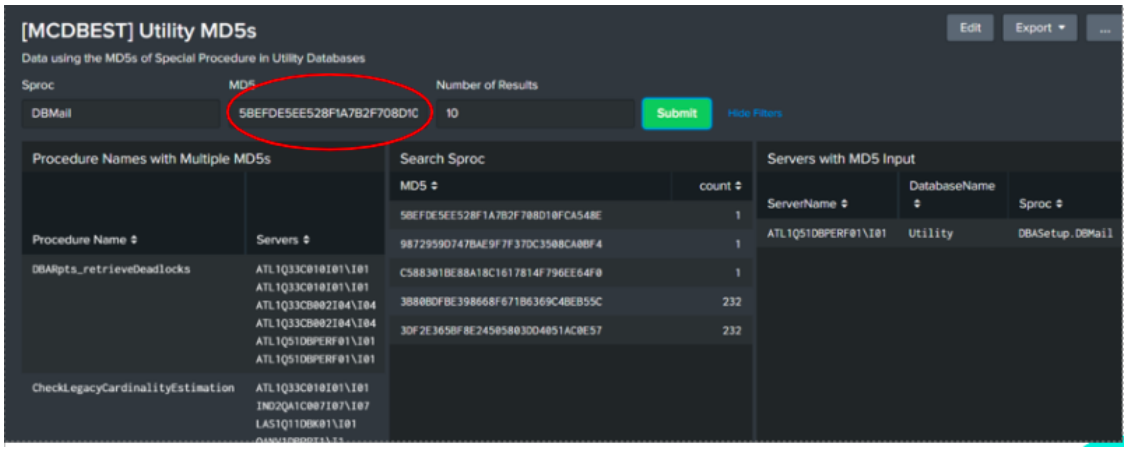


Figure 2: **Final product as demonstrated via Splunk**

The new program will be used for multiple databases across multiple teams. The product is complete in the sense that it does what it aimed to. However, there are plans to not only find the differences but also correct them automatically as well.

Even though I returned to the same team, I worked with a completely different set of technology. Summer of last year I primarily worked with C#. Last summer I worked with SQL, PowerShell, and Splunk. I could manage SQL and PowerShell reasonably well but I took the Splunk Fundamentals 1 course in order to familiarize myself with Splunk. Splunk was used because of its machine learning capabilities- it extracts intelligence from unstructured data. At first I was focused on passing the course and struggled with the actual implementation. Then I realized that I should be focusing on learning the material and my results improved. No one in my team of 20 had much experience with Splunk as well. I also worked on tickets related to the project I worked on the previous summer.

I also had to work with an external team to get Splunk functioning as intended. This was something I did not have much previous experience with. I experienced frustration when

progress slowed down drastically while waiting for others. Slack messages were often ignored and updates through the official Kanban dashboard were slow and sometimes done incorrectly. There were many bugs that took multiple days to solve. Most of them I resolved on my own by discovering fixes on the internet. Since the work was done remotely, I scheduled video calls for the remainder of them. There were design decisions and multiple languages considered and discussed with team members. For example, I had an implementation using exclusively SQL but was discarded so that Powershell would call SQL and then forward the data to Splunk.

Remote work was not as difficult as expected. Two of the greatest factors in finishing the project were the completion of previous internships at a startup, Amazon, and Salesforce. Just like the latest experience with Salesforce, two of the internships before it were both remote. The start-up internship was in-person but utilized ColdFusion which helps connect HTML to PostgresQL which gave the first exposure to database usage in industry. Salesforce had given me a generous work-from-home stipend which I used on a dual monitor setup, keyboard, and headphones. The internships gave valuable insight in how teams collaborate together to build a project used between teams. While the manager remained the same as last summer, another team member acted as the mentor, answering questions and problem solving when needed.

Even though I had a mentor, I also had multiple one-on-ones with different members of the team, either for help or seeing what they were working on. I also had meetings with external members to solve problems such as getting proper authentication or learning about particular features and practices exclusive to Salesforce. Overall, the project was challenging yet doable and I now have a better picture of the work I may be doing after graduation.

SOCIETY'S IMPACT ON AI AND VICE-VERSA

Artificial Intelligence (AI) is still in its very early stages, but it "is the most disruptive technology of the modern era" (Manheim & Kaplan, 2019). As of now, most AI systems have very narrow applications such as detecting faces or driving vehicles. Implementations less well-known but increasingly used include content analysis, medical robots, and autonomous warriors. As AI becomes more intelligent, we must not let the benefits of developing artificial intelligence obscure its drawbacks. Laymen are praising self-driving cars, Snapchat filters, etc., but there have already been cases of AI being used in controversial ways. One of the more well known examples was when Target sent high schooler coupons for cribs and baby clothes because the company inferred correctly that she was pregnant (Wagstaff, 2012). However, the main inspiration of the research topic came from a video of Elon Musk talking to Jack Ma, founder of the world's largest ecommerce and retail company, Alibaba, at the World Artificial Intelligence Conference ("Elon Musk," 2009). One of Musk's points was that policymakers do not generally have enough expertise on Artificial Intelligence to realize how powerful it is. Coincidentally, Jack Ma has a more optimistic view on the subject, stating "I'm not a tech guy. I think more about life" ("Elon Musk," 2009). Ma dismisses the threat of AI without examining the field more indepthly. It seemed peculiar that a man with such considerable power demonstrated such a restricted understanding of machine learning. Musk compared how humans see AI to how chimpanzees see humans; they do not necessarily see us as smarter and more capable, but as "strange aliens" ("Elon Musk," 2009). "Executives often overlook potential perils ('We're not using AI in anything that could 'blow up,' like self-driving cars') or overestimate an organization's risk-mitigation capabilities ('We've been doing analytics for a long time, so we already have the right controls in place, and our

practices are in line with those of our industry peers') It's also common for leaders to lump in AI risks with others owned by specialists in the IT and analytics organizations ("I trust my technical team; they're doing everything possible to protect our customers and our company") (Cheatham & Javanmardian, 2019).

Another complication of the development of AI may be the rise of emotional dependence of humans on such technologies. In fact, the feelings toward autonomous social robots will be similar to our affection for pets (Lin, Abney,& Bekey, 2014). Researchers claim that these robots will be so socially intelligent that they will have the ability to cause emotional harm (Lin et al., 2014). Thus, this type of dependence is different from other kinds of human dependencies on technology. Other topics the scientists cover will be discussed in the research project, including which ethical regulations robots should operate within on the battlefield, how society and ethics may change through robot development, and whether robots should have rights (Lin et al., 2014).

If society wants to make reasonable decisions concerning such technology, it would have to understand its capabilities, something only a small percentage of scientists can claim they do. Fortunately, awareness of AI's potential hazards is increasing due to organizations such as Centre for the Study of Existential Risk (CSER), and publications including Nick Bostrom's *Superintelligence*. CSER is a team dedicated to the mitigation of threats that could lead to either human extinction or civilization collapse. *Superintelligence* warns that AI could replace humans as the dominant life force on Earth.

Two responses to technological determinism, how technology influences societal culture and values, include the Social Construction of Technology (SCOT) and Actor Network Theory (ANT). SCOT argues that different social groups influence the development of technology. In

contrast, ANT argues that society is part of a network that contains various actors, all equal in their contribution to technological success. Figure 2 attempts to demonstrate a network of actants all contributing to the system. Car manufacturers have to abide by the regulations of policy makers. These manufacturers develop autonomous vehicles that are built by engineers. Customer perception and the media play a role in which vehicles get developed. It may be questioned how technology can have an equal impact as humans.

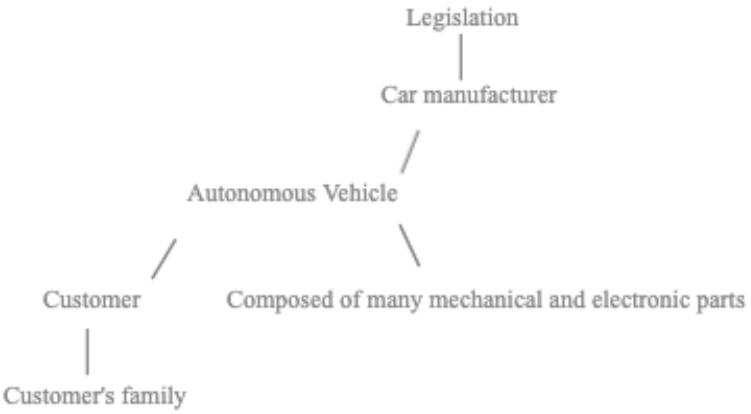


Figure 2: The actants of an autonomous vehicle, which is an actant for a bigger network. (Kumar, 2019)

For example, a technology can only be identified because we differentiate it from human actors. With the emergence of AI, technology more than ever before has the ability to communicate with humanity. Kiel Brennan-Marquez and Stephen Henderson (2019), law professors at the University of Connecticut and the University of Oklahoma respectively, question the impact humans should have in an increasingly AI driven criminal investigation (Brennan & Henderson, 2019). Role-reversibility is a notion that states that individuals making decisions should be subjected to the impact of the aforementioned decisions. The law professors believe that although some democratic traditions may change due to AI, role-reversibility is too

important to dismiss. They discuss how AI cannot be fairly subjected to role-reversibility as it does not have a consciousness. While a major component of the research project is concerning the impact of legislation on AI, the professors demonstrate how AI may be able to impact legislation (Brennan & Henderson, 2019). Another example of the role of AI development on legislation is demonstrated through autonomous vehicles. There is much debate on what role these vehicles should take in a lose-lose scenario (Hevelke & Nida-Rümelin, 2014). Figure 3 illustrates an example of a lose-lose situation. Three important questions emerge immediately. The first question is whether there should be a tort liability, a legal obligation of one party to a victim of an accident, with car manufacturers. If manufacturers are liable, it could deter the technology's development. The second question asks if "drivers" have a practical chance to intervene if the car is driving itself. Finally, Hevelke and Nida-Rümelin discuss if the driver should be solely financially responsible for the accident instead of criminally (2014). Revisiting laws in response to the growth of AI is not limited to the United States. Jaume-Palasi discusses how nations and international organizations are revising laws in response to the growing presence of AI (Jaume-Palasi, 2019). He explains how "every second week a new 'ethical code' seeks to present principles to fill a gap that public opinion fears has been opened by this technology" (p. 4). Sometimes ethicists posit situations in

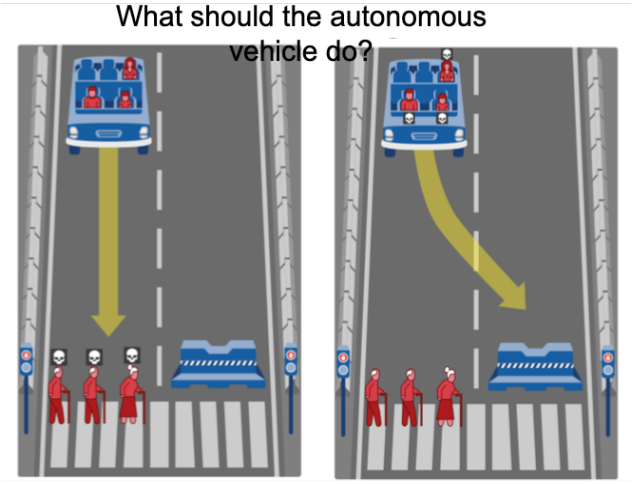


Figure 3: A lose-lose scenario an autonomous vehicle will have to take: It could hit the pedestrians or swerve to crash into another obstacle (Adapted by Dev Kumar from Edmond Awad, 2018)

which AI must choose between two actions with unpleasant consequences, and ask what the device should do. Since people in the same situation would face the same situations, these dilemmas were discussed long before computers existed (Parnas, 2017).

There have been high-profile cases of businesses's AI algorithms causing serious repercussions, including "VW's Dieselgate scandal with fines worth of $34.69B, Knight Capital's bankruptcy (~$450M) by a glitch in its algorithmic trading system, and Amazon's AI recruiting tool being scrapped after showing bias against women" (Koshiyama & Kazim, 2021). Possible solutions are both in use and being considered. Governments are starting to impose bans, regulators are fining companies, and the Judiciary may potentially treat algorithms as artificial "persons" in law. "A new industry is envisaged: Auditing and Assurance of Algorithms with the remit to professionalize and industrialize AI, ML and associated algorithms."

There is an argument that soft infrastructure, regulatory institutions, should focus on human values and be technology neutral so that technology cannot contribute to legal uncertainties. Society's role in the development of AI is critical in order to ensure AI is used constructively. Risse displays how artificial intelligence can affect "just about all rights" of the Universal Declaration of Human Rights through its discriminatory algorithms (Risse, 2019, p. 6). He argues that protection of human rights has an underlying assumption that human life should be valued more than other lives. However, the evolution of AI will bring entities superior to us intelligently and probably morally, and thus the research project will discuss how human rights should be preserved.

Another problem mentioned is the involuntary involvement of people as participants in machine learning algorithms. The expansion of computing technology will impact our

understanding of human rights. Roman V. Yampolskiy published a journal proving that it is impossible to consistently predict what choices a technology smarter than humans will make (Yampolskiy, 2019). There are fundamental theoretical limits to the ability to verify what a particular piece of code will carry out. Computer science is logical but an AI's algorithms are constantly changing, and so the results may be unforeseeable. The journal reveals that both AI's harmful effects and its decision making processes are unpredictable. Its unpredictability results in difficulty in making certain that its use abides by the law. Yampolskiy's paper is imperative as a major component of the research project is to investigate Elon Musk's assertion that AI will be too unpredictable to control.

There is also the danger of AI being used for nefarious purposes. Manheim and Kaplan argue that the "biggest social cost of the new technological era of AI is the erosion of trust in and control over our democratic institutions" (2019), noting the interference by Russia and voter manipulation by Cambridge Analytica and Facebook in the 2016 presidential election. The scientists go on to define "big nudging" as a form of "persuasive computing that allows one to govern the masses efficiently, without having to involve citizens in democratic processes." CEOs of large tech companies such as Google and Facebook may have a larger impact on our lives than the representatives we elect. We may not be at risk as a species but rather in terms of our democratic institutions and values. Because of the importance of data, technology companies will continuously push legal and ethical boundaries in pursuit of collecting more data to create models that make better predictions. There is little regulation preventing sharing this data to the government or private actors. This is a problem because the government may have to rely on the

technologies of these companies; Facebook's Deep Face facial recognition is much more powerful than the FBI's counterpart.

There have been some acts to contain the abuse of AI. In United States v. Jones, a majority of the Supreme Court signed on to or concurred in support of a "mosaic theory," under which long-term surveillance would be considered an illegal search in respect to the Fourth Amendment because of the detailed picture aggregated location information provided (Manheim & Kaplan, 2019). However, currently, there are "no regulations in the United States specific to artificial intelligence", but only vague ones of its applications. All states require that individuals be notified when their information has been compromised, usually through cyberattack, but state laws often have dissimilar and incompatible requirements. Three core principles of AI risk management include clarity, breadth, and nuance (Cheatham & Javanmardian, 2019). For clarity, we would "use a structured identification approach to pinpoint the most critical risks." This may be done by assembling a diverse cross section of leaders of business, IT, security, and risk management. For breadth, we would "institute robust enterprise-wide controls." We regulate and guide the development and use of "AI systems, ensure proper oversight, and put into place strong policies, procedures, worker training, and contingency plans." And finally, for nuance, we "reinforce specific controls depending on the nature of the risk", keeping in mind the "complexity of the algorithms, their data requirements, the nature of human-to-machine (or machine-to-machine) interaction, the potential for exploitation by bad actors, and the extent to which AI is embedded into a business process"(Cheatham & Javanmardian, 2019).

The rapid advancement of technology in the 21st century may be fruitful, but following precautions will be imperative in preventing unexpected consequences. Even simpler uses of machine learning such as the Database Drift Detector may evolve into something more intricate. As models become more powerful, it is of best interest to consider the significance of a world run by machines.

References

Awad, E. (n.d.). This question asks participants to decide in an emergency situation between staying on course and killing two elderly men and one elderly woman or hitting a concrete barrier and killing an adult man, an adult woman and a boy. Retrieved from https://www.pbs.org/newshour/science/in-a-crash-should-self-driving-cars-save-passengers-or-pedestrians-2-million-people-weigh-in

Brennan-Marquez, K., & Henderson, S. E. (2019). Artificial intelligence and role-reversible judgment. *Journal of Criminal Law & Criminology, 109*(2), 137–164. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=sih&AN=136927945&site=ehost-live&scope=site

Cheatham, B., Javanmardian, K., & Samandari, H. (2019). Confronting the risks of artificial intelligence. McKinsey Quarterly, 1-9

Elon Musk and Jack Ma disagree about AI's threat. (2019, August 29). Retrieved from https://www.bbc.com/news/technology-49508091

Harrison, G., & Feuerstein, S. (2006). MySQL stored procedure programming:. Sebastopol, Calif: OReilly.

Hevelke, A., & Nida-Rümelin, J. (2014, June 11). Responsibility for crashes of autonomous vehicles: An ethical analysis. Retrieved from https://link.springer.com/article/10.1007/s11948-014-9565-5

Jaume-Palasi, L. (2019). Why we are failing to understand the societal impact of artificial intelligence. *Social Research,* 86(2), 477–498. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=sih&AN=138503868&site=ehost-live&scope=site

Koshiyama, A., & Kazim, A. (2021). Towards algorithm auditing: A survey on managing legal, ethical and technological risks of AI, ML and associated algorithms. Retrieved from https://ssrn.com/abstract=3778998 or http://dx.doi.org/10.2139/ssrn.3778998

Lin, P., Abney, K., & Bekey, G. A. (2014). Robot ethics: the ethical and social implications of robotics. London, England: The MIT Press.

Manheim, K. M., & Kaplan, L. (2018). Artificial intelligence: risks to privacy and democracy.

Nwankpa, C., Ijomah, W., Gachagan, A., & Marshall, S. (2018). Activation functions: Comparison of trends in practice and research for deep learning. ArXiv.

Parnas, D. L. (2017). The real risks of artificial intelligence. Communications of the ACM, 60(10), 27-31.

Risks of Artificial Intelligence. (n.d.). Retrieved from https://www.cser.ac.uk/research/risks-from-artificial-intelligence/

Risse, M. (2019). Human rights and artificial intelligence: An urgently needed agenda. *Human Rights Quarterly*, *41*(1), 1–16. https://doi.org/10.1353/hrq.2019.0000

Wagstaff, K. (2012, February 17). How Target knew a high school girl was pregnant before her parents did. Retrieved from http://techland.time.com/2012/02/17/how-target-knew-a-high-school-girl-was-pregnant-before-her-parents/

Yampolskiy, & V., R. (2019, May 29). Unpredictability of AI. Retrieved from https://arxiv.org/abs/1905.13053.