# Введение

# Эпиграфика Древней Руси XI–XV веков

## - Когда и кем был создан проект?

Разработка методов, инструментов и процедур документирования и визуализации эпиграфических памятников Древней Руси ведется Лабораторией RSSDA с 2017 г. в содружестве с исследователями А. А. Гиппиусом (Лаборатория лингвосемиотических исследований НИУ «Высшая школа экономики») и С. М. Михеевым (Институт славяноведения РАН).

## - В чем цель проекта, его миссия? Кто его аудитория?

Результаты работы представлены на сайте, каждый может их посмотреть, но основная целевая аудитория – это люди, увлекающиеся историей.

В рамках проекта документировано 75 поверхностей с древнерусскими надписями на стенах храмов в Великом Новгороде, Переславле-Залесском и Галиче, 2 плиты-вставки в Можайске и 3 памятника монументальной эпиграфики, хранящихся в музейных коллекциях.

## - В чем медиальные особенности этого проекта? (Например, если это интерактивная карта, то каково ее устройство?) Если Вы учитесь на компьютерного разработчика, можете рассказать о технических деталях (какой стек технологий использовался?)

Надписи документируются фотограмметрическим способом на основе цифровых фотографий высокого разрешения. Для получения трехмерной полигональной модели поверхности с граффити необходимо от 100 до 1000 фотоснимков – в зависимости от глубины гравировки, высоты букв, размера поверхности. Прочтение надписей выполняется по трехмерным моделям и полученным на их основе картам высот. При этом используются математические алгоритмы, позволяющие различным образом окрасить модель поверхности с надписью, присвоив отдельным ее частям цвет, зависящий от их геометрии. Успех в прочтении плохо сохранившейся надписи в значительной степени обеспечивается за счет весьма высокой детальности моделирования.

## - На основе каких исторических источников (данных) создан этот проект? Что это за типы источников? В чем их особенности?

Граффити Древней Руси – надписи, процарапанные на стенах соборов и церквей домонгольского времени, – представляют собой наиболее сложный для исследования и публикации вид эпиграфического памятника. Воспроизведение граффити требует особого подхода. В большинстве случаев необходимо документировать комплекс небольших по размеру разновременных надписей на достаточно обширной поверхности. Как правило, граффити, в отличие от монументальных надписей, выполнены на штукатурке или мягком камне непрофессиональными резчиками и подручным инструментом. Буквы надписей нередко малы (иногда менее 3 мм), а глубина гравировки часто не превышает 0,2–0,3 мм. Шесть-десять веков, прошедших со времени появления этих надписей, не лучшим образом сказались на их сохранности. Надписи редко выделяются цветом и обычно видны лишь в боковом свете. Все это затрудняет их прочтение, прорисовку и публикацию.

## - Общее рассуждение: как этот цифровой проект влияет на наше восприятие прошлого? Он делает его сложнее, понятнее, нагляднее, эмоциональнее, интерактивнее etc.? Заставляет мыслить иначе? Каким образом?

## - Поделитесь Вашим личным мнением по поводу этого проекта или проектов: рекомендуете ли Вы его однокурсникам и почему?

## Источники

[https://rssda.su/]

[https://rssda.su/projects/ep-rus/]

# Digital Пётр

## - Когда и кем был создан проект?

На Artificial Intelligence Journey 2020 – ежегодной серии мероприятий, посвящённых искусственному интеллекту, ПАО Сбербанк провел соревнование «Digital Пётр: распознавание рукописей Петра I». На нём участникам предоставили датасеты из разбитых на строки документы, написанные Петром Великим. Задачей участников было создание модели, которая будет распознавать рукописный текст и преобразовывать в цифровой.

Организаторами конкурса выступили Российское историческое общество, фонд «История Отечества» и Сбер. Его участники решали задачу по распознаванию почерка первого российского императора.

## - В чем цель проекта, его миссия? Кто его аудитория?

Посетители сайта могут загрузить рукопись Петра I самостоятельно, а могут воспользоваться функцией расшифровки заранее заготовленных рукописей. Но всё-таки целевая аудитория – это люди, занимающиеся историей периода правления императора.

Зачем это всё? С 1872 по 2020 год было собрано 143 800 документов, в которых оставил след Пётр. Чтобы полностью изучить его личность, нужно многое из этого прочитать и понять. На расшифровку всех документов нужно очень много времени и человеческих ресурсов. Проект призван облегчить эту проблему.

## - В чем медиальные особенности этого проекта? (Например, если это интерактивная карта, то каково ее устройство?) Если Вы учитесь на компьютерного разработчика, можете рассказать о технических деталях (какой стек технологий использовался?)

Используемая искусственная нейронная сеть в определённой степени моделирует то, как сам человек воспринимает и анализирует информацию, поступающую в мозг через зрительную сенсорную систему. Изображение со скорописью сначала подается на вход свёрточным слоям для извлечения полезных признаков. Нейронная сеть сама выучивает характерные черты написания тех или иных букв. Принцип работы нейронной сети: распознает рукописный текст, настраивая нужным образом свои параметры (с помощью метода оптимизации – градиентного спуска). После этого каждый выученный признак (в виде вектора) подается на вход соответствующему рекурентному слою (GRU). Затем всё собирается. Количество настраиваемых параметров используемой нейросети составило 7 948 733.

Точность распознавания около 90%, что больше, чем у общих моделей, например transkribus. При этом время распознавания составляет 35 страниц в минуту, тогда как у историков на расшифровку каждой страницы столь сложного почерка уходит по несколько часов.

Преимущества описанного подхода – гибкость и способность модели быстро дообучаться на новом почерке, не требуя большого количества транскрибированных и размеченных данных: достаточно нескольких страниц с новым почерком. Создание отечественного программного обеспечения для распознавания и транскрибирования автографов Петра I откроет возможности для построения аналогичных моделей чтения русской скорописи XVI – начала XVIII веков, а также письма других эпох.

## - На основе каких исторических источников (данных) создан этот проект? Что это за типы источников? В чем их особенности?

Проект был создан на основе бумажных письменных источников: для их публикации требуется прочтение профессионалами, поэтому расшифровка очень замедлена.

До проекта «Digital Пётр» Российское историческое общество и ПАО Сбербанк создали научно-исследовательский проект «Автографы Петра Великого: Чтение технологиями искусственного интеллекта». В рамках проекта отбирались документы, охватывающие период с 1709 по 1713 год и включающие в себя от 2-3 слов до нескольких страниц, написанных Петром Великим. Это позволило выявить характерные особенности почерка, например, малую вариативность символов (1-2 вида написания), а также сильную неразборчивость надстрочных символов, а букв «кси» (ѯ), «пси» (ѱ), «ферт» (ф) вообще не было обнаружено.

## - Общее рассуждение: как этот цифровой проект влияет на наше восприятие прошлого? Он делает его сложнее, понятнее, нагляднее, эмоциональнее, интерактивнее etc.? Заставляет мыслить иначе? Каким образом?

Проект в большей степени помогает узнать противоречивую личность Петра – великого государственного деятеля, одновременно делает восприятие общей картины и сложнее, и проще.

## - Поделитесь Вашим личным мнением по поводу этого проекта или проектов: рекомендуете ли Вы его однокурсникам и почему?

{Если подготовленный прочитает эти документы, то вот уже он сможет сделать какие-то выводы и донести их массам}

## Источники

[https://historyrussia.org/sobytiya/avtografy-petra-velikogo-i-tekhnologii-iskusstvennogo-intellekta.html]

[https://rg.ru/2023/03/22/budet-li-prodolzhen-proekt-po-rasshifrovke-slozhnogo-pocherka-petra-pervogo.html]

[https://peterscript.historyrussia.org/]

[https://fusionbrain.ai/digital-petr]

# Не вошедший материал

## Рассказ про Лабораторию

- Когда и кем был создан проект?

RSSDA — это Remote Sensing and Spatial Data Analysis или дистанционное зондирование и анализ пространственных данных. У лаборатории 2 основателя: Юрий Свойский и Екатерина Романенко, она включает 16 постоянных и множество участников отдельных проектов, которые работают по всему миру.

- В чем цель проекта, его миссия? Кто его аудитория?

Цель этой деятельности — помочь исследователю создать цифровой образ объекта исследования, а затем «увидеть невидимое» там, где традиционными методами этого достичь невозможно. «Лаборатория RSSDA» внедряют цифровые технологии сбора и анализа пространственных данных в исторические и археологические исследования.

- В чем медиальные особенности этого проекта? (Например, если это интерактивная карта, то каково ее устройство?) Если Вы учитесь на компьютерного разработчика, можете рассказать о технических деталях (какой стек технологий использовался?)

Лаборатория комбинирует технологии лазерного сканирования (как наземного, так и воздушного), фотограмметрической обработки цифровых фотоснимков, спутниковой и классической геодезии. При этом применяются лазерные сканеры, фотоаппараты, GNSS-приемники, тахеометры, лазерные дальномеры, беспилотные летательные аппараты, мощные графические станции и иное измерительное и вычислительное оборудование.

- На основе каких исторических источников (данных) создан этот проект? Что это за типы источников? В чем их особенности?

Проект работает над множеством физических исторических источников:

* древних надписей на твердых носителях (камне, металле, дереве, керамике)
* наскальное искусство
* каменные статуи и изваяния
* картографируемые археологические памятники и ландшафты
* памятники архитектуры
* пещеры

- Общее рассуждение: как этот цифровой проект влияет на наше восприятие прошлого? Он делает его сложнее, понятнее, нагляднее, эмоциональнее, интерактивнее etc.? Заставляет мыслить иначе? Каким образом?

Проект оцифровывает многие источники, а затем позволяет восстанавливать утраченную со временем информацию. Каждый желающий может посмотреть результаты работы лаборатории, поэтому прошлое становиться нагляднее, понятнее и интерактивнее.

- Поделитесь Вашим личным мнением по поводу этого проекта или проектов: рекомендуете ли Вы его однокурсникам и почему?

{Цифровой музей}

## Альтернативная версия рассказа про Digital Пётр

В 1872 году у одного из основателей Императорского Русского Исторического общества зародилась идея создать единый комплекс источников с рукописями Петра I, этот проект был одобрен даже императором Александром II. В течение долгих лет велась кропотливая работа по выявлению документов везде: как в российских архивах, библиотеках, так и в зарубежных собраниях. Комиссии удалось собрать около 15 000 копий документов. В 1887 году увидел свет первый том «Писем и бумаг императора Петра Великого», в который вошли документы за 1688–1701 годы. С тех пор сбор документов продолжается и по сей день.

Но всё это нужно ещё и расшифровать. Российское историческое общество и ПАО Сбербанк стали решать эту проблему и создали научно-исследовательский проект «Автографы Петра Великого: Чтение технологиями искусственного интеллекта». В рамках проекта отбирались документы, охватывающие период с 1709 по 1713 год и включающие в себя от 2-3 слов до нескольких страниц, написанных Петром Великим. Это позволило выявить характерные особенности почерка, например, малую вариативность символов (1-2 вида написания), а также сильную неразборчивость надстрочных символов, а букв «кси» (ѯ), «пси» (ѱ), «ферт» (ф) вообще не было обнаружено.

И вот все полученные данные стали отправной точкой для участников соревнования «Digital Пётр: распознавание рукописей Петра I», проводимого Сбербанком в рамках Artificial Intelligence Journey 2020 – ежегодной серии мероприятий, посвящённых искусственному интеллекту. Участникам были предоставлены датасеты из нарезанных на строки документов. Размер обучающей выборки составил 6196 строк. Тестовой выборки - 3042 строк.

Используемая искусственная нейронная сеть в определённой степени моделирует то, как сам человек воспринимает и анализирует информацию, поступающую в мозг через зрительную сенсорную систему. Изображение со скорописью сначала подается на вход свёрточным слоям для извлечения полезных признаков. Нейронная сеть сама выучивает характерные черты написания тех или иных букв. Принцип работы нейронной сети: распознает рукописный текст, настраивая нужным образом свои параметры (с помощью метода оптимизации – градиентного спуска). После этого каждый выученный признак (в виде вектора) подается на вход соответствующему рекурентному слою (GRU). Затем всё собирается. Количество настраиваемых параметров используемой нейросети составило 7 948 733.

Преимущества описанного подхода – гибкость и способность модели быстро дообучаться на новом почерке, не требуя большого количества транскрибированных и размеченных данных: достаточно нескольких страниц с новым почерком. Создание отечественного программного обеспечения для распознавания и транскрибирования автографов Петра I откроет возможности для построения аналогичных моделей чтения русской скорописи XVI – начала XVIII веков, а также письма других эпох.