

## Авторство в эпоху искусственного интеллекта

Екатерина Волошина (образовательная программа «Фундаментальная и компьютерная лингвистика»)

### Аннотация

В работе обсуждается вопрос о роли автора в эпоху диджитализации. В течение последней половины XX века фигура автора проходила процесс деконструкции, начиная с работы Ролана Барта. Развитие искусственного интеллекта, который смог генерировать в том числе и художественные тексты, ставит по-новому вопрос о том, кто может быть автором.

**Ключевые слова:** авторство, смерть автора, философия искусственного интеллекта

В этой работе мы рассмотрим вопрос об авторстве в современную эпоху. Отношение к автору как источнику текста значительно изменилось за последнее столетие. Сначала в структуралистских работах критиковался подход анализа текстов через личность автора, что в дальнейшем нашло отражение в искусстве постмодернизма. В постмодернистском искусстве фигура автора оказывалась не отдельной личностью, а совокупностью унаследованных традиций. Развитие технологий, в том числе искусственного интеллекта, ставит вопрос об авторстве с другой стороны: кто может быть автором текста сегодня и что значит быть автором?



### 1. Деконструкция фигуры автора

Впервые о пересмотре фигуры автора заговорил Ролан Барт<sup>1</sup>. Барт утверждает, что бессмысленно приписывать конкретному тексту фигура автора как личность. Автор перестает существовать, как только начинается текст. Текст существует в вечности, актуализируется в тот момент, когда он читается и интерпретируется, текст существует тогда, когда уже не существует автор.

Кроме этого, Барт подчеркивает, что сам язык знает автора не как личность, а исключительно как

<sup>1</sup> Барт, Р. Смерть автора // Барт Р. Избранные работы: Семиотика. Поэтика. М., 1994. С. 384-391.

источник речи. В лингвистике автор текста – тот, кто порождает этот текст. К субъекту письма можно реферировать с помощью местоимений первого лица, но это всего лишь индикатором того, что речь идет о субъекте речи. В языке нет средств, чтобы отличать одного автора от другого. Мы можем отличать авторов по стилю письма, но мы не можем ничего сказать об авторе как о личности, основываясь на его письме. Хотя эти утверждения о языке верны по отношению к любому тексту, Барт пишет в первую очередь о художественном тексте. В модернизме, как пишет Барт, авторы намеренно передавали мысли своих персонажей, но не писали от себя (Барт приводит в пример Марселя Пруста). В таком случае интерпретация через личность автора теряет всякий смысл. Автор – источник мыслей, высказанных в тексте, но он не может навязать единственную интерпретацию, так как природа языка способствует множественной интерпретации и выражению смыслов, которых автор не подразумевал изначально. В работе Барта происходит отделение текста от фигуры автора. Как только автор заканчивает текст, текст становится независимым и существует по правилам художественных практик.

Похожие идеи высказывает и Мишель Фуко в лекции «Что такое автор»<sup>2</sup>. Как и Барт, Фуко говорит, что в тексте не существует личности автора, и попытки найти мнение автора в тексте – лишь интерпретации критиков. Фуко утверждает, что фигура автора нужна для классификации дискурсов и объединения их по источнику. Однако, в отличие от Барта, Фуко говорит о любом тексте, для которого важна функция автора: «все дискурсы, наделенные функцией-автор, содержат эту множественность Эго». Мы видим, что идеи об отсутствии личности автора в тексте расширяются на все дискурсы, доступные для такой классификации. По Фуко, автор не только отделен от текста, но и изначально ограничен социумом, в котором автор существует. То, что будет создано автором, уже частично определено культурным кодом общества, в котором живет автор. Дискурс оказывается выше личности и строится по внутренним правилам, которые автор, пишущий текст, сознательно или неосознанно соблюдает. Хоть и создается впечатление, что автору отведена важнейшая роль, так как текст сохраняет его имя, в действительности автор сохраняет только свое имя, но не свою личность в письме и таким образом служит лишь одной из характеристик текста. Автор нужен лишь как конструкт личности, которого каждый читатель представляет по-своему, основываясь на своей интерпретации.

---

<sup>2</sup> Фуко М. Что такое автор? // Фуко М. Воля к истине: по ту сторону знания, власти и сексуальности. М., Касталь, 1996. С. 7-46.

Идеи деконструкции фигуры автора и бесконечности интерпретаций стали одними из основополагающих идей постмодернизма. Невозможно приписать тексту единственное верную авторскую интерпретацию, а значит, текст открыт бесконечным интерпретациям<sup>3</sup>. В постмодернизме автор понимается как пересечение интертекстуальных практик<sup>4</sup> и не имеющий своей собственной воли, так как это его сознание предопределено текстами предыдущих эпох.

Согласно постмодернистским установкам, весь мир можно воспринимать как текст, систему знаков. Он регулируется культурными кодами, вторичными языками. Таким образом, автор в постмодернизме – это не только автор текста, имеющий дело с первичным, естественным языком, это любой создатель, источник чего-либо. Автор устраняется не только из художественного текста или речи, но в целом из каких-либо практик создания.

В современную эпоху взгляд на функцию автора, сформировавшаяся в постмодернизме, меняется в сторону большей волевости субъекта. В дальнейшем многие авторы пересмотрели свои ранние взгляды. Например, Умберто Эко<sup>5</sup> ограничивает свой тезис о бесконечном количестве интерпретаций и говорит, что интерпретации должны соответствовать каким-либо образом произведению, и роль автора – ограничивать круг интерпретаций. Однако другой проблемой оказывается диджитализация<sup>6</sup>. В современном мире тексты часто создаются без авторства (например, тексты Википедии), или же автора сложно атрибутировать к конкретному человеку, так как это лишь цифровой след человека. Кроме этого, не только человек, но и искусственный интеллект может писать тексты – например, нейросети могут писать стихи (проект “Нейролирика” в Школе лингвистики НИУ ВШЭ). Так, искусственный интеллект поднимает вопрос о том, должен ли пишущий субъектом быть волевым или разумным, чтобы исполнять функцию автора.

## 2. Искусственный интеллект

Вопрос об искусственном интеллекте стоит рассмотреть отдельно, в частности об оценке разумности вычислительных машин. История искусственного интеллекта начинается во второй половине XX века, когда в 1956 году в Дартмутском университете проводится первый научный семинар, посвященный искусственному интеллекту. На этом семинаре Джон Маккарти впервые

---

<sup>3</sup> Eco, U. The open work // Harvard University Press. 1989.

<sup>4</sup> Ильин И. П. Постструктурализм. Деконструктивизм. Постмодернизм. М.: Интрада, 1996. Т. 253. №. 2. С. 70.

<sup>5</sup> Eco U. The role of the reader: Explorations in the semiotics of texts. // Indiana University Press – 1979. – Т. 318.

<sup>6</sup> Павлов, А. В. Постпостмодернизм: как социальная и культурная теории объясняют наше время. //М.: Издательский дом “Дело” РАНХиГС, 2019. С. 423.

вводит сам термин “искусственный интеллект”. Тогда же было дано первое определение искусственного интеллекта. Марвин Минский определил искусственный интеллект как процесс выполнения задач, которые могли бы потребовать человеческих интеллектуальных усилий. В наши дни компьютерные модели могут выполнять задачи, которые требовали интеллектуальной работы для человека – например, вычисления или предсказания погоды, – поэтому в том смысле, как его описал Минский, искусственный интеллект уже существует, и его называют слабым искусственным интеллектом. Сильным искусственным интеллектом<sup>7</sup> называют такую компьютерную систему, которая осмысляет себя как мыслящее существо и обладает разумом. Сильный искусственный интеллект не был создан, и одним из главных вопросов, которым занимается философия искусственного интеллекта, является вопрос о возможности существования сильного искусственного интеллекта.

Другой проблемой является вопрос о том, какой алгоритм считать сильным искусственным интеллектом. Многие требования, предъявляемые к искусственному интеллекту, не могут быть измерены какими-либо тестами. Сильный искусственный интеллект должен обладать сознанием и квалиа, то есть чувственным восприятием. Однако наличие или отсутствие таких качеств сложно проверить.

Алан Тьюринг, утверждая, что невозможно узнать, что происходит внутри машины, предлагает оценивать машины только по их поведению<sup>8</sup>. Тьюринг разрабатывает тест, позже названный тестом Тьюринга. Алгоритм должен пройти следующий эксперимент: человек переписывается с другим человеком и компьютером, и по результатам испытуемый должен сказать, кто из его собеседников человек, а кто машина. При этом задача компьютера – обмануть экзаменатора. Если испытуемый регулярно принимает компьютер за человеческого собеседника, то этот компьютер прошел тест.

Джон Серль<sup>9</sup> возражает на этот аргумент, что порождение речи, которая обманет носителей, не может быть доказательством того, что искусственный интеллект обладает сознанием. В качестве примера Серль проводит мысленный эксперимент: человек сидит в комнате, получает китайские иероглифы и правила, по которым он должен написать другие иероглифы. Человек не знает китайского, поэтому иероглифы для него бессмысленны. При этом за некоторое время он

---

<sup>7</sup> Searle J. R. Minds, brains, and programs // Behavioral and brain sciences. 1980. Т. 3. №. 3. С. 417-424.

<sup>8</sup> Turing, A. Computing Machinery and Intelligence // Mind. 1950. LIX (236). С. 433–460.

<sup>9</sup> Searle J. R. Minds, brains, and programs // Behavioral and brain sciences. 1980. – Т. 3. №. 3. С. 417-424.

преуспевает в своей задаче настолько, что может обмануть носителя китайского. Серль утверждает, что этот человек не выучил китайского, и так как мы не можем проверить, что происходит внутри китайской комнаты, мы не можем утверждать на основании поведения, что машины понимают язык, а значит, обладают сознанием.

Аргумент Серля критикует тот факт, что порождение речи приравнивается к мышлению и сознанию. В философии XX века часто появлялась мысль об отождествлении языка и мышления, в частности благодаря лингвоцентризму. Например, Барт<sup>10</sup> утверждает, что язык, а не автор, – истинный источник текста.

Таким образом, сам язык определяет мышление и форму его выражения. Однако сложно утверждать обратное: порождение текстов не означает осознанное мышление, что подтверждает мысленный эксперимент Серля. Значит, если машины могут научиться языку (по крайней мере с точки зрения стороннего наблюдателя), то это не говорит, что они могут мыслить.

### **3. Искусственный интеллект как автор**

На данный момент искусственный интеллект умеет понимать тексты, в некоторых случаях делая это не хуже человека. Например, в тесте Стэнфордского университета предлагалось ответить на вопросы по смыслу текста, причем вопросы не использовали слова из текстов. Некоторые модели превзошли усредненный результат людей. Другим примером служит соревнование GLUE (General Language Understanding Evaluation). Языковая модель справилась со всеми заданиями на понимание языка лучше, чем средний носитель английского<sup>11</sup>. Кроме этого, как уже было упомянуто выше, нейросети могут порождать тексты, в том числе и художественные. Причем тексты, которые пишут нейросети, часто бывают грамматичными. Алгоритм обучения похож на описанный Серлем эксперимент о китайской комнате за тем исключением, что нейросети должны сами восстанавливать правила. Нейронные модели получают на вход языковые данные и правильные ответы, к которым они должны прийти.

Серль пишет, что мышление, приписываемое искусственному интеллекту, – это продукт интерпретации поведения машины человеком. Люди склонны метафоризировать и другие предметы, говоря, например, “холодильник работает”, хотя “работать” первоначально подразумевает волевого субъекта. Вернемся, однако, к пониманию автора, описанное в

---

<sup>10</sup> Барт, Р. Смерть автора // Барт Р. Избранные работы: Семиотика. Поэтика. М., 1994. С. 384-391.

<sup>11</sup> Sun Y. et al. Ernie 2.0: A continual pre-training framework for language understanding // Proceedings of the AAAI Conference on Artificial Intelligence. 2020. Т. 34. №. 05. С. 8968-8975.

разделе 1. Автор – источник текста, но не он задает интерпретацию, и интерпретация в первую очередь зависит от читателя.

Это рассуждение приводит к другому вопросу: кто является автором текста, порожденного нейросетью? Нейросеть обучается на данных, которые дает ей человек. Затем они переводятся в многовекторное пространство и из него - в символы. Что включает в себя создание текста? Язык – это сложно устроенная система знаков, и каждый знак имеет две стороны: означающее и означаемое. Тогда порождение текста – это приписывание означаемому означающее. В случае человека это перевод мысленных образов в знаки. Машины соотносят численные представления о мире, например, о словах с теми же самыми знаками, выражающими слова. В этом процессе уже не занят человек, так как этот процесс проходит как ряд сложных вычислений. Более того, человеческий ум не может представить себе то, как воспринимают машины слова: для машин это вектора в многомерном пространстве, человеческое же восприятие пространства гораздо более ограничено. Таким образом, нельзя сказать, что процессом порождения текста нейросетью руководит человек, так как человек не дает нейросети непосредственно правила. И нейросеть, и человек используют язык как систему знаков, однако множества означаемых у них разные.

Если мы признаем, что нейросеть может считаться таким же автором, как и человек, при этом мы не считаем, что машины способны к мышлению, то у текстов не может быть никакой первичной интерпретации. Личности автора просто может не существовать, а значит, у него нет никакого особенного отношения к тексту или смысла, вкладываемого в текст. Однако, если текст грамматичен, то он все еще интерпретируем, поэтому текст оказывается важнее, чем сам автор. Интерпретации возникают, даже если заведомо известно, что автор не мог вложить никакой смысл в этот текст.

Таким образом, технологии искусственного интеллекта завершили деконструкцию автора. При этом нельзя сказать, что фигура автора не влияет на текст или его восприятие, однако важным оказывается не личность автора, а то, что фигура автора добавляет к тексту: к примеру, если это литературная маска и личность автора неизвестна, эта аура загадочности распространится на произведение.

#### **4. Заключение**

В работе мы рассмотрели процесс деконструкции фигуры автора и вклад искусственного интеллекта в этот процесс. Уже во второй половине XX века автор стал восприниматься не как отдельная личность, чье восприятие мира отразилось в тексте, а как источник этого текста,

который не только не задает единственно верную интерпретацию текста, но и руководим более общими правилами дискурса и культурного кода своего общества и эпохи.

Деконструкция автора завершается в современную эпоху, если принять во внимание, что искусственный интеллект может писать стихи или создавать другие грамматичные тексты. При этом искусственный интеллект не обладает мышлением, поэтому не может вкладывать первоначальный смысл в текст. Таким образом, текст открыт к интерпретации читателя. Фигура автора лишь один из атрибутов текста, а текст оказывается важнее автора, так как тексты интерпретируемые, даже если точно известно, что автор не вкладывал в них никакого смысла.

#### Библиография

- Барт, Р.* Смерть автора //Барт Р. Избранные работы: Семиотика. Поэтика. М., 1994. С. 384-391
- Павлов, А. В.* Постпостмодернизм: как социальная и культурная теории объясняют наше время // М.: Издательский дом “Дело” РАНХиГС, 2019.
- Ильин И. П.* Постструктурализм. Деконструктивизм. Постмодернизм //М.: Интрада. 1996. Т. 253. №. 2.
- Фуко М.* Что такое автор? //Фуко М. Воля к истине: по ту сторону знания, власти и сексуальности. М., Касталь, 1996. С. 7-46.
- Eco U.* The role of the reader: Explorations in the semiotics of texts. // Indiana University Press – 1979. – Т. 318.
- Eco, U.* The open work // Harvard University Press. 1989.
- Searle J. R.* Minds, brains, and programs // Behavioral and brain sciences. 1980. Т. 3. №. 3. С. 417-424.
- Sun Y. et al.* Ernie 2.0: A continual pre-training framework for language understanding // Proceedings of the AAAI Conference on Artificial Intelligence. 2020. Т. 34. №. 05. С. 8968-8975.
- Turing, A.* Computing Machinery and Intelligence // Mind. 1950. LIX (236). С. 433. 60.