**Проблема перевода аббревиатур при помощи ИИ (на примере терминологии беспилотных летательных аппаратов)**

***Сабаева Юлия Сергеевна***

*Студент*

*Российский Университет Транспорта (МИИТ), Институт экономики и финансов, кафедра «Лингвистика», Москва, Россия*

*E-mail: sabaevayulia02@gmail.com*

В наши дни искусственный интеллект активно применяется в исследованиях в области лингвистики. Нейронные сети стали незаменимым помощником в обработке больших данных для дальнейшего анализа, распознавания текстов и разработке машинного перевода. Но несмотря на определенные преимущества, искусственный интеллект (ИИ) обладает рядом недостатков, что способствует искажению поступающей от него информации.

Во-первых, специальные тексты являются слабым местом у генеративных сетей, в результате чего они зачастую не могут справиться с терминологией. К примеру, в области беспилотных летательных аппаратов до сих пор возникает вопрос, являются ли аббревиатуры «БЛА» и «БПЛА» синонимами, означают ли они одно и тоже, ИИ начинает противоречить самому себе, путаясь в понятиях. Однако, ученые предпринимают попытки разобраться с терминологией в области беспилотных летательных аппаратов. Так, например, В.С. Фетисов и его соавторы в своей работе «Беспилотная авиация: терминология, классификация, современное состояние» попытались соотнести русские термины с английскими [2]. Однако, несмотря на попытки упорядочивания терминов в данной области, терминология до сих пор не устоялась, а значит ИИ не сможет получить репрезентативные данные для работы.

Во-вторых, существуют проблемы и в самой терминологии – терминологизация и детерминологизация слова, в результате чего возникают проблемы синонимии, полисемии и омонимии терминов. К примеру, важной частью терминологии являются аббревиатуры и сокращения, которые зачастую либо отсутствуют в базах данных, либо омонимичны аббревиатурам и сокращениям из других областей. Аббревиация является одним из наиболее популярных способов терминообразования. И в работе И. Хухуни были выявлены «две основные тенденции: 1. воспроизводить соответствующие единицы в их исходном виде, сохраняя при этом – когда речь идет о письменном тексте – написание латиницей, 2. создавать их русские эквиваленты» [5]. При этом в терминологии беспилотных летательных аппаратов «наблюдается процесс лексикализации, при котором они приобретают признаки обычного слова. С одной стороны, это может привести к появлению омонимов, с другой – в плане межязыковой и межкультурной коммуникации – к совпадению со словами другого языка, порой приводящего к лингвистическому шоку» [5].

К примеру, аббревиатура «БЛА» (беспилотный летательный аппарат) омонимична аббревиатуре «БЛА» (бета-лактамный антибиотик (биол.) или «БАК» (беспилотный авиационный комлекс), «БАК» (бомбардировочный авиационный корпус), «БАК» (биохимический анализ крови), где первые два термина относятся к области авиации. Также, термин «дрон» является омонимом к имени собственному «Дрон», в результате чего возникают ошибки при переводе с помощью искусственного интеллекта. Более того, термин «дрон», как один из самых известных терминов среди непрофессионалов, не соответствует требованиям к термину. Термин должен быть уникальным и не иметь полисемии, синонимии и омонимии. Гринев-Гриневич С.В. выделяет следующие свойства термина: «специфичность употребления, содержательная точность, дефинированность, независимость от контекста, однозначность, стилистическая нейтральность, конвенциональность и номинативный характер» [3]. Однако, “basically every UAV is a drone…but not every drone is a UAV” [6]. То есть «дрон» довольно широкое понятие, которое может «покрывать и другие типы подобных устройств, например, подводные, т.е. обладать полисемией или межтерминологической омонимией» [5].

В-третьих, существуют проблемы в самом корпусе текстов – неснятая омонимия и недостаток репрезентативных данных. На данный момент корпус дезаимбигуирован лишь на 1,5%, согласно Зобнянину А., а данные представлены минимум за 2019 год [4]. За столько лет многое поменялось, поэтому необходимо проводить активное снятие омонимии в корпусе и загружать новые данные в области беспилотных летательных аппаратов.

В нашем исследовании с помощью метода сплошной выборки был произведен анализ англоязычных и русскоязычных СМИ, где употреблялись термины, связанные с беспилотными летательными аппаратами. Были выявлены слова, которые чаще всего встречались в одних и тех же предложениях с данными терминами. При переводе предложений с терминами в области беспилотных летательных аппаратов в основном используются конкретные слова, поэтому их можно считать ключевыми и опираться при переводе на них. Для аббревиатур таким образом мы сможем определить область использования данной аббревиатуры (медицина или сфера беспилотников), а касаемо термина «дрон», к примеру, можно понять, перед нами именно термин или имя собственное.

Таким образом, машинный перевод не достиг уровня адекватности человеческого перевода, несмотря на облегчение процесса. Чтобы решить эту проблему, необходимо обучение ИИ, а именно необходимо связать термин с ключевыми словами, которые употребляются наиболее часто в одном предложении. Также необходимо постредактура перевода лингвистом-переводчиком, необходимо снятие омонимии в корпусе и загрузка, обработка новых данных. Активное развитие терминологии в сфере беспилотных летательных аппаратов несомненно упростит задачу по обучению ИИ. А составление глоссариев и терминологических словарей поможет разобраться в терминологии в данной области.

**Литература**

1. Ахманова О. С. Словарь лингвистических терминов / О. С. Ахманова. – М.: УРСС, 2005. – 455 c.
2. Беспилотная авиация: терминология, классификация, современное состояние / В. С. Фетисов, Л. М. Неугодникова, В. В. Адамовский, Р. А. Красноперов. – Уфа: ФОТОН, 2014. – 217 c.
3. Гринев-Гриневич С. В. Терминоведение. / С. В. Гринев-Гриневич. – М.: Академия, 2008. – 302 с.
4. Зобнин, А. Как устроен национальный корпус русского языка / А. Зобнин // YouTube: [сайт]. – URL: https://www.youtube.com/watch?v=QMUt9rFY\_nA (дата обращения: 11.02.2024).
5. Хухуни, И. Г. Формирование новой научно-технической терминосистемы (на материале беспилотных летательных аппаратов): специальность: 10.02.19 «Теория языка» диссертация на соискание ученоц степени кандидата флилологических наук / И. Г. Хухуни.; ГОУ ВО МО Московский государственный областной университет. – Мытищи, 2022. – 181 с.
6. Shawn, H. What is the Difference between a Drone, UAF and UAS? / H. Shawn. – Текст: электронный // [сайт]. URL: https://botlink.com/blog/whats-the-difference-between-a-drone-uav-and-uas (дата обращения 11.02.2024)