

WASP Software Engineering Course: Assignment 02

Arka Ghosh, PhD student at Department of Computing Science, Umeå University,
arkag@cs.umu.se

My PhD research concerns both foundational and applied research in the context of flexible and efficient management of large amounts of richly structured geospatial-temporal data by relying on the paradigm of Virtual Knowledge Graphs (VKGs), also known as Ontology-based Data Access (OBDA) in literature. A lot of data is now being generated on the internet, social networks, sensors and so on, where typically neither its structure nor its meaning is explicitly specified; moreover, data coming from such sources is typically highly incomplete and differently formatted. Ontologies can help to overcome these problems by providing well-defined semantics and background knowledge. Specifically, I am investigating how to use the knowledge graph framework to integrate large amounts of geospatial-temporal data from various sources (e.g. satellite sensors, ground sources, historic data) and in different formats under common semantics but we have realized that using this knowledge graph might cause performance issues, therefore, we are concentrating on the virtual approach of the proposed framework, hence the virtual knowledge graph (VKG). Geospatial data plays an important role in many real-world application domains nowadays such as GPS, urban planning, agriculture, remote sensing, tracking sea-level variation, monitoring global climate change, etc. Thus, it is of great interest to study how to enrich the VKG paradigm with Geospatial data.

Topic 1: Requirement Engineering

Another morning in some Silicon Valley in some large city. A client is walking toward a big software company with a rough idea of desired software which will change the client's business output massively. The client has approached the company's designated people and pitched the idea of potential software. Now the client is neither a technical expert nor a software engineer. So, the client explains what is desired in the end product in natural language. This includes the general notion of all the required functions the end product must have and all the features the client wishes to see in the software. Now the requirements expressed in simple language are not enough to build software but it's a mere starting point. Now the people in the company specifically system analysts try to formulate and gather relevant information out of what the client said and document them in a proper technical manner. This process is called software requirement engineering. The goal of requirement engineering is to develop and maintain a sophisticated and descriptive 'System Requirements Specification or SRS' document which marks the beginning of the software life cycle and determines what software is going to be produced at the end. SRS document contains how the intended software will interact with hardware, user interfaces, speed of operation, the response time of the system, portability of software across various platforms, maintainability, speed of recovery after crashing, security, quality assessments and technical limitations etc.

In my research, I still do not have a plan to build full software. I am trying to formulate a new way to represent geospatial data in a knowledge graph paradigm with a virtual approach so that future software products can rely on this virtual data structure to reduce time and space complexity in terms of data storage, processing, analysing and inferring new knowledge.

Topic 2: Human Factors

In the movie Terminator 2: Judgment Day it is shown what could happen if humans are not considered as a part development of superior AI in software engineering and let machines do all the jobs autonomously. Now, this was a science fiction movie dating back to the year 1991. It

seems impossible back then but after 31 years, at present with the massive advancement of AI in software engineering the involvement of humans is getting limited. Recent advancements in AI and software engineering have led to building self-coding applications like DeepMind's AlphaCode AI and OpenAI's Codex engine. They have been proven to write code at a competitive level and can beat an average programmer in code tests, pretty awesome right but somewhat scary. All of these advancements are limiting the involvement of humans as data providers only. Humans are not as perfect as autonomous machines in case of repetitive jobs or solving complex problems but machines need instruction from humans to perform. Humans are frequently assumed to act perfectly in the world of machine learning, with the ability to simply provide correct data to machines. So, it's a possible research direction where one can explore how to increase the active involvement of humans in the development and advancements in AI and software engineering. Humans are important and all these advancements and developments are meant to improve their quality of living and make life easy than before. The rapid progressions in software engineering and ML/AI are necessary but it is also important to ask to what extent what we do represents real life. Because artificial intelligence and machine learning are becoming more prevalent in our daily lives, we must keep the human perspective in mind.

In my research, the use of the virtual knowledge graph (VKG) paradigm will free the end-users from the need to know the precise structure of geo-spatial-temporal data sources. In this way, the integrated data will be much easier to access, process, and analyse by the human users which will help them in critical decision-making problems in future business and industry.

Topic 3: Human-Computer Interaction (HCI)

Human-Computer Interaction (HCI) deals with the interactive connection between humans (users) and computers. It is an interdisciplinary field that spans computer science, computer engineering, psychology, human factors, and cognitive science. It was more related to personal computing in the past but nowadays Virtual Reality (VR) and Augmented Reality (AR) have come into play and they are narrowing the gap between humans and computers as time passes. A suitable example of this is as follows, I had the opportunity to go to "Epiroc" a Swedish manufacturer of mining and infrastructure equipment as a part of the WASP tour. There I fixed a battery of a huge mining truck inside a summation using VR gadgets, pretty cool right I did not have to go inside a factory or a mine (mine are dangerous as a workplace) to learn how to fix a battery, I just got training experience through VR. They have just created a training simulation inside a computer-generated virtual world equipped with actual world scenarios and people can train there without breaking a sweat. It's proven at the Epiroc that the people who went through this VR training program are performing better in the real-world mining scenario than those who didn't go through it.

So, I think it would be cool to have these kinds of simulation systems based on VR and AR in every aspect of human life such as teaching, programming, medical science, the software industry, child care, entertainment etc. This would make humans' life so much easier.

I can think of one possible connection that Human-Computer Interaction (HCI) has with my PhD research. Don't know for sure now since I have just started my PhD. In my research people can use a VR based visualized system to view, and interact with a 3D earth system to check about its ever-changing climate, and temperature variation or teach kids about the different aspects of geodata without worrying about integrating geo data from various sources as it will be taken care of with virtual knowledge graph paradigm.

Future trends in Software Engineering and AI/ML

My research project is connected with geospatial technology and satellite remote sensing where data is being generated in a huge amount (severe terabytes of data coming from satellites). These data are used to monitor climate conditions, sea surface rise, urban planning, traffic mapping, coastal mapping and many more. It will be pretty cool if there is a knowledge inference system over this data that infer and predict new knowledge from the real-time geospatial data which will help to make business decisions based on geolocations and other geo data. Let's take an industrial example, Volvo uses the archived digital elevation model (DEM) data, and real-time location to monitor and manoeuvre their autonomous trucks between source and destination. In my research, I am trying to integrate geodata using the knowledge representation technique of AI so it can be used in later advancements in Machine Learning and AI.

Recent works on this have been going on. Recently I read a keynote about **GeoAI**, Google geospatial initiative that uses the Google Earth Engine platform and machine learning techniques to process, analyse, and train big geo-data to maximize social benefits. The World Geospatial Industry Council (WGIC) outlines a few use case scenarios that demonstrate how the combination of artificial intelligence (AI) and machine learning (ML) with geospatial and Earth observation data can provide insights to produce practical and useful solutions for attaining global good.

The new trend in AI and software engineering is the creation of DALL-E v2. It's the AI system that generates original images and artworks from a natural language description as an input. It has taken the creative world and machine world by storm. Many people are using this to create their own artwork by typing it in simple English language.

Similarly, in programming and natural language processing, similar kinds of software are also coming into play. GitHub's co-pilot is an example of an AI Programmer that helps the coder to write code faster by writing trivial programming things and repetitive code by analysing the coder's coding preferences. I also think there will be more cloud-based IDE like Google Colab to automate and run ML models online in real-time.

The development and engineering in the software industry are going through rapid advancements in AI and ML fields. The industry is in the midst of an AI/ML revolution and we will see a more increasing focus on areas that include AI and ML in the future, creating job opportunities for AI and ML engineers globally.