

## 2020 LM Ethics in Engineering Competition Case



With some friends from college, Eduardo Guadalupe started ResQ Inc. to bring Artificial Intelligence (AI) and Machine Learning (ML) technology to support humanitarian disaster relief. ResQ's vision is to "Rescue the World," which has been an attractive draw for young engineers to join the company.

ResQ's first product, GRID, is a Quick Reaction Capability (QRC) system for disaster relief search and rescue (SAR) missions. ResQ markets GRID's capabilities to save lives while significantly reducing the financial and personnel strain on non-governmental organizations (NGO) and government relief organizations.

At the heart of GRID is an advanced AI and ML software algorithm that uses large-scale data analytics and situational awareness of both live and recorded data to define rescue priorities, and then develop real-time complex mission rescue plans as natural disasters unfold. GRID uses airborne UAVs to fly over disaster regions to collect data, assess damaged areas, identify people in need, and then develop a rescue strategy involving multiple platforms simultaneously. GRID's open-system architecture integrates with its customers' land, air, and sea resources to carry out SAR missions.

The cornerstone of GRID is its ability to use social media, crowd sourcing, government databases, and collected live data to gather information to best identify and analyze the most impacted areas to determine priorities for the most rapid, effective, and impartial rescue mission.

GRID uses ML to generate a growing database of information from different scenarios and events to more precisely direct responses. Thanks to its numerous successful US pilot programs to date, the system has been trained with years of data, continuously improving itself to identify highly accurate patterns in different disaster relief situations.

ResQ's demonstrated success in the US has resulted in strong international interest for GRID. ResQ's European business development teams are in final negotiations with three large European Union (EU) countries, with an option for full EU deployment.

With the business now expanding to other countries and the growth of AI and ML across innovative industries, ResQ established an ethics board to help govern the development of new products.



An undisclosed Asian-Pacific country (UAP) has expressed strong interest in a complete GRID system. In pursuit of a potential major contract, ResQ deployed a prototype system with the mutual understanding that if all tests were successfully passed, UAP would purchase a full GRID system.

In the contract negotiations UAP identified a risk with governing export-control laws and requested the ability to modify the input data parameters and data storage methodology of the software to tailor the platform to their specific geographical location, natural disasters, and country's needs. UAP highlighted to ResQ that its own country's social media platform would work in parallel with GRID to help expedite SAR missions and mitigate the risk in data sharing. UAP informed ResQ that they wouldn't finalize a contract without this capability.

ResQ's leadership put the challenge to the engineering team. They found a way to partition a customer's proprietary data (which is encrypted on ResQ's servers) from the rest of the datasets, allowing a customer to tailor their needs while benefitting from the rest of ResQ's huge database. Additionally, ResQ added an interface to allow the customer to modify the social media platform data sourcing implementation. UAP stated that the change would help aid in data collection and rescue strategy development. With this requirement met, UAP entered into the contract with ResQ.

During in-country testing, the system had an unexpected deviation in its rescue strategy and prioritization. On the final set of tests, GRID continually failed to allocate sufficient rescue resources to a geographically-specific group of individuals. To debug this issue, the engineering team moved the location of this group to an area that was always included in the rescue strategy in all previous tests, but the group was still excluded from the mission plan. ResQ quickly called off the demonstration to attempt to minimize any concerns, citing that the system had a small bug that needed to be resolved.

ResQ's engineering team said that to truly identify whether the errors are a systematic problem or simply a coincidence, they would need to analyze the data going into the system. However, UAP refused to provide the data. Instead, UAP's engineers said that the failure was only a coincidental anomaly, and they would accept the system as is. In fact, UAP was so anxious for full deployment they informed ResQ that any modifications to the system that was tested in-country would be rejected, and UAP would deem ResQ in breach of its contract and subject to significant penalties.

While the business development team was working through these issues with UAP, back at headquarters ResQ initiated an internal Root Cause Analysis (RCA) to determine what caused the unexpected issue with the algorithm. Jack Jonas, the lead software engineer, strongly advocated against deploying GRID until the deviated behavior had been fully solutioned. Jack theorized that the original algorithms and ML framework were developed, tested, and proven using the extensive data collected in US-based missions. As a result, the system could have

implemented a bias towards “Western” cultures and environments which led to the deviation in behavior in UAP.

Nicole Nickels, the Engineering Project Manager (EPM), pushed back and stated that the issue isn’t the algorithm, but rather biased data entering the system from the country’s social media and information systems, which was intentionally causing the system to not prioritize the individuals in the rescue strategy.

Shari Samson, the AI Subject Matter Expert (SME) for ResQ, stated that this small deviation in behavior is simply due to the fact that the US-based system was extensively trained over time using a bottoms-up ML approach and that due to the data partition agreed upon in the contract, initial deployment of the system in a foreign country would need time until it had conducted enough missions to learn and correct itself.

Due to the lack of system output data from the testing, these three experts could not conclusively decide on the formal cause of the problem. When they presented their analyses to ResQ’s executive leadership team and the ethics board, there was strong support for Shari’s claim based on her years of experience and personal credibility. They dismissed the possibility that there could be a cultural bias in the system’s algorithm, calling it an unsubstantiated accusation against the product. Word of a cultural bias would create a public-relations nightmare that could lead to grounding all GRID systems, putting the US at risk if a natural disaster occurs.

Additionally, the ethics board, contracts, and legal all dismissed Nicole’s theory. The data entering the system is not the responsibility of ResQ and that the system and company are legally compliant with US laws and regulations, satisfying all the requirements of the system.

ResQ went ahead and agreed to UAP’s acceptance criteria. Following Shari’s recommendation, ResQ immediately deployed GRID to begin teaching the system to quickly correct the deviated behavior.

Soon after deployment, a major cyclone hit UAP, causing significant and widespread damage. Within 24 hours, UAP’s news service reported that GRID worked perfectly, and casualties were minimal.

However, independent news sources discovered that many heavily impacted areas with large non-indigenous populations did have high casualty rates, despite GRID being deployed in those areas. The Western media called this a failed rescue due to unjust bias against these residents.

Upon hearing the reports coming out of UAP, the European customers froze negotiations, demanding clarification as to why ResQ would permit racial profiling and other bias in its GRID system. These actions prompted officials from the EU to contact ResQ with the warning that if GRID violates EU Anti-Discrimination Laws, ResQ would be precluded from doing business within the EU.

Eduardo Guadalupe does not know what to do. He is not sure how to proceed with UAP, as well as with the European prospects. ResQ's ethics board has been unable to come to a consensus.

Eduardo contacts your consulting firm to provide an unbiased recommendation on the situation. Your team is tasked with analyzing the ethical, engineering, and business issues at stake. ResQ is seeking a clear path forward that will continue to keep its business profitable and its values intact.

Due to security requirements, your team will not get access to GRID's proprietary intellectual property during your review. Eduardo has asked that you state any technical assumptions you have made in developing your recommendations.

