



Фото предоставлено Т.Базаровой

“
Сегодня любой пользователь со своего домашнего компьютера может войти на сайт «Digital Петр», загрузить туда цифровую фотографию или отсканированную копию автографа и прочесть то, что собственноручно написал первый российский император.

цифровой копии и компьютерном наборе присвоили номер. Затем загрузили их в web-приложение Computer Vision Annotation Tool (CVAT) - здесь была выполнена их построчная разметка, соответствующая компьютерному набору. Д.Димитров и М.Потанин применили подход, основанный на глубоком обучении и использовании нейронных сетей. Они определяют и запоминают особенности написания разных букв, что позволяет «настроить» параметры сетей. Так искусственный интеллект научился расшифровывать почерк царя. Созданный для машинного чтения автографов Петра Великого алгоритм безошибочно распознает свыше 97% текста.

Правда, пока искусственному интеллекту удается не все. Скажем, если часть слова оказывается на другой строке, то обе его части он воспринимает как разные слова. Нужна помощь человека, чтобы внести необходимые коррективы и расставить знаки препинания.

Например, первый слог слова «самому» написан в конце первой строки, а два последних - в начале второй. Искусственный интеллект распознал «самому» как два разных слова.

И все же результат очень хороший. Приступая к работе, мы такого эффекта не ожидали. Безусловно, опытный историк, много лет работающий с почерком Петра, справился бы с этим лучше, но для специалиста, впервые взявшегося за его прочтение, помощь, безусловно, неоценимая. Замечу, что и скорость у машины намного выше нашей. А что касается погрешностей расшифровки, то допустить их могут и машина, и человек. Поэтому важно, что теперь появилась возможность сравнивать и проверять качество прочтения. Двойной контроль бывает очень полезен. Сегодня любой пользователь со своего домашнего компьютера может войти на сайт «Digital Петр» (<https://sber.ru/digital-petr>), загрузить туда цифровую фотографию или отсканированную копию автографа и прочесть то, что собственноручно написал первый российский император. И это только начало. Мы рассчитываем продолжить работу по распознаванию скорописи XVII-XVIII веков. На очереди письма сподвижников Петра I: Федора Головина, Петра Шафирова и других (многие из них обладали еще более сложным для прочтения почерком).

Сейчас по инициативе и при поддержке Российского исторического общества СПбИИИ начинает новый проект - создание цифровой библиотеки автографов Петра I (peterscript.historyrussia.org). Появится возможность ознакомиться с оригиналами писем царя-реформатора, их публикациями, которые можно сравнить с расшифровками искусственного интеллекта, а также узнать, при каких обстоятельствах был написан тот или иной автограф.

Отмечу, что искусственный интеллект способен обучаться и совершенствоваться. В частности, он учится распознавать не только отдельные буквы, но и слова, и делать это намного быстрее, чем сейчас. Благодаря его помощи перед научным сообществом открываются огромные перспективы. ■

В преддверии 350-летия со дня рождения царя-реформатора Российское историческое общество и ПАО «Сбербанк» предложили «поручить» искусственному интеллекту разобрать петровский почерк. И тем самым сделать доступным эпистолярное наследие царя для историков и всех интересующихся этой неординарной личностью, а также эпохой его правления. Так появился научно-исследовательский проект «Автографы Петра Великого: чтение технологиями искусственного интеллекта». За его осуществление взялись специалисты Сбербанка по анализу данных во главе с Д.Димитровым и М.Потаниным, а также ученые Санкт-Петербургского института истории РАН, в Научно-историческом архиве которого хранятся копии автографов Петра I. Было это в июне 2020 года, и уже осенью появились первые удачные результаты машинного прочтения письменного наследия Петра. Рассказывает один из ведущих участников проекта, кандидат исторических наук Татьяна БАЗАРОВА (СПБИИ РАН) (на снимке).

За рубежом уже созданы и размещены в Интернете программы, позволяющие читать западноевропейские рукописи эпохи Средневековья и Нового времени, но у нас ничего подобного до настоящего времени не было, а потребность огромная. Архивы, как известно, проводят массовую оцифровку документов, стараясь сделать их доступными для широкого круга всех интересующихся историей страны. Чтобы обработать колоссальный объем информации, и потребовалась помощь искусственного интеллекта. Айтишники Сбербанка создали программу машинного прочтения рукописей Петра. Сначала они разработали методику передачи текста автографов для компьютерной обработки, добившись точной передачи букв царской скорописи. Каждой строке в

XVIII века, делают это довольно быстро. Историкам и сегодня приходится заново изучать петровские автографы - так появляется возможность сравнить копии с оригиналом, оценить качество расшифровки, а иногда и исправить ошибки.

Большинство писем царя прочитали еще в конце XIX века, когда созданная по инициативе академика Афанасия Бычкова и министра народного просвещения Дмитрия Толстого Комиссия по изданию писем и бумаг Петра Великого приступила к изучению наследия первого российского императора. Была скопирована основная масса материалов - письма, указы - так началась подготовка к изданию многотомного труда. В 1887 году вышел из печати первый том «Писем и бумаг императора Петра Великого», в который вошли материалы 1688-1701 годов. Всего до революции успели издать шесть томов. Первый выпуск седьмого вышел в 1918 году, а второй увидел свет только в 1946-м. Труд этот продолжают до сих пор. В 2022 году был опубликован первый выпуск четырнадцатого тома, охватывающий события января-июня 1714 года. И хотя значительная часть уникального проекта выполнена, еще не опубликованы многочисленные материалы, относящиеся к последним десяти годам жизни Петра I.

Из первых рук

Ну-ка, разберись

Искусственный интеллект расшифровывает скоропись Петра I

Юрий ДРИЗЕ

► Похоже, царю-реформатору было не до степенных, вышколенных дяков-писарей, к тому же вопрос: сильно ли он им доверял? Мысли в голове Петра теснились так, что рука за ними не поспевала. Тут уж не до каллиграфии и знаков препинания, да и какими они были в конце XVII - первой четверти XVIII века! Конечно, официальные документы - указы и манифесты первого российского императора - написаны аккуратным писарским почерком и заверены подписью Петра. Заметим, что до него русские государи редко держали перо сами. Скажем, Иван Грозный вел обширную переписку, однако автографов не оставил - царское ли это дело? К тому же в старину русские самодержцы считали унизительным ставить свою подпись под обращением к

адресату: выходило, что тот стоит выше их - властителей всея Руси. Европа этот предрассудок изжила намного раньше, чем Россия, ожидая Петра - человека нового времени. Он легко переступил через давние традиции и сам писал многочисленным своим сподвижникам, а также жене Екатерине Алексеевне, царевичу Алексею, другим членам семьи.

На первый взгляд почерк Петра кажется необыкновенно трудным для прочтения и имеет свои особенности - определенное препятствие для тех, кто впервые с ним сталкивается. Однако, как считают многие историки, изучающие его автографы, это лишь начальное впечатление - все дело в привычке и опыте. Затратив определенные усилия, разобрать царскую скоропись не составляет особого труда. Специалисты, знакомые с традициями письма конца XVII - начала