UNIVERSITY OF WINDSOR

COMPUTER SCIENCE SOCIETY

MONTHLY NEWSLETTER

FEBRUARY 2022 • ISSUE 1 • VOLUME 1

WHAT'S INSIDE

WIKIPEDIA'S 6 DEGREES OF SEPARATION

BY NIKO JONES

Have you heard of 6 degrees of separation before? It describes how anyone can be connected to anyone with 6 people in between. You technically are hitting it off with Bill Gates. Pretty much best friends with Elon Musk.

Well now we will get to test this through the only means we have. WIKIPEDIA!

On **February 18th**, we will be hosting a Wikipedia race for the UWindsor Computer Science Society. Currently we are aiming to host this event in-person, but nothing has been finalized and whether it is online or in person is still up in the air. We will be finding out how on Earth Shaquille O' Neil is related to the Theory of Relativity and how Pizza Hut is related to Google.

Come on out for a night of fun and wiki races!
More information is to come. Please keep an eye on the Discord server for new information and sign-ups.

THE GRIND FOR INTERNSHIPS

BY ABDULAZIZ KHAN

Internships. We all want one to get experience. Some of us get interviews and some don't. Everyday students at the university are constantly applying to various internships either through coop or their own personal websites to land an interview. These internships might be with big companies or even smaller upcoming companies.

The one question that we always ask is why it is so hard to get an internship. Sometimes it's because of the work that the company might be doing that they need someone who is a perfect fit, other times it might be because our application does not parse well.

Regardless of the reason, we push forward and continue to apply to various postings of various companies. Some tips that I thought could help with the grind for internships would be the following.

Number one: make sure to first connect with someone that works as a University Recruiter for the company you are applying for because your resume will most likely land on their table.

Next would be to keep applying even if you don't have all the skills that are stated in the job description, most of the time companies will teach you the skills regardless of experience.

The final tip would be that you don't feel defeated if you get a rejection, move on and continue to apply to various other postings. Remember that all you need is one internship to swing open the door to other internships and connections.

I hope these tips help to make the internship grind a bit easier. If you need assistance with your resume or preparing for interviews, the CSS has many resources available to you on their website, YouTube channel and discord server! Also feel free to contact me via discord if you want a personal resume critique or need some questions answered about the job market.

Thanks for reading!

INTRODUCTION TO OPEN DATA

BY RYAN PRAIRIE

Imagine you have an idea for a piece of software. Let us say it is an app to find the most efficient route between your classes. You spend the weekend working on your idea. After some long nights, scribbled designs on paper and a dozen crushed Red Bulls in the garbage, you run it and it works!

Hearing about your success, friends come asking to try out this app. You then explain, "To get the app to work, I had to do a bunch of rough estimates of the distance between buildings and rooms". For them to be able to use your app, they must put in what classes they are taking, where the classes are, and the rough distances between each location. That is very time-consuming and a lot of work.

Now imagine a world where you make your app, you plug it into an existing UWindsor Open Data API, and all the users must do is sign in and put in which classes they are taking. The rest is handled by the UWindsor Open Data API.

Open Data or an Open Data Initiative (or Policy) is a service that a group provides that gives their data as data sets or an API to be easily accessible by the public.

Right about now, you may be asking, "Why do I care? What does it mean for me?" That is a particularly good question. Open Data lets us create tools that we have not even thought of yet.

The University of Waterloo has an Open Data Initiative, and they have a tremendous number of apps made by students. The point is empowering students to create exactly what they want, how they want, with amazing tools.

Do you want an Open Data Initiative at the University of Windsor? CSS agrees! Now how do you get it? Well, you can help CSS convince the university to create this initiative. If interested, send an email to css@uwindsor.ca and let us know exactly how much you want it.

If you want to read more about it, you can read our proposal.

USING SYNTHETIC DATA TO IMPROVE MACHINE LEARNING

BY JUSTIN BORNAIS

Artificial intelligence is becoming an increasingly prominent field of computer science. The idea of computers being able to think for themselves is both amazing and is what motivates me to learn more about programming.

There is only one drawback to AI (Artificial Intelligence): data is needed for the model to learn and mature. A lot of data in fact, especially for sophisticated models. For example, Inception V3 (a model for image classification developed from Google) requires over a million datapoints for training.

That is a lot of data! Collecting the data can be expensive and incredibly time consuming. If only it were possible to generate your own data instead of having to collect it organically.

That is where synthetic data comes in. Synthetic data can solve lots of problems regarding AI. For one thing, it allows developers to work with more data without having to overcome various obstacles they may face. Some of these obstacles include the cost of collecting information, as well as potential privacy concerns.

It also will preserve relationships between variables in a model, as the data will be created intelligently instead of randomly. Not only that, but synthetic data can simulate conditions not yet encountered in your organic dataset. This can be highly beneficial for the health industry, robotics, security and other areas.

Now of course, any synthetic models generating data from already existing models can only go so far. They can only simulate general trends and reproduce specific properties derived from their organic counterparts. Though, a study at MIT in 2017 showed that training a model from real versus artificial data showed no significant performance difference 70% of the time.

Synthetic data is a growing reality in the field of artificial intelligence. As developers discover new and improved ways to generate data, we will see an increasing amount of synthetic data in various machine learning applications.

I hope this excites you as much as it excites me!