



УДК 378.147

DOI: <http://dx.doi.org/10.21686/1818-4243-2024-1-9-20>

Е.А. Косова, К.И. Редкокош, П.О. Михеев

Физико-технического института Крымского федерального университета
им. В.И. Вернадского, Симферополь, Россия

Использование нейронной сети для генерирования изображений при обучении студентов разработке альтернативного текста

Цель исследования: разработать и проверить подход к обучению составителей цифрового контента в части создания альтернативного текста, точно описывающего оригинальное изображение, с использованием нейронной сети для генерирования контрольных изображений, реконструируемых по тексту. Отсутствие в веб-ресурсе текстовых описаний к визуальному контенту ограничивает цифровую доступность, особенно для пользователей с нарушением зрения. Для обеспечения доступности каждое информативное изображение должно сопровождаться альтернативным текстом. Известно, что текстовые альтернативы, сгенерированные с помощью автоматических инструментов, уступают по качеству описаниям, выполненным человеком. Следовательно, составитель цифрового контента должен уметь разрабатывать альтернативный текст к изображениям. Выдвинуто предположение, что нейронная сеть, способная генерировать изображения по текстовым описаниям, может выступать в роли инструмента, служащего для проверки релевантности составляемых текстовых альтернатив.

Материалы и методы. Исследование выполнялось в апреле-мае 2023 года. 17 обучающихся бакалавриата изучили требования к разработке текстовых альтернатив, выполнили первичные текстовые описания к трем предложенным фотографиям, а затем откорректировали текст с использованием нейронной сети Kandinsky 2.1 согласно алгоритму: генерирование изображения по описанию; визуальное сравнение полученного изображения с оригиналом; возвращение к редактированию описания или завершение процесса. По первичным и итоговым описаниям исследователи воссоздали изображения с использованием той же нейронной сети. Дальнейшая работа заключалась в оценке качества всех текстовых описаний и сходства всех сгенерированных изображений с оригинальными. Результаты исследования (текстовые описания; оценки, выставленные экспертами;

ссылки на сгенерированные изображения) опубликованы в виде набора данных в репозитории Mendeley Data. Для анализа данных использовали t-тест, корреляцию Пирсона и многомерную регрессию (при заданном уровне значимости $p = 0,05$).

Результаты. Установлено, что средние оценки качества первичных и итоговых текстовых описаний значимо не отличались ($p > 0,05$), также не было выявлено значимых отличий для длины текста ($p > 0,05$). При этом существенно ($p < 0,05$) возросло сходство сгенерированных изображений с оригинальными фотографиями после использования обучающимися нейронной сети. Следовательно, тренировка в нейронной сети способствовала повышению качества (сходства с оригиналом) изображений, сгенерированных по измененным текстовым описаниям, без потери качества описаний. Обнаружено также, что качество итоговых текстовых альтернатив тем выше, чем больше их размер в пределах отведенного лимита, чем лучше и короче первичные описания ($p < 0,05$). Таким образом, лаконичные и точные альтернативные описания к изображениям после тренировки обучающихся в нейронной сети могут быть преобразованы в не менее качественные текстовые альтернативы, релевантность которых повышается за счет добавления в описание деталей сюжета.

Заключение. Нейронные сети для генерирования изображений могут быть применимы в качестве программного инструмента, стимулирующего потенциальных авторов контента к созданию более точного и полного альтернативного текста при сохранении его лаконичности. Представляется важным продолжить исследования, распространив их на изображения других типов, с использованием различных нейронных сетей.

Ключевые слова: цифровая доступность, альтернативный текст, нейронные сети, электронное обучение, цифровые компетенции.

Yekaterina A. Kosova, Kirill I. Redkokosh, Pavel O. Mikheyev

Institute of Physics and Technology, V.I. Vernadsky Crimean Federal University, Simferopol, Russia

Using A Neural Network to Generate Images When Teaching Students to Develop an Alternative Text

The purpose of research. The purpose of the study is to develop and test an approach to training digital content compilers in creating alternative text that accurately describes the original image, using a neural network to generate reference images reconstructed from the text. The lack of textual descriptions of visual content in a web resource limits digital accessibility, especially for users with visual disorders. To ensure accessibility, each informative image should be accompanied by the alternative text. Text alternatives generated by means of automated tools are known to be lower in quality to human-generated descriptions. Therefore, a digital content compiler must be able to develop the alternative text for images. It has been

suggested that a neural network for generating images from text descriptions can act as a tool for checking the relevance of the developed text alternatives.

Materials and methods. The study was carried out in April-May 2023. 17 undergraduate students studied the requirements for developing text alternatives, completed initial text descriptions for three proposed photographs, and then corrected the text using the Kandinsky 2.1 neural network according to the algorithm: generating an image from the description; visual comparison of the resulting image with the original; returning to editing the description or ending the process. Based on the initial and final descriptions, the researchers

reconstructed the images using the same neural network. Further work consisted of assessing the quality of all text descriptions and the similarity of all generated images to the original ones. The results of the study (text descriptions; expert evaluations; links to generated images) were published as a data set in the Mendeley Data repository. The t-test, Pearson correlation and multivariate regression were used to analyze the data (at the specified significance level $p = 0,05$).

Results. It was found that the quality scores of the initial and final text descriptions were not significantly different ($p > 0,05$), and also there were no significant differences for the length of the text ($p > 0,05$). At the same time, the similarity of the generated images and original photographs after students used the neural network has increased considerably ($p < 0,05$). Therefore, training in the neural network contributed to improving the quality (similarity to the original) of images generated from modified text descriptions, without

losing the descriptions' quality. It was also shown that the quality of the final text alternatives was higher the larger their size within the allotted limit, the better and shorter the initial descriptions ($p < 0,05$). Thus, concise and accurate alternative descriptions for images after training students in a neural network can be converted into equally high-quality text alternatives, the relevance of which is increased by adding plot details to the description.

Conclusion. Neural networks generating images can be applied as a software tool to encourage potential content authors to create more accurate and complete alternative text while keeping it concise. It seems important to continue the research by extending it to other types of images and using a variety of neural networks.

Keywords: digital accessibility, alternative text, neural networks, e-learning, digital competencies.

Введение

Цифровая доступность рассматривается как совокупность свойств электронных ресурсов, определяющая пригодность последних для максимально широкого диапазона пользователей (в том числе людей с ограничениями здоровья) с учетом персональных возможностей, потребностей и предпочтений [1]. В частности, электронный ресурс не может считаться доступным, если пользователи, лишенные возможности нормально видеть, не способны воспринимать и понимать размещенный на нем контент [2, 3]. В настоящее время в мире насчитывается около 253 миллионов лиц с глубокими нарушениями зрения (слепотой и слабовидением) [4], что составляет более 3 % от общей популяции. Все эти люди относятся к одной из самых уязвимых категорий с точки зрения цифровой доступности [5-7].

Отсутствие в электронном контенте текстовых описаний к визуальному содержанию провоцирует цифровое неравенство в отношении лиц с нарушением зрения [8]. Тем не менее, по данным оценки цифровой доступности одного миллиона лучших веб-сайтов мира [9], 22,1 % изображений, размещенных на главных страницах, не имеют альтернативного текста. При этом веб-контент с течением времени становится все более графичным. Так, количество изображений на

главных страницах веб-сайтов в 2023 году возросло на 9,1 % по сравнению предыдущим годом [9]. По данным [10], на ноябрь 2023 года 82,1 % всех веб-ресурсов глобальной сети включали изображения в формате PNG, 77,9 % – JPG, 54,0 % – SVG, 32,3 % – в прочих форматах. Каждое такое изображение (за исключением декоративных) должно сопровождаться альтернативным текстом, содержащим вербальную расшифровку визуальной презентации.

Для автоматизации разработки альтернативных текстовых описаний к изображениям используются методы: программного извлечения подходящих текстов из пула текстовых описаний; компьютерного зрения; обработки естественного языка; глубокого обучения на основе нейронных сетей [11-15]. Вместе с тем перечисленные технологии имеют свои недостатки в отношении точности, полноты и корректности (в том числе этической) генерируемых описаний [11-17] по сравнению с текстовыми альтернативами, созданными человеком [14, 15, 17], которые также не всегда совершенны из-за недостаточной компетентности авторов [17, 18].

Создание качественного альтернативного текста предполагает участие обученных людей [14, 17], которые знают принципы разработки текстовых описаний и умеют использовать программные

инструменты и шаблоны для оптимизации своей деятельности [14, 19, 20]. Навыки разработки альтернативных текстовых описаний к изображениям входят в состав компетенций цифровой доступности как отдельный элемент [1].

Таким образом, обучение составителей цифрового контента (в том числе веб-программистов и веб-дизайнеров) методам разработки альтернативного текста, а также поиск новых программных решений, оптимизирующих создание текстовых описаний, относятся к актуальным задачам.

Согласно нашей гипотезе, нейронные сети, основанные на алгоритмах генерирования изображений по текстовым описаниям (например, Midjourney [21], DALL-E [22], Kandinsky [23]), могут быть использованы как вспомогательный инструмент в обучении, служащий для проверки релевантности составляемого альтернативного текста посредством сравнительной оценки сходства сгенерированных и оригинальных изображений с целью дальнейшего его улучшения. Однако анализ литературы не обнаружил работ, подтверждающих или опровергающих это предположение.

Цель настоящего исследования — разработать и проверить подход к обучению составителей цифрового контента в части создания альтернативного текста, точно описывающего оригинальное изображение, с использованием нейронной

сети для генерирования контрольных изображений, реконструируемых по тексту.

Текстовые альтернативы к изображениям

Альтернативный текст (англ. alternative (alt) text) представляет собой текстовый эквивалент нетекстового контента (например, изображений), размещенного в цифровом документе или на веб-странице [24]. В литературе термины «альтернативный (замещающий) текст», «текстовое описание», «текстовые альтернативы» используются как взаимозаменяемые [17]. Текстовое описание озвучивается программами экранного доступа или воспроизводится в формате Брайля, что позволяет понять нетекстовый контент людям с глубокими нарушениями зрения и некоторыми когнитивными ограничениями. Кроме того, альтернативный текст отображается на веб-странице в режиме отключения изображений и улучшает индексацию веб-ресурса в поисковых системах [25].

Положение «1.1. Текстовые альтернативы» Руководящих принципов доступности веб-контента (англ. Web Content Accessibility Guidelines, WCAG) [2, 3] требует предоставления текстового эквивалента любому нетекстовому содержимому цифрового ресурса для обеспечения восприятия и понимания контента, а также преобразования его в доступные для пользователя форматы. Соответствующий проверяемый признак «1.1.1. Нетекстовый контент» имеет уровень соответствия А, устанавливающий наивысший приоритет исполнения. Отсутствие на цифровом ресурсе текстовых альтернатив к изображениям означает, что некоторые люди (например, с глубокими нарушениями зрения) не смогут им воспользоваться. Под нетекстовым содержи-

мым понимают не только статические изображения, но и медиа (видео или аудио), элементы управления, интерактивные объекты, симуляции, CAPTCHA (от англ. Completely Automated Public Turing test to tell Computers and Humans Apart – Полностью автоматизированный публичный тест Тьюринга, позволяющий отличить компьютеры от людей) [2, 3].

Создание текстовых альтернатив для изображений поддерживается офисным программным обеспечением (например, функцией «Замещающий текст» во всех продуктах Microsoft Office [26]). В веб-ресурсах для добавления альтернативного текста к изображению используется атрибут гипертекстовой разметки *alt* тега *img* (* 0,05$). В случаях повторного обращения к нейронной сети для редактирования первичного текста ($n = 42$) сходство с оригиналом изображений, сгенерированных нейронной сетью, также возрастало значительно (оценка сходства первичных изображений – $2,56 \pm 0,72$; оценка сходства итоговых изображений – $3,03 \pm 0,61$; $t = 4,056$; $p < 0,05$).

Таким образом, использование в обучении нейронной сети способствовало повышению качества изображений, сгенерированных по откорректированным текстовым описаниям, без потери качества самих описаний.

С помощью корреляционного анализа установлено

Результаты оценки первичных текстовых описаний и сгенерированных изображений в группах обучающихся, изменявших и не изменявших описания

Results of evaluation of primary text descriptions and generated images in groups of students who changed and did not change the descriptions

Оцениваемые параметры	Оценки в выборке с неизменными описаниями, $a \pm \sigma (n)$	Оценки в выборке с измененными описаниями, $a \pm \sigma (n)$	t -тест независимых выборок, $t (p)$
Качество первичных текстовых описаний, баллы	$3,39 \pm 0,61 (9)$	$3,49 \pm 0,52 (42)$	-0,48 (0,642)
Сходство с оригиналом изображений, сгенерированных по первичным описаниям, баллы	$3,25 \pm 0,59 (9)$	$2,56 \pm 0,72 (42)$	3,07 (0,008)*
Длина первичных текстовых описаний, количество символов	$124,22 \pm 51,75 (9)$	$138,29 \pm 43,61 (42)$	-0,76 (0,464)

Примечание: a – среднее арифметическое, σ – среднеквадратичное отклонение, n – количество описаний, * – значимое отличие ($p < 0,05$).

Note: a – arithmetic mean, σ – standard deviation, n – number of descriptions, * – significant difference ($p < 0,05$).

Таблица 2 (Table 2)

Результаты оценки текстовых описаний и сгенерированных изображений до и после использования обучающимися нейронной сети

Results of evaluating text descriptions and generated images before and after students use the neural network

Оцениваемые параметры	Первичные оценки, $a \pm \sigma$	Итоговые оценки, $a \pm \sigma$	Парный зависимый t -тест, $t(p)$
Качество текстовых описаний, баллы			
– фотография № 1, $n = 17$	$3,50 \pm 0,48$	$3,63 \pm 0,75$	-0,951 (0,356)
– фотография № 2, $n = 17$	$3,35 \pm 0,59$	$3,49 \pm 0,51$	-1,705 (0,108)
– фотография № 3, $n = 17$	$3,57 \pm 0,54$	$3,63 \pm 0,48$	-0,460 (0,651)
– все фотографии, $n = 51$	$3,48 \pm 0,53$	$3,58 \pm 0,58$	-1,612 (0,113)
Сходство сгенерированных изображений и оригиналов, баллы			
– фотография № 1, $n = 17$	$2,34 \pm 0,70$	$2,90 \pm 0,48$	-3,040 (0,008)*
– фотография № 2, $n = 17$	$3,19 \pm 0,79$	$3,53 \pm 0,45$	-1,636 (0,121)
– фотография № 3, $n = 17$	$2,51 \pm 0,44$	$2,79 \pm 0,63$	-2,612 (0,019)*
– все фотографии, $n = 51$	$2,68 \pm 0,75$	$3,07 \pm 0,61$	-3,988 (<0,001)*
Длина текстовых описаний, количество символов			
– фотография № 1, $n = 17$	$137,8 \pm 39,3$	$137,8 \pm 52,9$	0,007 (0,994)
– фотография № 2, $n = 17$	$124,2 \pm 46,9$	$121,7 \pm 43,0$	0,652 (0,524)
– фотография № 3, $n = 17$	$145,4 \pm 48,1$	$135,8 \pm 48,5$	1,228 (0,237)
– все фотографии, $n = 51$	$135,8 \pm 44,9$	$131,8 \pm 47,9$	1,018 (0,314)

Примечание: a – среднее арифметическое, σ – среднеквадратичное отклонение, * – значимые отличия ($p < 0,05$).

Note: a – arithmetic mean, σ – standard deviation, * – significant difference ($p < 0,05$).

(табл. 3), что качество текстового описания оценивалось экспертами тем выше, чем больше его длина в пределах рекомендованного размера ($p < 0,05$), в то время как сходство сгенерированного изображения с оригиналом не зависело от размера соответ-

ствующего альтернативного текста ($p > 0,05$). Эти закономерности справедливы как для первичных, так и для итоговых текстовых описаний, откорректированных после использования обучающимися нейронной сети.

С целью уточнения взаи-

Таблица 3 (Table 3)

Результаты проверки связей между длиной текстовых описаний, оценкой качества описаний и оценкой сходства изображений, $n = 51$
Results of testing the relationships between the length of text descriptions, assessment of the quality of descriptions and assessment of image similarity, $n = 51$

Тип описания	Переменная-фактор	Результирующие переменные	Корреляция Пирсона, $r(p)$
Первичное	Длина текстовых описаний	Оценка качества текстовых описаний	0,460 (0,001)*
		Оценка сходства изображений	0,006 (0,967)
Итоговое		Оценка качества текстовых описаний	0,656 (< 0,001)*
		Оценка сходства изображений	-0,177 (0,214)

Примечание: * – значимые отличия ($p < 0,05$).
Note: * – significant difference ($p < 0,05$).

мосвязей между оценкой итоговых текстовых описаний и прочими переменными-факторами, способными оказать влияние на качество разработанного альтернативного текста, построена модель множественной линейной регрессии: $y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \theta_4 x_4 + \varepsilon$, где y – оценка качества итогового текстового описания, $\Theta = \{\theta_0, \theta_1, \theta_2, \theta_3, \theta_4\}$ – вектор коэффициентов модели, ε – отклонение прогнозируемого значения от фактического значения для выборочной модели, x_1 – оценка качества первичного текстового описания, x_2 – оценка сходства с оригиналом изображения, сгенерированного по первичному текстовому описанию, x_3 – количество символов в первичном текстовом описании, x_4 – количество символов в итоговом текстовом описании.

В результате обучения модели получены следующие значения её коэффициентов: $\theta_0 = 0,986$ ($p = 0,002 < 0,05$); $\theta_1 = 0,614$ ($p = 6,65 \cdot 10^{-8} < 0,05$); $\theta_2 = 0,018$ ($p = 0,763 > 0,05$); $\theta_3 = -0,008$ ($p = 2,43 \cdot 10^{-5} < 0,05$); $\theta_4 = 0,012$ ($p = 2,59 \cdot 10^{-9} < 0,05$). Положительные значения коэффициентов (при $p < 0,05$) свидетельствуют о наличии прямой связи между соответствующими переменными-факторами и результирующей переменной модели, отрицательные – о наличии

обратной связи. На основании полученных данных вычислен коэффициент множественной детерминации $R^2 = 0,739$ и скорректированный коэффициент детерминации для линейной модели $R^2_{adj} = 0,716$ ($p = 7,189 \cdot 10^{-13} < 0,05$). Найденные R^2 и R^2_{adj} принадлежат интервалу (0,7; 0,9), что, согласно шкале R.E. Chaddock [33], соответствует высокой тесноте связи.

Таким образом, имеем значимую ($p < 0,05$) линейную регрессию с коэффициентами детерминации, свидетельствующими о наличии сильной взаимосвязи между переменной, определяющей качество итоговых текстовых описаний, и независимыми переменными модели. Единственной незначимой ($p > 0,05$) переменной является x_2 (оценка сходства сгенерированного и оригинального изображений), остальные – оказывают значимое влияние на оценку качества итоговых текстовых альтернатив, причем самый существенный вклад в объяснение результирующей переменной y дает переменная x_1 (оценка качества первичного текстового описания), имеющая в уравнении линейной регрессии наибольшее по модулю значение коэффициента.

Интерпретировать результаты регрессии можно следующим образом: качество итоговых текстовых описаний тем

выше, чем лучше соответствующие первичные описания и меньше их длина; одновременно, чем длиннее итоговые текстовые описания, тем выше их качество. Другими словами, лаконичные и точные альтернативные описания к изображениям после тренировки обучающихся в нейронной сети могут быть преобразованы в не менее качественные текстовые альтернативы, релевантность которых повышается за счет добавления в описание деталей изображения.

Обсуждение

Исследование применимости нейронных сетей, генерирующих изображения, в обучении составителей текстовых альтернатив к нетекстовому контенту выполнено впервые.

Мы сосредоточились на разработке тестовых описаний информативных изображений человеком, так как методы искусственного интеллекта, машинного обучения и глубокого обучения все еще не совершенны в отношении точности и детализации альтернативного текста [11–17]. В то же время использование некоторых инструментов искусственного интеллекта, в частности нейронных сетей для генерирования изображений, может быть полезно для оттачивания авторами качества созданных ими текстовых описаний.

В результате исследования установлено, что тренировка обучающихся в нейронной сети способствует повышению точности сгенерированных изображений, то есть увеличению их сходства с оригиналом, без ущерба для качества альтернативного текста. При этом результаты не позволяют говорить о статистически значимом улучшении качества итоговых текстовых описаний по сравнению с первичными. Вместе с тем показано, что альтернативный текст значительно улучшается с увеличением

его длины (за счет добавления деталей, поясняющих сюжет), что справедливо как для первичных, так и для итоговых описаний. Последний результат согласуется с опубликованным в работах [14, 17], где полнота (включение в описание всех важных атрибутов изображения) определена как ключевая характеристика качества альтернативного текста.

Результаты регрессионного анализа показали, что использование нейронной сети в качестве инструмента для самооценки является продуктивным. Тренировка в нейронной сети понуждает составителей работать над улучшением текста, сохраняя лаконичность описания (в пределах заданного лимита) и одновременно добавляя в него детали сюжета.

Достаточно высокое качество текстовых описаний (не ниже 3,5 баллов в среднем) может быть связано с тем, что обучающиеся изначально разрабатывали, а затем корректировали собственные текстовые альтернативы, не прибегая к использованию черновых описаний, созданных автоматическими инструментами и другими людьми. Это предположение находит косвенное подтверждение в работе [14], показавшей, что качество альтернативного текста, разработанного составителем на основе автоматически сгенерированного описания, значительно ниже, чем при полностью самостоятельном написании.

Согласно [17], разработка программных решений, оптимизирующих создание человеком альтернативных описаний к изображениям, относится к актуальным задачам. Кроме того, известно, что интерфейсы для разработки текстовых альтернатив, содержащие шаблоны, подсказки и уточняющие вопросы, позволяют создавать более

релевантные описания [14, 16]. В нашем исследовании нейронную сеть для генерирования изображений можно рассматривать как учебный тренажер с визуальным интерфейсом, работа в котором стимулирует авторов к совершенствованию текстовых описаний (подчиненных шаблонным требованиям к содержанию и размеру) путем поступательного улучшения сходства сгенерированных изображений с оригиналом.

Созданию качественных текстовых альтернатив способствует обратная связь от пользователей из целевой аудитории; вместе с тем, оперативно получить такие отзывы не всегда возможно [17]. Нейронная сеть в нашем исследовании выступает в роли программного инструмента для моментальной обратной связи, который позволяет автору самостоятельно оценивать точность и полноту текстового описания по характеристикам сгенерированного изображения.

В настоящем исследовании составители обучались создавать текстовые описания только для информативных изображений. Вместе с тем инструкции по разработке альтернативного текста могут различаться в зависимости от типа и назначения изображения [14, 25]. В перспективе целесообразно исследовать применимость нейронных сетей в обучении составлению текстовых описаний для всего спектра изображений, включая функциональные, сложные, сгруппированные, подвижные изображения, изображения текста, карты изображений.

Научное направление, изучающее возможности генерирования изображений из текста (от англ. text-to-image), развивается, появляются новые модели нейронных сетей, обладающие улучшенными свойствами распознавания, интерпретации и визуализации

текстовых данных [34]. Дальнейшие исследования могут быть связаны с изучением эффекта от использования нейронных сетей, основанных на разных алгоритмах генерирования изображений, в обучении и повседневной практике составителей текстовых альтернатив.

Исследование имеет **ограничения**: не использовались инструменты для автоматизации оценки — сходство изображений и качество текстовых описаний оценивались субъективно; использовалась одна нейронная сеть, работающая с русскоязычными текстами, — не исключено, что применение других нейронных сетей и языков описаний даст иные результаты; изображения, подлежащие оценке, генерировались исследователем — вероятно, изображения, сгенерированные обучающимися на своих компьютерах, выглядели иначе; исследование выполнено на небольшой выборочной совокупности участников и изображений — возможно, увеличение объема исследованных случаев позволит выявить более выраженные статистические закономерности.

Заключение

Наличие в цифровом контенте альтернативного текста к изображениям является одним из признаков доступности электронного ресурса. При этом текстовые альтернативы, полученные с помощью автоматических инструментов, уступают по качеству описаниям, выполненным человеком. Современный автор цифрового контента должен уметь разрабатывать текстовые альтернативы к изображениям, в том числе с помощью программного обеспечения, оптимизирующего такую разработку.

В результате исследования был разработан и проверен подход к обучению авторов цифрового контента составле-

нию текстовых описаний к информативным изображениям с использованием нейронной сети. Установлено, что нейронные сети для генерирования изображений могут быть

применимы в качестве программного инструмента, стимулирующего авторов к созданию более точных и полных описаний при сохранении их лаконичности.

Представляется важным продолжить исследования, распространив их на изображения других типов, с использованием различных нейронных сетей.

Литература

1. Web Content Accessibility Guidelines (WCAG) 2.1. [Электрон. ресурс]. 2018. Режим доступа: <https://www.w3.org/TR/WCAG21/> (Дата обращения: 22.11.2023).
2. Web Content Accessibility Guidelines (WCAG) 2.2. [Электрон. ресурс]. 2023. Режим доступа: <https://www.w3.org/TR/WCAG22/> (Дата обращения: 22.11.2023).
3. World Blind Union [Электрон. ресурс]. 2023. Режим доступа: <https://worldblindunion.org/> (Дата обращения: 22.11.2023).
4. Gill K., Sharma R., Gupta R. Empowering visually impaired students through e-learning at higher education: problems and solutions // IOSR Journal of Humanities and Social Science. 2017. Т. 22. № 8. С. 27–35. DOI: 10.9790/0837-2208072735.
5. Marghalani A. Online courses accessibility for low-vision [Электрон. ресурс] // In 2020 AECT Convention Proceedings. 2020. С. 1–37. Режим доступа: https://members.aect.org/pdf/Proceedings/proceedings20/2020/20_03.pdf (Дата обращения: 22.11.2023).
6. Jung C., Mehta S., Kulkarni A., Zhao Y., Kim Y.-S. Communicating Visualizations without Visuals: Investigation of Visualization Alternative Text for People with Visual Impairments // In IEEE Transactions on Visualization and Computer Graphics. 2022. Т. 28. № 1. С. 1095–1105. DOI: 10.1109/TVCG.2021.3114846.
7. The WebAIM million: Images and alternative text [Электрон. ресурс]. 2023. Режим доступа: <https://webaim.org/projects/million/#alttext> (Дата обращения: 22.11.2023).
8. Usage statistics of image file formats for websites [Электрон. ресурс]. 2023. Режим доступа: https://w3techs.com/technologies/overview/image_format (Дата обращения: 22.11.2023).
9. Sharma H., Agrahari M., Singh S.K., Firoj M., Mishra R.K. Image Captioning: A Comprehensive Survey // In 2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC), Mathura, India. 2020. С. 325–328. DOI: 10.1109/PARC49193.2020.236619.
10. Hanley M., Barocas S., Levy K., Azenkot S., Nissenbaum H. Computer Vision and Conflicting Values: Describing People with Automated Alt Text // In Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES '21). Association for Computing Machinery, New York, NY, USA. 2021. С. 543–554. DOI: 10.1145/3461702.3462620.
11. Lee J., Peng Y.H., Herskovitz J., Guo A. Image Explorer: Multi-Layered Touch Exploration to Make Images Accessible // In Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21). Association for Computing Machinery, New York, NY, USA. 2021. Article 69. С. 1–4. DOI: 10.1145/3441852.3476548.
12. Mack K., Cutrell E., Lee B., Morris M.R. Designing Tools for High-Quality Alt Text Authoring // In Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21). Association for Computing Machinery, New York, NY, USA. 2021. Article 23. С. 1–14. DOI: 10.1145/3441852.3471207.
13. Jeong H., Chun V., Lee H., Oh S.Y., Jung H. WATAA: Web Alternative Text Authoring Assistant for Improving Web Content Accessibility // In Companion Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23 Companion). Association for Computing Machinery, New York, NY, USA. 2023. С. 41–45. DOI: 10.1145/3581754.3584127.
14. Salisbury E., Kamar E., Morris M. Toward Scalable Social Alt Text: Conversational Crowdsourcing as a Tool for Refining Vision-to-Language Technology for the Blind // Proceedings of the AAAI Conference on Human Computation and Crowdsourcing. 2017. № 5(1). С. 147–156. DOI: 10.1609/hcomp.v5i1.13301.
15. Edwards E.J., Gilbert M., Blank E., Branham S.M. How the Alt Text Gets Made: What Roles and Processes of Alt Text Creation Can Teach Us About Inclusive Imagery // ACM Trans. Access. Comput. 2023. № 16(2). С. 1–18. DOI: 10.1145/3587469.
16. Chintalapati S.S., Bragg J., Wang L.L. A Dataset of Alt Texts from HCI Publications: Analyses and Uses Towards Producing More Descriptive Alt Texts of Data Visualizations in Scientific Papers // In Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '22). Association for Computing Machinery, New York, NY, USA. 2022. С. 1–12. DOI: 10.1145/3517428.3544796.
17. Morash V.S., Siu Y.-T., Miele J.A., Hasty L., Landau S. Guiding Novice Web Workers in Making Image Descriptions Using Templates // ACM Trans. Access. Comput. 2015. Т. 7. № 4. С. 1–21. DOI: 10.1145/2764916.

18. Gleason C., Pavel A., Liu X., Carrington P., Chilton L.B., Bigham J.P. Making Memes Accessible // In The 21st International ACM SIGACCESS Conference on Computers and Accessibility (Pittsburgh, PA, USA) (ASSETS '19). Association for Computing Machinery, New York, NY, USA. 2019. С. 367–376. DOI: 10.1145/3308561.3353792.
19. Midjourney [Электрон. ресурс]. 2023. Режим доступа: <https://www.midjourney.com/home?callbackUrl=%2Fexplore> (Дата обращения: 22.11.2023).
20. DALL-E [Электрон. ресурс]. 2023. Режим доступа: <https://openai.com/dall-e-2>
21. Kandinsky [Электрон. ресурс]. 2023. Режим доступа: <https://rudalle.ru/kandinsky2> (Дата обращения: 22.11.2023).
22. Alternative Text [Электрон. ресурс]. 2023. Режим доступа: <https://webaim.org/techniques/alttext/> (Дата обращения: 22.11.2023).
23. Images Tutorial [Электрон. ресурс]. 2022. Режим доступа: <https://www.w3.org/WAI/tutorials/images/> (Дата обращения: 22.11.2023).
24. Добавление замещающего текста к фигуре, картинке, диаграмме, рисунку SmartArt или к другому объекту [Электрон. ресурс]. 2023. Режим доступа: <https://support.microsoft.com/ru-ru/office/83-44989b2a-903c-4d9a-b742-6a75b451c669> (Дата обращения: 22.11.2023).
25. ГОСТ Р 57891-2022. Тифлокомментирование и тифлокомментарий. Термины и определения: дата введения 2022-01-01 / ФГБУ «РСТ», НУ ИПРПП ВОС «Реакомп» [Электрон. ре-

сурс]. 2022. Режим доступа: <https://nd.gostinfo.ru/document/6880129.aspx> (Дата обращения: 15.12.2023).

26. Об установлении порядка обеспечения условий доступности для инвалидов по зрению официальных сайтов государственных органов, органов местного самоуправления и подведомственных организаций в информационно-телекоммуникационной сети «Интернет» [Приказ Министерства цифрового развития, связи и массовых коммуникаций РФ от 12 декабря 2022 г. N 931] [Электрон. ресурс]. 2022. Режим доступа: <https://www.garant.ru/products/ipo/prime/doc/405916637/> (Дата обращения: 15.12.2023).

27. Lunch atop a Skyscraper [Электрон. ресурс]. 2023. Режим доступа: https://en.wikipedia.org/wiki/Lunch_atop_a_Skyscraper (Дата обращения: 15.12.2023).

28. Apollo 11 [Электрон. ресурс]. 2023. Режим доступа: https://en.wikipedia.org/wiki/Apollo_11 (Дата обращения: 15.12.2023).

29. Drysdale J., Regan M. Our Peaceable Kingdom: The photographs of John Drysdale. New York: St. Martin's Press, 2000. 112 с.

30. Chaddock R.E. Principles and methods of statistics. Boston: Houghton Mifflin Company, 1925. 471 с.

31. Maerten A.-S., Soydaner D. From paintbrush to pixel: A review of deep neural networks in AI-generated art. 2023. DOI: 10.48550/arXiv.2302.10913.

References

1. Web Content Accessibility Guidelines (WCAG) 2.1. [Internet]. 2018. Available from: <https://www.w3.org/TR/WCAG21/> (cited 22.11.2023).
2. Web Content Accessibility Guidelines (WCAG) 2.2. [Internet]. 2023. Available from: <https://www.w3.org/TR/WCAG22/> (cited 22.11.2023).
3. World Blind Union [Internet]. 2023. Available from: <https://worldblindunion.org/> (cited 22.11.2023).
4. Gill K., Sharma R., Gupta R. Empowering visually impaired students through e-learning at higher education: problems and solutions. IOSR Journal of Humanities and Social Science. 2017; 22; 8: 27-35. DOI: 10.9790/0837-2208072735.
5. Marghalani A. Online courses accessibility for low-vision [Internet]. In 2020 AECT Convention Proceedings. 2020: 1–37. Available from: https://members.aect.org/pdf/Proceedings/proceedings20/2020/20_03.pdf (cited 22.11.2023).
6. Jung C., Mehta S., Kulkarni A., Zhao Y., Kim Y.-S. Communicating Visualizations without Visuals: Investigation of Visualization Alternative Text for People with Visual Impairments. In IEEE Transactions on Visualization and Computer Graphics. 2022; 28; 1: 1095-1105. DOI: 10.1109/TVCG.2021.3114846.

7. The WebAIM million: Images and alternative text [Internet]. 2023. Available from: <https://webaim.org/projects/million/#alttext> (cited 22.11.2023).

8. Usage statistics of image file formats for websites [Internet]. 2023. Available from: https://w3techs.com/technologies/overview/image_format (cited 22.11.2023).

9. Sharma H., Agrahari M., Singh S.K., Firoj M., Mishra R.K. Image Captioning: A Comprehensive Survey. In 2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC), Mathura, India. 2020: 325-328. DOI: 10.1109/PARC49193.2020.236619.

10. Hanley M., Barocas S., Levy K., Azenkot S., Nissenbaum H. Computer Vision and Conflicting Values: Describing People with Automated Alt Text. In Proceedings of the 2021 AAI/ACM Conference on AI, Ethics, and Society (AI/ES '21). Association for Computing Machinery, New York, NY, USA. 2021: 543–554. DOI: 10.1145/3461702.3462620.

11. Lee J., Peng Y.H., Herskovitz J., Guo A. Image Explorer: Multi-Layered Touch Exploration to Make Images Accessible. In Proceedings of the 23rd International ACM SIGACCESS

- Conference on Computers and Accessibility (ASSETS '21). Association for Computing Machinery, New York, NY, USA. 2021; 69: 1–4. DOI: 10.1145/3441852.3476548.
12. Mack K., Cutrell E., Lee B., Morris M.R. Designing Tools for High-Quality Alt Text Authoring. In Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21). Association for Computing Machinery, New York, NY, USA. 2021; 23: 1–14. DOI: 10.1145/3441852.3471207.
13. Jeong H., Chun V., Lee H., Oh S.Y., Jung H. WATAA: Web Alternative Text Authoring Assistant for Improving Web Content Accessibility // In Companion Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23 Companion). Association for Computing Machinery, New York, NY, USA. 2023: 41–45. DOI: 10.1145/3581754.3584127.
14. Salisbury E., Kamar E., Morris M. Toward Scalable Social Alt Text: Conversational Crowdsourcing as a Tool for Refining Vision-to-Language Technology for the Blind. Proceedings of the AAAI Conference on Human Computation and Crowdsourcing. 2017; 5(1): 147–156. DOI: 10.1609/hcomp.v5i1.13301.
15. Edwards E.J., Gilbert M., Blank E., Branham S.M. How the Alt Text Gets Made: What Roles and Processes of Alt Text Creation Can Teach Us About Inclusive Imagery. ACM Trans. Access. Comput. 2023; 16(2): 1–18. DOI: 10.1145/3587469.
16. Chintalapati S.S., Bragg J., Wang L.L. A Dataset of Alt Texts from HCI Publications: Analyses and Uses Towards Producing More Descriptive Alt Texts of Data Visualizations in Scientific Papers. In Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '22). Association for Computing Machinery, New York, NY, USA. 2022: 1–12. DOI: 10.1145/3517428.3544796.
17. Morash V.S., Siu Y.-T., Miele J.A., Hasty L., Landau S. Guiding Novice Web Workers in Making Image Descriptions Using Templates. ACM Trans. Access. Comput. 2015; 7; 4: 1–21. DOI: 10.1145/2764916.
18. Gleason C., Pavel A., Liu X., Carrington P., Chilton L.B., Bigham J.P. Making Memes Accessible. In The 21st International ACM SIGACCESS Conference on Computers and Accessibility (Pittsburgh, PA, USA) (ASSETS '19). Association for Computing Machinery, New York, NY, USA. 2019: 367–376. DOI: 10.1145/3308561.3353792.
19. Midjourney [Internet]. 2023. Available from: <https://www.midjourney.com/home?callbackUrl=%2Fexplore> (cited 22.11.2023).
20. DALL-E [Internet]. 2023. Available from: <https://openai.com/dall-e-2>
21. Kandinsky [Internet]. 2023. Available from: <https://rudalle.ru/kandinsky2> (cited 22.11.2023). (In Russ.)
22. Alternative Text [Internet]. 2023. Available from: <https://webaim.org/techniques/alttext/> (cited 22.11.2023).
23. Images Tutorial [Internet]. 2022. Available from: <https://www.w3.org/WAI/tutorials/images/> (cited 22.11.2023).
24. Dobavleniye zameshchayushchego teksta k figure, kartinke, diagramme, risunku SmartArt ili k drugomu ob'yektu = Add alt text to a shape, picture, chart, SmartArt, or other object [Internet]. 2023. Available from: <https://support.microsoft.com/ru-ru/office/83-44989b2a-903c-4d9a-b742-6a75b451c669> (cited 22.11.2023). (In Russ.)
25. GOST R 57891-2022. Tiflocommentation and Tiflocommentary. Terms and definitions: date of introduction 2022-01-01 / FGBU “RST”, NU IPRPP VOS “Reacomp” [Internet]. 2022. Available from: <https://nd.gostinfo.ru/document/6880129.aspx> (cited 15.12.2023). (In Russ.)
26. On establishing a procedure for ensuring accessibility conditions for visually impaired persons of official websites of state bodies, local governments and subordinate organizations on the Internet information and telecommunications network [Order of the Ministry of Digital Development, Communications and Mass Communications of the Russian Federation dated December 12, 2022 N 931] [Internet]. 2022. Available from: <https://www.garant.ru/products/ipo/prime/doc/405916637/> (cited 15.12.2023). (In Russ.)
27. Lunch atop a Skyscraper [Internet]. 2023. Available from: https://en.wikipedia.org/wiki/Lunch_atop_a_Skyscraper (cited 15.12.2023).
28. Apollo 11 [Internet]. 2023. Available from: https://en.wikipedia.org/wiki/Apollo_11 (cited 15.12.2023).
29. Drysdale J., Regan M. Our Peaceable Kingdom: The photographs of John Drysdale. New York: St. Martin's Press; 2000. 112 p.
30. Chaddock R.E. Principles and methods of statistics. Boston: Houghton Mifflin Company; 1925. 471 p.
31. Maerten A.-S., Soydaner D. From paintbrush to pixel: A review of deep neural networks in AI-generated art. 2023. DOI: 10.48550/arXiv.2302.10913.

Сведения об авторах

Екатерина Алексеевна Косова

К.п.н., доцент, заведующая кафедрой прикладной математики

Физико-технический институт

Крымский федеральный университет

им. В.И. Вернадского, Симферополь, Россия

Эл. почта: lynx99@inbox.ru

Кирилл Игоревич Редкокош

Аспирант

Крымский федеральный университет

им. В.И. Вернадского, Симферополь, Россия

Эл. почта: kirillf13@yandex.ru

Павел Олегович Михеев

Студент

Крымский федеральный университет

им. В.И. Вернадского, Симферополь, Россия

Эл. почта: pavel0990848502@gmail.com

Information about the authors

Kosova A. Yekaterina

Cand. Sci. (Pedagogical), Associate Professor,
Head of the Department of Applied Mathematics

Institute of Physics and Technology,

V.I. Vernadsky Crimean Federal University,

Simferopol, Russia

E-mail: lynx99@inbox.ru

Kirill I. Redkokosh

Postgraduate student

Institute of Physics and Technology, V.I. Vernadsky

Crimean Federal University, Simferopol, Russia

E-mail: kirillf13@yandex.ru

Pavel O. Mikheyev

Student

Institute of Physics and Technology, V.I. Vernadsky

Crimean Federal University, Simferopol, Russia

E-mail: pavel0990848502@gmail.com