# Essay for WASP Software Engineering and Cloud Computing

Ahmet Balcioglu[1, *]

[1] *Department of Computer Science and Engineering,*
*Chalmers University of Technology, Gothenburg, Sweden*

## I. INTRODUCTION TO YOUR MY RESERACH

My research focuses on learning high-level and human interpretable representations (so-called "casual" representations) through black-box methods, such as neural networks. One of the main goals of my research is to be able to reliably generate such representations and make them configurable (Imagine giving the computer two images of different frames of a video and the computer being able to reliably interpolate between the two, that is unlike openAI's SORA it can understand the physics behind as well; hence causal representation.). Another aspect of my research is to use the available techniques for causal representation learning in real life applications, particularly for more sample efficient solutions. As such my research focus is primarily theoretical, and in my day-to-day work, I predominantly deal with abstract concepts. Most of the experiments I conduct is on simulated data.

Most causal representation learning and related areas focus on well-defined and rigorous problems and try to form well defined mathematical assumptions in exchange for the benefit of learning representations that are causal, and human interpretable. This trade-off also comes with a higher demand on data: in order to learn meaningful representations the data quality, not so much quantity, needs to be higher.

Also, as part of my research group, we are interested in machine learning applications for healthcare and take the view for every research question as its pertinence to particular applications in healthcare. One application I am currently working on is for personalized healthcare recommendations for conditions with multiple competing treatments. Here the idea is to learn a common latent representation for the population and recommend treatments for new patients based on previous patients with similar latent characteristics.

## II. PRINCIPLES FROM LECTURE

### A. Hidden Technical Debt

Just as every real life machine learning application has a hidden technical debt [1], machine learning research itself also experiences the same difficulties. There are multiple steps in my research that starts from theoretical modelling, and ends in empirical evidence using simulations. However, in practice it would preferable to have real-life data as opposed to simulations or at least have higher quality simulations. In addition, the theoretical model needs baselines to check and justify its contributions. In fact, most of research time is spent not on developing a theoretical model, but instead on justifying it, finding good baselines, finding good datasets, and running the experiments. As a result there is a hidden technical debt that is not strictly part of my research that needs to be paid.

This technical debt is not wholly similar to the software engineering case as the source of the debt is not maintaining a code base but iterative justification of the theory. In my opinion, there is no way around paying this technical debt; however, I will argue in the next section that paying of this debt could be made to, at least in my case, be part of an instructive and more positive research experience.

[*] Electronic address: ahmet.balcioglu@chalmers.se

### B. Agile as Research

As I mentioned, my research is in a relatively theoretical area, and currently a paper life-cycle goes on by starting a theoretical modelling of the problem that we can solve, then proving theoretical results on that model, and finally creating simulations for data gathering and testing the model to show empirically that it works. I would argue that this is similar to a waterfall model for research and there a few problems with it. First, once we're done with the theory we have no idea if that theory is useful empirically. Second, if the experimental results are not necessarily satisfactory, then we have no opportunity change the theory, instead what I have to do is to either publish it as a negative result or find a different dataset that works. Either of which - in my opinion - should not be part of good quality research. Lastly, I end up with a research that is guided only by positive examples, as opposed to combination of positive and negative examples.

Now, I want to turn this around and imagine an "agile" research model and how it would apply in my case. Again, I might have to start with a theoretical modelling of the problem; however, it not be the final model but rather a small quick theoretical model the next thing I do would be to implement the a simulation and try to see how it behaves empirically. This way, I can have a better understanding of the model and have better guesses as to its empirical performance. If the empirical performance is bad, then there is no need to provide theoretical results. Instead, I have a counter example to think about, and I have time to change the theoretical model and iterate. Once I find a satisfying model, as a final step I can prove theoretical results – which may have become easier with the insight gained from all the counter examples.

Working this way would allow me to spot blind spots early on, allow my research to be guided by practical considerations as well as theoretical justifications, and finally would give me negative examples to think about and try to understand. Even though software engineering and core machine learning research solve very different problems, in both cases it is important to catch and understand problems early on, in addition research questions can benefit greatly from negative examples.

## III. PRINCIPLES FROM GUEST LECTURES

I could not find a good angle within which to relate these principles to my research, these were the points in lectures that inspired me the most.

### A. Organizational Power

- **Reward**: administer desirable outcomes (rewards e.g. pay, promotion)
- **Coercive**: to give or recommend punishments (reprimands, undesirable work tasks...)
- **Legitimate**: from formal position, comply out of obligation/responsibility
- **Expert**: from superior or special knowledge or skill => respect/credibility
- **Referent**: based on individual characteristics => attention/respect/admiration

The organizational power structure suggested here is something that I recognize from my day-to-day experience in academia and previously in the industry. Naturally, we are all familiar with the legitimate forms of power and the formal authority to provide rewards and punishments. However, more soft forms of power was something that every one of us noticed but never named. In the daily conversations, whether it was about academic work or industrial work, we all notice and respect the person with more experience and expertise. More interestingly, this person did not need to be an expert but rather needed only to have more authority behind what they say and defend it well. In that sense, it was anybody's game to have this kind of soft power.

There is another kind of power I notice which is not mentioned here, especially in industry, job descriptions are vague and depending on the person occupying the position the job position increases in importance and gains a more key role in the company. If the work environment is toxic, people use this power to impede tasks that they do not want.

### B. Kotter's 8-step Organisational Change Model

Step 1: Increase Urgency

Step 2: Build Guiding Team

Step 3: Develop the Vision

Step 4: Communicate for Buy-in

Step 5: Empower Action

Step 6: Create Short Term Wins

Step 7: Don't Let Up

Step 8: Make Change Stick

Hearing about the 8-step to change was quite instructive and important. At first, it was very surprising that, before anything, the first step to change is to increase urgency. I would have imagined that, for instance, building a team and finding support among ones colleagues is more important then to increase the urgency of the problem. However, with some reflection, these steps and the importance and precedence of creating urgency is something I see in my academic life as well. Especially in collaborative work, changing research direction loosely follows these steps where we first argue about what task takes the most precedence, why it is the most important, who should take it, and how to get results and feedback about the change.

## IV. RELEVANT PAPERS

### A. Defining Quality Requirements for a Trustworthy AI Wildflower Monitoring Platforms [2]

*a. Describe the core idea(s) of the paper and why it/they are important to the engineering of AI systems*

The goal of the paper is to shine an insightful and holistic light into the maintenance and design of AI systems through the lens of AI models developed for wildflower counting. The paper uses the wildflower example to justify the eight categories under ISO 25000 for a trustworthy AI – human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, environmental and societal well-being, and accountability – each of which the paper argues are pivotal for the deployment of a healthy machine learning model. The flower counting model also provides an opportunity for the authors to interview individuals who work in different parts of the ML-development life-cycle to flesh out the reasons and justifications for each of the eight categories.

Standardization of AI model development and establishing necessary standards for a model is crucial for fair and safe use of these systems. In accordance with the regulations having proper testing procedures and guidelines would ensure that all players in the industry adhere to eight standards. This is important as most "professions" in society today are regulated in some way; doctors swear an oath, engineering professions go through quality insurance procedures, accountants go need licensing and, in some cases, swear an oath, and the list could go on. All these professions are in some way "profess" to do something, whereas programming in general, and machine learning in particular professionals do not need a degree or do not have to obey by any specific regulations. This makes ML development a notable exception in today's job market.

*b. How the paper relates to your own research*

The interpretable models offered by causal representations and the explainability requirements fall in together nicely. If we understand how the representation of the model changes on different data points, we could understand how the model representation diverges from human understanding of those features. It would be ideal to have some unusual test cases in the model, in order to test for model robustness which is defined as resilience towards invalid or perturbed data points. Controllability does not directly pertain to model implementation but to human-AI interaction. In essence my research does not diverge too greatly from other ML research as the last decision should be made by humans – especially in healthcare scenarios where the well-being of patients are at stake. Causal representations are perhaps most useful in collaboration effectiveness since, when deployed to ts full potential- the user might be able to plug-and-play with the system using human-interpretable features. Model correctness, privacy, and unfair bias are all an important part of data collection process. My research would benefit from, or would need these for a successful deployment, just like any ML project.

*c. How your research and its results would fit into a larger AI-intensive software project where one of the core ideas from the paper would benefit the project if applied. Describe both how the paper could help improve the project and how your WASP research would fit into the project.*

From the description given in the paper, the wildflower counting project works by taking photographs of meadows using a camera attached to a drone. As far as the basic inference model is concerned, it is not easily justifiable that a different representation would lead to better prediction accuracy. However, causal representations may help the model to extract and disentangle information related to the background and foreground which would lead to a more robust model. For instance, the new could make better use of images that are taken in different weather conditions and have more accuracy in unusual weather conditions.

*d. Discuss briefly how your research could be potentially adapted/changed to make AI engineering in the project based on the idea of the paper even better/easier.*

I think robustness under differing background conditions would be the main contribution my research could contribute to this project. In order to make better use of my research a few elements of the data gathering process would need to change. For instance in order to get a wide variety of data images taken from the same location with different weather conditions would be required, likewise it may be beneficial to equip to drones with torches to simulate the effect of sun and shadow. Another beneficial effect would be in deployment, as the model learns features about various backgrounds we could also use those features during deployment to for secondary prediction tasks – for instance to check weather the model prediction of the background fits with human observation.

### B. Improving Generalizability of ML-enabled Software through Domain Specification [3]

*a. Describe the core idea(s) of the paper and why it/they are important to the engineering of AI systems*

The paper is in the intersection of software engineering and machine learning research. The main goal is to try to generalize concepts that are generally interesting to machine learning such as text and visual data in terms of object oriented programming relations in differing domains. For instance a pedestrian, or a cyclist can be well-defined through "is-a" (person), "does-a" (walking, or moving), and "has-a" (bicycle). Through this kind of definitions they try to generalize potential concepts that are missing in the data and suggest a structure for future data collection.

*b. How the paper relates to your own research*

The relations suggested in the paper between the objects are in essence causal relations that my research aims to extract. In that sense, the structure of the data collection process suggested in this paper and the necessities of causal learning are very tightly aligned. The paper also acknowledges this: "When there is a causal relationship between the dependent and the independent type of variable. The level of detail and completeness of the specifications, in addition to their correctness, significantly impacts the MLSC performance." [3]

*c. How your research and its results would fit into a larger AI-intensive software project where one of the core ideas from the paper would benefit the project if applied. Describe both how the paper could help improve the project and how your WASP research would fit into the project.*

My research would benefit very directly from the suggested data collection method. Especially, if it is in a test or simulated environment. It would be possible for instance to learn the causal differences between pedestrians and cyclists through changing the simulated environment and learning contrastive models. It would also allow for asking different kinds of causal and counterfactual questions.

*d. Discuss briefly how your research could be potentially adapted/changed to make AI engineering in the project based on the idea of the paper even better/easier.*

Most works in causal representation learning do not work on image data or use big models such as image transformers. However, in a project like this in which a big image dataset is gathered with structure and with regard to causal changes in the system. My research are could be potentially integrated to train transformer models and accompany this kind of data gathering with causal questions.

[1] David Sculley, Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips, Dietmar Ebner, Vinay Chaudhary, Michael Young, Jean-Francois Crespo, and Dan Dennison. Hidden technical debt in machine learning systems. *Advances in neural information processing systems*, 28, 2015.

[2] Petra Heck and Gerard Schouten. Defining quality requirements for a trustworthy ai wildflower monitoring platform. In *2023 IEEE/ACM 2nd International Conference on AI Engineering–Software Engineering for AI (CAIN)*, pages 119–126. IEEE, 2023.

[3] Hamed Barzamini, Mona Rahimi, Murteza Shahzad, and Hamed Alhoori. Improving generalizability of ml-enabled software through domain specification. In *Proceedings of the 1st International Conference on AI Engineering: Software Engineering for AI*, pages 181–192, 2022.