Professional Interview

# Background

The person I interviewed has been working as a data scientist for six years. She spent the first two years after graduate school at an energy consulting company and spent most of her time making automated dashboards for higher ups and working on two projects. For the purpose of this paper, we’ll call this company Energy Consultants. The first was a project for predicting failures and maintenance for electric submersible pumps. The second was detecting methane leaks using satellite images and infrared drone footage in an image classification model.

Her second and current position is as a higher-level data scientist at a petrochemical company – we’ll call this PetroChemical Inc. She is on the Upstream Analytics team and has done work on production forecasting for the company’s wells in the Permian Basin and created a Bayesian optimization pipeline for well spacing and completion parameters.

# First Issue

The first ethical issue she came across in her career was at her first job at Energy Consultants. She was working on detecting methane leaks using satellite images and infrared drone footage. These images would be put into an image classification model to determine if there were leaks or not. The satellite images functioned as a base map and the infrared drone footage was used to detect potential methane leaks.

Some of the satellite images included residential neighborhoods near pipeline infrastructure. As she was looking at some of these images, she realized that the resolution was high enough to capture people outside, types of cars, and addresses. She became concerned about the privacy of the people who were living in these neighborhoods because their daily activities were observable to anyone who had this data. Additionally, if the data and results of this project were made public, it could indicate potential health risks of individuals and devalue their property.

My interviewee wanted to make sure that any data was anonymized and that any results found would be addressed by the client (a pipeline company), both in fixing issues in their product and in notifying residents of any issues discreetly and respectfully. Because the residents had not consented to the project, and therefore their homes and lives outdoors being photographed and analyzed, it was especially important to anonymize the data.

She went to her boss and presented her concerns. Her boss was grateful she came to him with the issue and organized a meeting with the rest of their team to discuss their options. They also included personnel from their legal department to help them be wary of any legal issues this could cause. My interviewee said several possibilities were thrown around as to how to deal with the situation. They ended up choosing a few options.

The first thing they did was notify the client of the privacy concern and what they intended to do about it. They chose to encrypt a mask onto the drone footage so that the images could not be made easily decipherable to the human eye but were readable by the computer. The team also aggregated the location data of the images to census tract level for use outside of the team and client, including for internal presentations. The true location data was put into a password protected folder that only the team members and client had access to. Finally, the legal team drew up some documents that bound the responsibility for data privacy to the client with the exception of the time period of the project.

The interviewee noted that she felt uneasy at first because she was fresh in her career and wasn’t sure how seriously the concerns would be taken. However, she was relieved when her boss immediately responded positively. The client was annoyed at first that they wanted to make a big deal out of this since it meant they had to bill more hours. However, once the legal concerns were made clear and that Energy Consulting might actually be saving them legal fees in the future, the annoyance was mostly resolved. The interviewee said she was happy with the outcome of the issue, but it didn’t make the stress of the process any easier.

# Second Issue

At PetroChemical Inc., just a few months ago, she was asked to prepare production forecasts for a set of newly drilled wells. Some people in the executive suite were planning on presenting this to a potential investor of importance. The VP of Finance had reached out to her and suggested only showing the upper bound predictions to highlight the best-case scenario. She was immediately uncomfortable with the idea because she knew she wouldn’t want someone to do that to her.

In this case, she was a team leader and had more autonomy to make decisions. However, she still reached out to her manager and asked for advice on the situation. The two of them agreed on the importance of representing her team’s findings accurately and avoiding misleading figures. She ultimately decided to include mentions of the best case and worst case scenarios but focus the majority of the figures on the most likely scenario, which fell between the best and worst cases.

The VP didn’t comment on her contribution to the presentation in relation to his previous request but overall did not seem unhappy with the results. This is probably because the most likely scenario was still pretty positive and the investor ended up making a deal with her company. Although my interviewee found this resolution a little less than satisfactory, she said it was the least dramatic and most ethical solution to the issue. My interviewee said she doesn’t like confrontation and definitely didn’t want to put her job on the line, so she felt uncomfortable with addressing the issue head on with anyone in the executive suite but also didn’t want to do something she felt wasn’t right. She says that, looking back, the VP probably said the initial request as just a comment, and not something he really thought all that hard about.

Reflecting on this experience, still fresh in her mind, she said it has definitely helped inform her own leadership approach and how she interacts with her team. She’s found that clarity and intention are useful for avoiding uncomfortable situations like that one and getting the best work out of her team.

# Reflection

I found the first issue surprising, most in that I would have never guessed that personal privacy could be at risk from satellite imagery. I know there used to be concerns during the Cold War about the Soviet Union spying on Americans with satellites, but I never considered how true that could actually be. I hope that I will have the wherewithal to identify a problem like that in the future like my interviewee. I am impressed with her courage and awareness in her early career. It seems that the issue was taken care of quickly and effectively. Even the client’s annoyance is understandable, but they seemed fairly reasonable throughout the whole process from what my interviewee said.

The issue she faced more recently at PetroChemical Inc. is not as surprising because I have faced the exact same issue at a previous job. It seems the issue of wanting to manipulate numbers to make your organization look as good as possible goes across industry lines. In my experience, I ignored pressure from coworkers and managers to misrepresent our data and simply included true numbers in my reporting. When giving information to board members and applying for grants, I highlighted the numbers that were best (but true) and made sure that explanations were given for underperforming areas (including how they could help in their respective positions). I think the way my interviewee handled the situation was quite elegant. Not only did she include all the relevant information, but she made sure that the most likely scenario was the main focus, which seems the best way to present predictive data to anyone, including investors.