Миша, привет. Я решил всерьез написать тебе про мои текущие карьерные планы, чтобы посоветоваться.

Начну с близких планов. Сейчас я прохожу специализацию по алгоритмам от Тима Рафгардена на курсере; правда, я немного застрял на третьем курсе из четырех, но к концу мая я планирую его пройти. После этого я бы хотел начать Advanced Machine Learning от ВШЭ и Data Science Specialization от John Hopkins University. Не проходил ли ты их? Считаешь ли ты их качественными и хорошими? Также я думаю, что прошел бы специализацию по биоинформатике от университета San Diego, потому что я нахожу механизм транскрипции-трансляции ДНК очень красивым и хочу узнать о нем и о других подобных процессах подробнее. Может, где-то пригодится.

Как ты наверняка знаешь, наш любимый карьерный гид 80000hours.org [советует](https://80000hours.org/career-reviews/) начать карьеру AI Safety researcher с получения PhD в области машинного обучения. В дальнейшем, по-видимому, будет работа где-то в MIRI, FHI, CHAI, DeepMind, OpenAI… Если у меня есть именно такая цель, то где бы ты посоветовал мне пройти магистратуру и куда пойти потом? Я знаю, что в Сколтехе есть магистратура Data Science, по описанию она мне нравится (особенно тем, что там фокус именно на технологии ИИ, а не на работу с большими данными). Другой вариант – МФТИ, там очень много магистерских программ по теме, и туда будет легко попасть, если стать победителем в олимпиаде «Я - Профессионал». Я знаю, что в Сколтехе учится Саша Лыжов. Какое отношение с магистратурой у тебя, я не знаю. Считаешь ли ты, что, быть может, было бы лучше искать что-то вне России? Где, по твоему мнению, получать PhD? Про PhD в computer science я сегодня интересный текст прочитал с сайта 80К часов. Может быть, ты видел. Звучит как по мне.

Embedded Agency читать я начал. Я заметил, что, чем больше читать про AI Safety, тем становится интереснее. Я пока осилил только первую часть. У меня очень мало знаний в этих областях (как и у всех). Стал читать текст [An Introduction to Löb’s Theorem in MIRI Research](https://intelligence.org/files/lob-notes-IAFF.pdf). Этот текст мне не понравился, потому что он чересчур обзорный и нестрогий. Потом я прочитал One Decade of Universal Artificial Intelligence. Вот это мне очень понравилось, хотя строгих определений дано не было.

Кажется, что я бы хотел строить свой диплом вокруг текста Embedded Agency или какой-то другой статьи на похожую тематику, которая меня еще больше заинтересует. Написал об этом своему научному руководителю. Я уверен, за лето я больше буду разбираться в поле и лучше понимать свои интересы. Сейчас я читаю книгу Ника Бострома Superintelligence, хотя ты и Алексей Подгайнов мне ее не советовали.

Помимо ссылок на источники с веб-страницы, где лежит Embedded Agency, я на 80K часов нашел этот [список](https://humancompatible.ai/bibliography) рекомендованных материалов. Всего очень много. В ближайшие месяцы (годы?) я буду читать материалы из него и из Embedded Agency.

Во время магистратуры я хочу посетить MLSS, [AI Safety Camp](https://aisafetycamp.com/), школу [Deep Bayes](https://deepbayes.ru/) , [воркшоп](https://intelligence.org/ai-risk-for-computer-scientists/) MIRI. Что ты сам думаешь про эти места? Знаешь еще какие-то похожие? Также необходимо уже к концу магистратуры иметь какие-то публикации, чтобы взяли на PhD с большей вероятностью. Ты уже писал сам что-нибудь?

Сейчас я готовлюсь к поступлению в ШАД и к собеседованию на летнюю стажировку в Яндекс. Стажировка в Тинькофф Банке 17 мая заканчивается. Я там строил рекомендательную систему. В будущем я заниматься этим не хочу, потому что, на мой взгляд, они решают проблему, которая, не будучи решенной, не наносит никому вреда.

Есть ли у тебя какие-то победы на каггле, в хакатонах? Участвовал ли ты в студенческих олимпиадах («Я – Профессионал»)?