Evolution of Knowledge in Social Media and Their Relationship to an Evolving Real World

Calton Pu
School of Computer Science
Georgia Institute of Technology
Atlanta, USA
calton.pu@cc.gatech.edu

Abhijit Suprem
School of Computer Science
Georgia Institute of Technology
Atlanta, USA
asuprem@gatech.edu

Aibek Musaev
School of Computer Science
Georgia Institute of Technology
Atlanta, USA
aibek.musaev@gatech.edu

Joao Eduardo Ferreira
Department of Computer Science
University of Sao Paulo
Sao Paulo, Brazil
jef@ime.usp.br

***Abstract*— The physical world evolves. The cyber world evolves and grows with big data, with social media as a major component of information growth. Classic ML models are limited by their static training data with implicit Complete and Timeless Knowledge assumptions. In an evolving world, static training data suffer from *knowledge obsolescence* due to truly novel timely information. Knowledge obsolescence introduces a widening distance between static ML models and the evolving world, called *cyber-physical gap*. Periodic retraining of new models may restore their accuracy temporarily, but subsequently their performance will deteriorate with widening cyber-physical gap. Knowledge obsolescence affects statically trained models of any size, including LLMs. Two major research challenges arise from cyber-physical gap: (1) collection and incorporation of space-time aware ground truth training data, and (2) understanding and capturing of the varying speed of information and knowledge evolution when the physical and cyber worlds evolve.**

Keywords—machine learning, dynamic model performance, evolution of knowledge, evolving social media

I. INTRODUCTION (*HEADING 1*)

The COVID-19 pandemic has been a stark reminder of the unpredictable evolution of our world. Large natural disasters such as hurricanes, earthquakes, and landslides all change our environment and require our adaptation to the evolution of our physical world. Since the advent of information technology (IT) in the 20th century, we have also created a cyber world from bits of information. Due to the ease of creating and manipulating bits, the cyber world changes and grows at a rate much faster than the physical world. According to one estimate (IDC [1]), the world data will reach 64ZB in 2020, and 175ZB by 2025. The growth is exponential: roughly doubling every two years.

Although the main sources of big data growth are usually attributed to IoT sensor data on the physical world, those data are often produced and consumed by automated systems. In contrast, human generated social media data in the cyber world (e.g., the estimated more than 1 trillion tweets sent since the founding of Twitter in 2006. Social media often have higher social impact, since they are produced and consumed by humans (originally). There is an increasing suspicion that some parts of social media may be produced (and consumed) by bots, but that speculation is beyond the scope of this paper.

Before the advent of large language models (LLMs [2]) and foundation models [3], the gold standard of classic machine learning (ML) research consisted primarily of training and evaluating classifiers from closed ground truth data sets (e.g., MNIST [4] and CIFAR [5]). On the positive side, this self-contained model of ML research and AI (artificial intelligence) enabled a fair comparison among the competing algorithms based on the same (static) ground truth data.

However, two significant issues were left out of this classic ML approach. First, the applicability of the classifiers to other data sets (i.e., generalizability) was left to the reader as a separate exercise, despite well-known overfitting problems. Second, the relationship between the classifier (as well as the ground truth it was derived from) and the real world became outside the scope. In other words, it is OK for the ML classifiers constructed in the cyber world to be disconnected from the physical world. In this paper, the word "knowledge" denotes the cyber world, and "world" denotes the physical world. The distance between our knowledge and the physical world will be called *cyber-physical gap*.

Before the arrival of ChatGPT [6] in November 2022, and the recognition of successful LLMs more generally, any suggestions of alternatives beyond annotated ground truth would have been routinely dismissed as irrelevant (see discussion in Subsection III.B), or even heresy, by the ML mainstream. Fortunately, LLMs have opened the door to training data beyond manual annotations. While we are not advocating any specific sources of ground truth training data, we hope our readers will keep an open mind on potentially fruitful possibilities beyond the exclusive orthodoxy of closed and static training/evaluation data sets.

This paper will discuss some of the significant research and practical challenges arising from the relaxed assumptions on new training data. Concretely, two major issues arise from an evolving world with evolving knowledge: generalizability and cyber-physical gap. First, static knowledge is unaware of the

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

changes, becoming out of date when the world evolves. Therefore, generalizability becomes an important measure of the decaying performance of static classifiers. Since generalizability depends on the distance between the physical world and knowledge of classifiers in the cyber world, it is also important to study the cyber-physical gap, which affects classifier performance as both the physical world and the cyber knowledge in classifiers evolve.

II. Basic Assumptions of This Paper

A. Illustrative Applications of an Evolving World

The classic ML approach to focus on classifiers trained from closed, static training data sets. They would work well in a static, or slowly changing world. In the 21st century, new technologies such as the Internet are bringing changes to the modern society at unprecedented fast speeds. Changes are happening and growing both in the cyber world (e.g., growth of social media) and in the physical world (e.g., modern infrastructure in transportation, smart energy generation and distribution, and global supply chain). In our research, we have been working with social media data on natural disasters such as landslides [7][8], and man-made hazards such as fake news [9].

Natural disasters such as landslides, earthquakes, and hurricanes are quite unique in their space/time coordinates. Each of them is also truly novel when they happen. Man-made hazards such as fake news are also novel by construction when they are introduced for the first time. There is a small amount of repetition due to laziness of both attack and defense, but the professional fake news producers are very inventive, e.g., claims of Bill Gates being responsible for the COVID-19 pandemic.

As an illustration of evolving cyber world, Figure 1 shows the prevalence of two fake news campaigns on the origin of COVID-19 virus from April 2020 to April 2022 (Figure 1 from [10]). The virus was attributed to 5G networks and Bill Gates were launched and reinforced at specific points in time (the high peaks in the normalized graph).

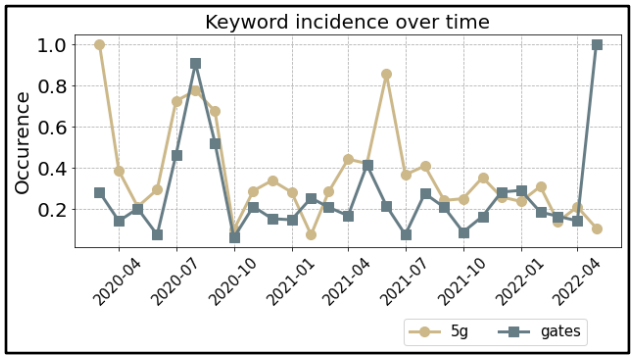

Figure 1 Illustrative Sample of Fake News Campaigns: 5G and Bill Gates Origin of COVID-19 Virus (Figure 1 from [10])

Truly novel events such as natural disasters and man-made hazards present significant challenges for static ML classifiers due to their inherent and/or intentional novelty. While excellent at distinguishing apples from oranges, static ML classifiers trained from past knowledge are often reduced to guessing when facing truly novel information. Even recurrent events such as elections pose significant challenges when new candidates appear, or the same candidates introduce new campaigns by changing their party affiliation.

B. Analysis of Assumptions in Static ML Classifiers

Novel events such as natural and man-made disasters appear and evolve quickly. Physically localized events such as landslides and tornados last a few minutes or hours. Large hurricanes last a few days, and their effects usually would pass in a few weeks as recovery efforts repair the damages and the affected population return to normal life. For different reasons, but similar motivations, man-made hazards such as fake news also have a relatively short expiration date, after which they lose most their value. Typical fake news campaigns only last for a few weeks. They are dropped and forgotten when the novelty wears off and fact checkers confirm the campaigns as fake news.

The knowledge expiration timeframe of a few days or a few weeks would be infeasible for classic ML knowledge acquisition process based on manually annotated ground truth training data, which takes months to years to complete. This mismatch has constrained most of classic ML research on physical and man-made events to retrospective studies of (yesteryear's) news and fake news. The recent development of LLMs have added new authoritative training data, but each new generation of LLM (e.g., the GPT-x family) still take (on the order of) a year to create. A typical example is GPT-3 and ChatGPT [6], explicitly constrained to knowledge from 2021 or earlier.

Given the impact of various (and often implicit) assumptions in classic ML/AI work, we make them explicit in Table 1. As an example, much of classic ML research assume closed ground truth training data sets (e.g., MNIST and CIFAR) without regard to the physical world. The justification is that the ground truth data were gathered from the physical world, and they are often called "real world data". While a valid snapshot of the physical world at the time of data collection, the continuity of snapshot validity in an evolving world has remained outside of classic ML work. Table 1 starts from these common assumptions of a static world in squire {1,1}.

The three columns of Table 1 represent a gradual relaxation of static knowledge assumption in {1,1}. The first column consists of the classic static knowledge trained from closed annotated data. The second column consists of static knowledge from varied and open sources, as done in LLMs. The third column denotes evolving knowledge from truly new sources with data on the evolving world).

Table 1 Static vs. Evolving Training Knowledge and Testing

Cyber knowledge & Physical World	Static Knowledge (annotated training data)	Static Knowledge (varied/open sources)	Evolving Knowledge (new training data sets)
Static world, in-distr. test	{1,1} Classic ML evalua.	{1,2} GPT-2 (1.5B par.)	{1.3} Static refinements
BIBO world, OOD test	{2,1} Virtual concept drift	{2,2} GPT-3 (175B par.)	{2,3} Contin. refinements
Evolving world & test	{3,1} Real concept drift	{3,2} GPT-4 (1.76T par.)	{3,3} This paper's focus

The rows in Table 1 represent a gradual relaxation of the static (and in-distribution) assumption of the test data set {1,1} in classic ML work. The first row illustrates the adaptation of different training strategies in a static world to improve classifier accuracy in large, closed test data sets. The second row relaxes the static test data assumption through continuous variations (often in sensor data streams). However, the variations are typically bounded (as in BIBO – bounded input and bounded output) models, allowing ML models with finite number of nodes to achieve good coverage of a constrained input state space. These assumptions often fall into the area of virtual concept drift [11] {2,1} and can be covered reasonably with relatively large LLMs {2,2} and {3,2}. The third row represents an evolving world without pre-defined bounds, by removing the two original assumptions (static knowledge in a static world). The area {3,3} is the focus of this paper, which follows the trend of pioneering work started from real concept drift [8][11] {3,1} and growing LLMs {3,2}.

III. Generalizability in an Evolving World

A. Generalizability of Static Models

The original definition of generalizability refers to a statically trained classifier's performance (e.g., accuracy) when tested with out-of-distribution (OOD) {2,1} input data. A representative area of these studies is virtual concept drift, where the OOD test data sets are also closed and static. Virtual concept drift studies differ from classic ML evaluations {1,1} in only one aspect: classic ML perform in-distribution testing (e.g., using k-fold validation), while virtual concept use OOD tests.

More directly relevant to the physical world is the area of real concept drift [8][11] {3.1}, where the OOD test data can change dynamically, following the evolution of the physical world. These unbounded variations of test data cause static classifiers to exhibit unpredictable accuracy over time. This situation is inevitable for static classifiers (column 1 of Table 1). In LLMs (column 2 of Table 1), the loss of performance is reduced by significant growth of training data.

B. From Variety of Knowledge to LLMs

In addition to annotated ground truth training data, ML have incorporated a variety of knowledge sources. For instance, unsupervised learning methods are able to find knowledge by exploiting existing properties in concrete data sets. An early example consists of k-means clustering algorithms finding association rules when data points contain clusters. Skipping through the historical development of alternative methods such as weakly supervised learning [12], few-shot learning [13], and other similar methods, we find the current generation of Large Language Models (LLMs) [2] as representative successful models that incorporate the most knowledge from a variety of trustable sources.

LLMs showed that the limitations of annotated training data sets could be lifted by incorporating other sources. As a concrete example, GPT-x series] integrate an increasingly large numbers of parameters from alternative trustable sources: GPT-2 (1.5 billion parameters), GPT-3 (175 billion), and GPT-4 (1.76 trillion). As result, LLMs have been able to provide useful answers to a much wider set of queries, corresponding to the size of their training data. The 3 rows in column 2 of Table 1 illustrate the three orders of magnitude parameter differences. LLMs represent significant advances over static annotated knowledge because of their significantly larger number of parameters and their design enabling improvements through refinements (column 3 of Table 1).

The growth and success of GPT-x series bought good news and bad news. The good news is the significant improvements we see in each generation. The bad news is that those improvements required orders of magnitude increases in their parameters. These scalability concerns can be seen clearly in the concrete implementations of GPT-x, with orders of magnitude increases every generation. Similar scalability concerns apply to LLMs more generally, since most LLMs are generated by "brute force" additions of more training data.

Postponing the scalability concerns to the next section (cyber-physical gap), the performance and wider applicability of current LLMs already have proven to exceed the classic (gold standard) ML models by far. We assume it is clear to (most of) ML community that the incorporation of alternative trustable sources is already an acceptable and promising practice.

IV. Cyber-Physical Gap

A. The Static Approximation Gap

We start from a recap of closed samples: ML has been described as *a function approximation of a sample*. The sample is the ground truth training data and the models and classifiers generated from the training data are evaluated against the ground truth. Quantitative criteria such as accuracy and precision measure the distance between the models and the training data during inference and testing. Let us call this distance function the *static approximation gap*. It is static for typical models (without randomization) because the models and the training data are static and closed, which could be called the Complete Knowledge assumption, since data points outside of sample (OOD) are ignored during evaluation. Under the Complete Knowledge assumption, questions such as generalizability became out of scope.

In addition to Complete Knowledge, a second significant assumption can be called Timeless World, where the knowledge contained in the training data is considered valid forever. Indeed, some concepts are timeless, e.g., digits representing Arabic numerals. However, the physical world is evolving rapidly, particularly due to recent technological advances. For example, the concepts of "telephone", "vehicle make", and "new COVID variant", all evolve with time, sometimes rather quickly.

Under the Complete Knowledge and Timeless World assumptions, a classic ML model has better performance if its static approximation gap is closer to zero, with a perfect score of zero gap being optimal. From this perspective, we have good news and bad news. The good news: this definition is well-defined and allows evaluation experiments with controlled variables. The bad news: both the Complete Knowledge and Timeless World assumptions are very far from an evolving physical world, rendering the models achieving (near-)zero static approximation gap criteria rapidly inapplicable to practical problems as the physical world evolves.

B. Cyber-Physical Gap

The changes from an evolving physical world can be modeled conceptually as a distance between the sample data and the state of interest in the physical world. We call this distance *cyber-physical gap*. Complementing the "internal" static approximation gap in classic ML work, cyber-physical gap is "external" to the training data. As an example, cyber-physical gap is present in IoT applications connected to the physical world, e.g., the distance between digital twins.

The conceptual linkage to the physical world raises a hypothetical question: Is it theoretically possible to measure the actual state of the physical world continuously? In principle, the answer would be no. Measurements using digital scientific and engineering instruments cannot achieve arbitrary frequency required by continuity. Fortunately, our models of the physical world, e.g., in physics and chemistry, show that sufficiently precise sampling can provide reproducible (discrete) measurements. As concrete examples, fine-grain sampling is realized by many control systems every day in practical, real-world applications.

This observation helps us divide the cyber-physical gap into two components: (1) the physical sampling gap, and (2) the cyber approximation gap. The physical sampling gap refers to the appropriate granularity of discrete sampling that will allow the system to model the continuous physical world with sufficient fidelity. This is feasible since it is achieved routinely by many control systems. The cyber approximation gap refers to the distance between a sample of the physical world with sufficient fidelity and the ML ground truth training data.

From the perspective of linking conceptually the cyber and physical worlds, we can see that the cyber approximation gap is a generalization of static approximation gap through the relaxation of Complete Knowledge and Timeless World assumptions. Acknowledging the incompleteness of ML ground truth training data in the physical world would allow the addition of (new and old) knowledge beyond the annotated training data. Furthermore, tracking and incorporating truly novel information from the real world, e.g., new COVID variants, will help us constrain the distance between the evolving physical world and new (and evolving) ML training data.

C. Cyber Approximation Gap Challenges

The growth of LLMs such as the GPT-x series demonstrates an increasing recognition of the limitations of the Complete Knowledge assumption in classic ML. Each generation of GPT-x increases the number of parameters by orders of magnitude. This meteoric rise of new knowledge highlights the significant challenges to the Timeless World assumption. We will briefly discuss some of the research challenges in building future AI/ML systems that attempt to relax one or both the Complete Knowledge and Timeless World assumptions.

The discussion on research challenges can benefit from the inclusion of another implicit assumption: Location Independence. Analogous to the Timeless World assumption, closed training data often imply the applicability of their models to any location. The limitations of the Location Independence assumption have been shown in applications such as autonomous driving, e.g., differences in pedestrian density and their behavior in diverse environments such as New York City and Arizona suburbs. The identification of Timeless World and Location Independence assumptions lead us towards a higher-level understanding of AI/ML models in the context of space-time continuum.

The time dimension spans the spectrum from Timeless (static models valid through the entire history) to Time-Dependent (dynamic models that vary as a function of time). Analogously, the space dimension spans the spectrum from Location Independent (static models valid everywhere) to Location Dependent (space-sensitive models that vary as a function of spatial coordinates). Since our laws of physics are often dependent on space-time coordinates, it should be intuitively acceptable that more sophisticated would evolve towards similar awareness. This direction of model evolution is supported by the requirements of IoT applications such as autonomous driving.

The first significant challenge of Space-Time Aware model development is the collection and insertion of space-time coordinates information into the (old and new) ground truth training data. Those coordinates were unnecessary under the Timeless and Location Independent assumptions, but the relaxation of those assumptions introduce the possibility of variations depending on space-time coordinates. This challenge explains the ongoing deployment difficulties of autonomous driving, where the Space-Time Awareness has yet to be made explicit. This challenge is even more important for man-made hazards such as fake news, where the changes are dynamically and intentionally introduced to confuse static ML models trained under the Timeless and Location-Independent assumptions.

The second significant challenge for Space-Time Aware models consists of a detailed understanding of the rate (speed) of knowledge variation as space-time coordinates change. If the sample data vary slowly and in small increments, then a coarse grain approximation may suffice. This is very useful for applications that interact with stable physical world properties such as autonomous driving. Understanding the speed of variation also will impose practical requirements on the speed of dynamic model deployment to achieve practical impact. For example, ML-based fake news filters must be developed and deployed before the entire fake news campaigns to have run their course.

In general, the knowledge variations include the waning of their applicability and rising of truly novel knowledge. The waning can be called *knowledge obsolescence*, and acquisition of novel knowledge called *knowledge rejuvenation*. An example of rapid changes would be fake news, where each new campaign requires knowledge rejuvenation, and their end signifies knowledge obsolescence (for practical filtering). Autonomous driving examples of knowledge obsolescence and rejuvenation include slow and long-term changes (e.g., construction of new roads and highways), fast and short-term changes (e.g., obstruction due to accidents), and legal variations (e.g., the blanket prohibition of red-light right turns in Manhattan without signs). From Space-Time Awareness perspective, the obsolescence and rejuvenation of knowledge in IoT applications (e.g., autonomous driving) should be treated explicitly with high priority, instead of a minor occasional nuisance.

The two significant challenges from Space-Time Awareness of ML models affect the development and evaluation of ML models at both knowledge and technical levels. For example, in the optimization dimension, approaches that reduce the size and complexity of ML models typically exploit statistical properties in the training and evaluation sample data; under the assumptions of Timeless and Location Independence the selection of most important parameters would be good forever and everywhere. As we take Space-Time Awareness into account, optimizations would become contingent on the statistical distributions at specific space-time coordinates, turning into a significant technical challenge.

Compared to the events and changes in the physical world, the evolution of knowledge is much faster in social media, where topics of interest flare up and die down in a matter of hours or days. This is the case for both true news and fake news. So far, much of research on social media (and media more generally) have been retrospective studies due to the Timeless (and Location Independence) assumptions in data collection. An example of blurring the physical sampling gap is the famous Sakaki paper [14] (5000+ citations), a pioneer retrospective study of historical tweets from a log. The paper contains a good study of the cyber approximation gap between earthquakes and tweets on them, but their study overlooked the physical sampling gap. Subsequent retrospective studies on social media as a reflection of real-world events often have followed this presumed omission of the physical sampling gap.

D. Need for New Research on Cyber-Physical Gap

From a classic ML perspective, a straightforward approach would be carrying out periodic retraining (e.g., done in the GPT-x series) to incorporate new information and knowledge. This "catch-up" approach can be seen as a relaxation of the Timeless Knowledge assumption by replacing it with a "Slowly Changing" Knowledge assumption. This scenario is illustrated in Figure 2 with data from a preliminary study on fake news.

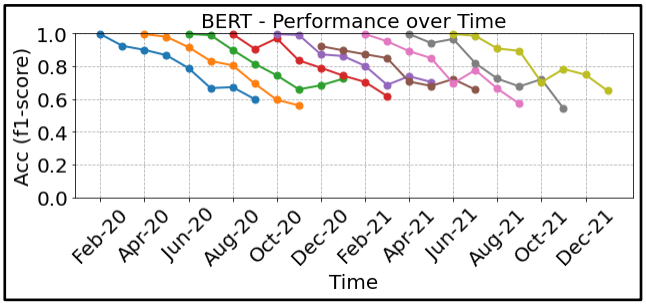

Figure 2 Performance Deterioration of Static ML Models Due to Knowledge Obsolescence (preliminary experimental study)

Figure 2 shows results from a retrospective study of BERT-based refinement fake news classification models generated through bi-monthly retraining. Each curve in Figure 2 represents a new model, trained from collection of COVID-related fake news tweets [16] since 2020. Each new model incorporates a modest amount of annotated fake news tweets from two previous months. When tested with data up to those months (in-distribution), the new model starts with an excellent F1 score (always very near 1.0). However, when tested on (novel) fake news tweets from subsequent months (OOD), the F1 score of static ML models deteriorate visibly with the increasing amount of novel fake news every two months.

Through a history of almost two years, Figure 2 shows consistent deterioration happens with all static models. Further studies suggest positive correlations between model performance deterioration and the amount of truly novel fake news (OOD). These deterioration results are consistent with the model performance deterioration observed in large-scale evaluations of many COVID pandemic predictive models. As an example, Fig. 3 in Friedman et al [17] shows general deterioration of model performance in 12 weeks for all predictive and prognostic models studied. These (and other) performance deteriorations due to the rise of truly novel information can be plausibly and reasonably explained by knowledge obsolescence.

We hope that by discussing the cyber-physical gap more carefully, with measurably bounded distances in space and time between the cyber and physical worlds, we (and the community at large) will be able to link more concretely the fast-changing social media (through Space-Time Aware ML models) to evolving events in the real world. Such concrete linkage would benefit applications such as social media-based detection of natural disasters such as COVID-19 pandemic and new variants, and man-made hazards such as constantly invented fake news.

V. SUMMARY

Classic (gold standard) ML models created from static annotated ground truth training data have struggled with generalizability issues when used in evolving environments such as the world in pandemic. This is due to the inherent Complete and Timeless Knowledge assumptions in closed training data. Classic ML model evaluations, e.g., using k-fold validation, are defined as distances between different subsets of a closed ground truth (in-distribution). Since out-of-distribution (OOD) data would fall outside of the Complete Knowledge, the generalizability of models to OOD test data became out of scope.

We call attention to the distance (called the cyber approximation gap) between ML models on the one side, and OOD data arising from an evolving world (e.g., IoT sensors and growing social media) on the other side. When application data evolve in space and time, static ML models would inevitably suffer from widening cyber approximation gap. To adapt to real-time changes in the real world, the ML models also need to bridge the physical sampling gap, which separates the (physical and social) sensor sampling of the evolving real world on the one side, and the actual real-world state on the other side. The combination of cyber approximation gap (between model and sensor data) and physical sampling gap (between sensor data and real world) form the cyber-physical gap between dynamic ML models and evolving real world applications.

There are (at least) two significant research challenges for bridging the cyber-physical gap in evolving applications that connects with the real world, e.g., IoT applications such as autonomous driving, and detection of natural and man-made hazards such as fake news filtering. The first challenge is

collecting Space-Time Aware dynamic training data and creating dynamic ML models that incorporate new knowledge as they arise from evolving applications. The second challenge is compensating for dynamic knowledge obsolescence and rejuvenation as the existing and novel knowledge wane and rise. The techniques for bridging the cyber-physical gap are part of our current research and beyond the scope of this paper. We encourage the readers to contact the authors for further details.

ACKNOWLEDGMENT

This research has been partially funded by National Science Foundation by CISE/CNS (2026945), SaTC (2039653), SBE/HNDS (2024320) programs, and gifts, grants, or contracts from Fujitsu, and Georgia Tech Foundation through the John P. Imlay, Jr. Chair endowment. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or other funding agencies and companies mentioned above.

REFERENCES

[1] David Reinsel, John Gantz, John Rydning. The digitization of the world from edge to core. Framingham: International Data Corporation, 2018, 16.

[2] Chang, Yupeng, Xu Wang, Jindong Wang, Yuan Wu, Kaijie Zhu, Hao Chen, Linyi Yang et al. "A survey on evaluation of large language models." *arXiv preprint arXiv:2307.03109* (2023).

[3] Bommasani, Rishi, Drew A. Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S. Bernstein et al. "On the opportunities and risks of foundation models." *arXiv preprint arXiv:2108.07258* (2021).

[4] LeCun, Yann. "The MNIST database of handwritten digits." *http://yann. lecun. com/exdb/mnist/* (1998).

[5] Krizhevsky, Alex, and Geoffrey Hinton. "Learning multiple layers of features from tiny images." (2009): 7.

[6] ChatGPT by OpenAI [https://openai.com/chatgpt]. Accessed September 12, 2023.

[7] LITMUS: A multi-service composition system for landslide detection, by Aibek Musaev, De Wang, and Calton Pu. IEEE Transactions on Services Computing, 8(5): 715-726 (2015).

[8] Suprem, Abhijit, Aibek Musaev, and Calton Pu. "Concept Drift Adaptive Physical Event Detection for Social Media Streams." In World Congress on Services, pp. 92-105. Springer, Cham, San Diego, July 2019.

[9] A. Suprem and C. Pu, "Event Detection in Noisy Streaming Data with Combination of Corroborative and Probabilistic Sources," *2019 IEEE 5th International Conference on Collaboration and Internet Computing (CIC)*, Los Angeles, CA, USA, 2019, pp. 168-177, doi: 10.1109/CIC48465.2019.00029.

[10] Suprem, Abhijit, Joao Eduardo Ferreira, and Calton Pu. "Continuously Reliable Detection of New-Normal Misinformation: Semantic Masking and Contrastive Smoothing in High-Density Latent Regions." *arXiv preprint arXiv:2301.07981* (2023).

[11] Gama, João, Indrė Žliobaitė, Albert Bifet, Mykola Pechenizkiy, and Abdelhamid Bouchachia. "A survey on concept drift adaptation." *ACM computing surveys (CSUR)* 46, no. 4 (2014): 1-37.

[12] Zhou, Zhi-Hua. "A brief introduction to weakly supervised learning." *National science review* 5, no. 1 (2018): 44-53.

[13] Wang, Yaqing, Quanming Yao, James T. Kwok, and Lionel M. Ni. "Generalizing from a few examples: A survey on few-shot learning." *ACM computing surveys (csur)* 53, no. 3 (2020): 1-34.

[14] Sakaki, Takeshi, Makoto Okazaki, and Yutaka Matsuo. "Earthquake shakes twitter users: real-time event detection by social sensors." In *Proceedings of the 19th international conference on World wide web*, pp. 851-860. 2010.

[15] Pu, C., Abhijit Suprem and Rodrigo Alves Lima. “Challenges and Opportunities in Rapid Epidemic Information Propagation with Live Knowledge Aggregation from Social Media.” In *Proceedings of 2020 IEEE International Conference on Cognitive Machine Intelligence (CogMI)*, Virtual Conference, December 2020.

[16] Suprem, Abhijit, Sanjyot Vaidya, Joao Eduardo Ferreira, and Calton Pu. "Time-Aware Datasets are Adaptive Knowledgebases for the New Normal." *arXiv preprint arXiv:2211.12508* (2022).

[17] Friedman, J., Liu, P., Troeger, C.E. *et al.* Predictive performance of international COVID-19 mortality forecasting models. *Nat Commun* **12**, 2609 (2021). https://doi.org/10.1038/s41467-021-22457-w