1. **What technical skills would you bring to your team?**

We would like responses to focus on expertise relating to methodology and coding frameworks i.e your experience in specific programming languages (e.g. Python, C++, Julia) or with types of modelling (e.g. time series models, Bayesian statistics) and frameworks (e.g. Tensorflow, Stan, R-INLA). If you have specific skills related to a challenge you're interested in, do mention them.

As a Data Science student, I have been building up my skills in formal mathematics, probability theory, and statistics through coursework at Harvard University and the Massachusetts Institute of Technology. These are complemented well with my undergraduate education in computer science, and professional software development experience. I am most comfortable using Python and Java for my programming needs, and am also proficient in using Cython, R, C, git, shell scripts, and Javascript.

My current research focuses on counterfactual explanations, and algorithmic recourse. I thus have extensive experience using publicly available toolkits for generating counterfactual explanations, and have also worked on custom software extending and building upon these approaches. I have been studying the ability of counterfactual explanations to improve the understanding of decision-makers of model behaviour as a whole, before deployment, as opposed to explaining individual predictions made by machine learning models on single test data points after deployment.

I recently presented a small portion of this work ("Recourse for Humans") at the Participatory Approaches to ML workshop (ICML 2020), concerning the inclusion of human feedback in order to generate context-appropriate counterfactual explanations. Previously, I have also worked on computational social science, network analysis, and influence maximisation problems.

1. **Please provide a link (or links) to evidence the technical skills stated in the previous question.**

i.e. if you have written code that is openly available on GitHub, add the link to the repository here. If you have a paper that showcases relevant skills, include a link, ideally in conjunction with the corresponding code. We are particularly interested in projects undertaken on real-world data, especially if the raw data was messy. You do not need to provide a link for every aspect of your earlier response, but without links, we are unable to verify your answer. If you prefer to send a private dropbox link, we assure you it will only be used to assess your coding skills and the content will not be made available to anyone outside of the Data Study Group executive team.

All application materials have been collated and can be found at this [github repository README](https://github.com/kaivalyar/DSG/). Some papers are still under double-blind review, and are shared as a compressed zip-file, which requires a password to open. The password is: “\*\*\*\*\*\*\*” (the word “\*\*\*\*\*\*\*”, in lowercase, without quotation marks). This has been verified to unzip correctly on Linux machines, but should work elsewhere too. It is requested that these works not be shared.

1. **What experience do you have of working in a collaborative environment?**

This question is attempting to capture everything outside of the technical skills question above. Possible responses might include examples of working in teams across disciplines/skill levels/sectors etc., project management (using GitHub/GitLab is a plus), or experience in writing and in team leadership.

Perhaps you have an example of where you've helped someone or a team to achieve more. You may have URLs for evidence of open source collaboration or written summaries of research for general or specialised audiences, scientific communication via social media. Maybe you can point to experience participating in, teaching or coordinating similar events. You can be creative here.

I have experience both in participating and in leading teams working on educational projects and software development projects. In college, I set up a student group called the Open Source Development lab, initiating the creation of a series of introductory web tutorials teaching fundamental programming concepts to promote open-source software development. I led the team designing the curriculum, recording the tutorials, and publishing them online, culminating in weekly viewership of ~500 students (around 80%) of the target class of students. Finally, we were even featured on the university’s introductory programming course website.

I am also comfortable working in teams writing code to build software and analyse data. I have worked with researchers at the Institute for Artificial Intelligence, University of Bremen, optimising the pracmln library, which performs statistical relational learning using Markov logic networks. Apart from bringing a 30% speedup to the code, I used my experience working on this project to learn and share my knowledge publicly with the Python User Group in Singapore. In the following weeks, this led to an extremely productive and beneficial exchange wherein I resolved Cython queries, and in turn learnt about other parts of the Python ecosystem from the group.

1. **What do you hope to personally achieve through your participation in the Data Study Group?**

One our goals at The Alan Turing Institute is to train and support the next generation of data scientists. The Data Study Groups are an excellent opportunity for you to apply your skills to real world problems, learn new data science techniques, and improve your problem solving skills in a collaborative environment. Use this space to describe what you hope to achieve through your participation, and/or what you hope to contribute.

Over the course of my research, I have extensively studied fairness and interpretability in machine learning, both from a technical and a social perspective. I have had the fortune of studying the legal aspects of machine learning interpretability from the Turing institute’s Prof. Sandra Wachter, and have been particularly fascinated by her seminal paper “Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR”, introducing counterfactual explanations and algorithmic recourse. While I have been able to study counterfactuals in an academic context, I would like to understand and witness their use in a real world business setting.

My recent research focuses on the use of counterfactual explanations and the incorporation of human feedback to engender trust in machine learning models. I am thus particularly interested in verifying whether these ideas actually hold water with real consumers and businesses, outside of a theoretical setting. I am personally invested in seeing the fruits of my academic labour translate into tangible real world impact, and believe my experience in studying and implementing various counterfactual explanation techniques will bring immense value to the catsAi project team.

1. **Please describe a time that you used your interpersonal skills to support a team.**

When I first arrived in the United States, I had to work with a team to build a music recommender system project as part of a course at university. This was especially challenging since all the team members were of different nationalities, age groups, and possessed different technical proficiencies. There was also significant disagreement among the team members about the best methods to use for such a project, and on appropriately designating individual components.

Since we were all students studying data science, and working on a data science project, I recommended at this stage that we make future decisions based not on hidden, personal preferences, but on shared data about our individual skills. Thus, we were soon able to identify common and distinctive proficiencies among the team members. We then made a list of all the tasks required for successful project completion, and their corresponding skills, and apportioned the work such that each task had one proficient, and one beginner member assigned to it. This way, we balanced the pedagogical and result-oriented needs of the class, and successfully completed the project.