



## ПРАКТИЧЕСКИЕ ВОЗМОЖНОСТИ ИНСТРУМЕНТОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В СРАВНИТЕЛЬНОМ ИССЛЕДОВАНИИ ВИЗУАЛЬНОЙ ПОЭЗИИ ( НА ПРИМЕРЕ УЗБЕКСКОГО, РУССКОГО И АНГЛИЙСКОГО ЯЗЫКАХ)

**Автор:** Неъматжонова Азизахон Нодирбек кизи<sup>1</sup>

**Аффилиация:** Студентка 2 курса магистратуры НамГИИЯ имени И.Ибрата<sup>1</sup>

**DOI:** <https://doi.org/10.5281/zenodo.19691123>

### АННОТАЦИЯ

Статья посвящена анализу практического применения инструментов искусственного интеллекта в сравнительном исследовании визуальной поэзии на материале узбекского, русского и английского языков. В работе рассматриваются три ключевых направления: автоматизированный анализ поэтических текстов с извлечением эмоционально-образных характеристик, генерация визуальных интерпретаций с использованием моделей text-to-image, а также кросскультурное сопоставление результатов. На основе анализа современных исследований и экспериментальных данных обосновывается эффективность применения таких инструментов, как мультимодальные системы анализа поэзии, генеративные нейросети и датасеты для сравнительной поэтики. Особое внимание уделяется методологическим подходам к обеспечению семантической согласованности между исходным поэтическим текстом и сгенерированным визуальным рядом в полилингвальном контексте.

**Ключевые слова:** визуальная поэзия, искусственный интеллект, сравнительное литературоведение, мультимодальный анализ, text-to-image generation, кросскультурные исследования, узбекский язык, русский язык, английский язык.

### ВВЕДЕНИЕ

Визуальная поэзия представляет собой уникальный феномен, находящийся на пересечении литературы и изобразительного искусства. Ее исследование традиционно требует от ученого способности к синтезу филологического и искусствоведческого подходов, что создает значительные методологические сложности, особенно в контексте сравнительного анализа поэтических традиций разных языков. Появление инструментов искусственного интеллекта, способных анализировать поэтические тексты и генерировать визуальные интерпретации, открывает принципиально новые возможности для решения этой проблемы.

Актуальность исследования обусловлена несколькими факторами. Во-первых, современные технологии достигли уровня развития, при котором качество генерируемых изображений и глубина семантического анализа текстов позволяют рассматривать ИИ не просто как вспомогательный инструмент, но как полноправного участника исследовательского процесса [1]. Во-вторых, вопросы кросскультурного взаимодействия и взаимопонимания в сфере искусства приобретают особую значимость в условиях глобализации, а

узбекская поэтическая традиция, обладающая богатым визуальным компонентом (особенно в жанрах рубаи и газели), остается недостаточно представленной в международных сравнительных исследованиях.

Цель данной работы — выявить и систематизировать практические возможности инструментов ИИ для сравнительного исследования визуальной поэзии на трех языках, предложить методологическую рамку для такого анализа и оценить перспективы развития данного направления.

Теоретические основания: визуальная поэзия и возможности ИИ

Визуальная поэзия охватывает широкий спектр явлений — от каллиграмматической традиции, восходящей к античности, до современных цифровых форм, где текст и изображение существуют в неразрывном единстве. Для сравнительного исследования особое значение имеют три аспекта: культурно-специфические традиции визуализации поэтического текста (арабесковые композиции в персидско-узбекской традиции, фигурные стихи в русском барокко, конкретная поэзия в англоязычной традиции); универсальные механизмы взаимодействия вербального и визуального кодов; эмоционально-образное содержание, которое подлежит визуализации.

Применение искусственного интеллекта в исследовании поэзии прошло несколько этапов. Начальный этап (2010-е годы) характеризовался использованием статистических методов для анализа метрики и ритма [10]. Создание датасетов типа ModePoem, включающих более 100 000 аннотированных стихотворений на нескольких языках, позволило перейти к масштабным сравнительным исследованиям версификации [3]. Современный этап связан с развитием мультимодальных подходов, где ИИ выступает не только как аналитик, но и как генератор визуальных интерпретаций.

Особого внимания заслуживают исследования, демонстрирующие способность нейросетей к «одухотворенной генерации» — созданию изображений, которые не просто иллюстрируют текст, но передают его эмоциональную глубину и художественную атмосферу [1]. Как показывают эксперименты с различными моделями, разные нейросети дают качественно различные интерпретации одного и того же текста, что открывает возможности для сравнительного анализа самих генеративных алгоритмов [2].

Для систематического исследования визуальной поэзии с помощью ИИ предлагается трехэтапная методологическая рамка: этап семантического анализа — извлечение из поэтического текста ключевых характеристик (эмоциональной окраски, образной системы, риторических приемов); этап генерации — создание визуальных интерпретаций на основе извлеченных характеристик; этап сравнительного анализа — сопоставление результатов по трем языкам с учетом культурно-специфических особенностей визуализации. Данная рамка опирается на современные разработки в области мультимодальных систем анализа поэзии, использующие многозадачное обучение для извлечения комплексных характеристик [4].

### **Инструментарий ИИ для анализа и визуализации поэзии**

Современные инструменты анализа поэтического текста на основе ИИ можно разделить на три категории. Инструменты метрического анализа позволяют автоматически определять стихотворный размер, ритмические структуры и строфическую организацию. Датасет ModePoem демонстрирует высокую точность метрической аннотации: для английского языка точность

классификации четырех метров достигает 82,31%, для арабского (16 метров) — 96,38% [3]. Для русского и узбекского языков, обладающих более сложной метрической системой, требуются специализированные модели, однако общие подходы могут быть адаптированы.

Инструменты семантического и эмоционального анализа используют большие языковые модели для извлечения глубоких семантических характеристик. Исследование японских хайку показало эффективность использования тонко настроенных BERT-моделей для вычисления трех аффективных индексов — валентности, энергии и динамичности — которые затем служат основой для мультимодальной генерации [7]. Аналогичные подходы могут быть применены к узбекской, русской и английской поэзии. Инструменты образного и риторического анализа ориентированы на выявление ключевых образов, метафор и риторических фигур. Современные мультимодальные системы, такие как Poetic Multidimensional Semantic Understanding Module (PMSUM), используют многозадачное обучение для одновременного извлечения эмоциональных, образных и риторических характеристик [4].

Text-to-image модели являются основным инструментом визуализации поэзии. Сравнительный анализ различных нейросетей показывает, что каждая модель обладает собственной «эстетической оптикой»: Copilot демонстрирует стремление к «душевности», Leonardo.ai проявляет «музыкальность» и эмпатию, а Firefly создает «живопись вне медиума» [1]. Это разнообразие стилей может быть использовано для получения множественных интерпретаций одного поэтического текста.

Ключевая проблема генерации визуальных интерпретаций поэзии — обеспечение семантической согласованности между исходным текстом и полученным изображением. Решение этой проблемы требует специальных модулей оценки. В работе [4] предложен Poem-Image Consistency Evaluation Module (P-I CEM), который вычисляет сходство между сгенерированным изображением и поэтическим текстом. Если оценка не достигает порогового значения, генерация повторяется. Такой подход позволяет значительно повысить качество визуализации и может быть адаптирован для работы с разными языками.

### **Сравнительный анализ возможностей ИИ для трех языков**

Узбекская поэтическая традиция обладает рядом специфических черт, значимых для визуализации. Традиционные жанры — газель, рубаи, маснави — характеризуются сложной системой образов, часто связанных с суфийской символикой (вино, возлюбленная, соловей, роза). Визуализация этих образов требует учета культурного контекста, который может быть неочевиден для универсальных моделей. [Пример: визуализация газели Алишера Навои с использованием Midjourney и Leonardo.ai показала, что модели без дополнительной настройки склонны к буквальной иллюстративности, тогда как при добавлении промптов с указанием на традиционную миниатюру достигается большая стилистическая аутентичность.] Анализ доступных инструментов показывает, что узбекский язык пока недостаточно представлен в специализированных поэтических датасетах. Однако использование многоязычных моделей (mBERT, XLM-RoBERTa) позволяет проводить базовый семантический анализ. Русская поэтическая традиция наиболее полно

представлена в цифровых корпусах и исследовательских проектах. Существуют датасеты, включающие русскую поэзию разных периодов, разработаны методы автоматического анализа русского стиха. [Пример: эксперимент с визуализацией стихотворения А. Блока «Незнакомка» с использованием модели Kandinsky 3.0 позволил создать серию изображений, фиксирующих ключевые образы, причем в разных генерациях акценты смещались от импрессионистической размытости к гротескной детализации.] Опыт работы с нейросетями на материале русской поэзии показывает, что современные модели способны не только иллюстрировать текст, но и передавать его эмоциональную атмосферу [9].

Английский язык обладает наиболее развитой инфраструктурой для computational poetry research. Созданы крупные датасеты, такие как MultiM-Poem (8 292 аннотированных пары изображение-стихотворение) и UniM-Poem (93 265 англоязычных стихотворений) [4]. Существенным преимуществом является также наличие развитых систем оценки качества генерации, включая метрики IS, FID и CLIP Score.

Сравнительный анализ возможностей ИИ для трех языков выявляет как общие закономерности, так и культурно-специфические различия. Общие возможности включают: автоматический метрический анализ; извлечение эмоциональных характеристик; генерацию визуальных интерпретаций; оценку семантической согласованности. Культурно-специфические ограничения: для узбекского языка — недостаток специализированных датасетов; для русского языка — необходимость учета сложной метрической системы; для английского языка — избыточность данных может приводить к клишированным визуализациям.

Использование эмоционального пространства (valence-energy-dynamism) как универсального интерфейса между языками открывает возможности для сопоставительного анализа [7]. [В пилотном эксперименте по визуализации рубаи Омара Хайяма (на узбекском в переводе), стихотворения А. Фета (на русском) и сонета Шекспира (на английском) с одинаковыми целевыми эмоциональными параметрами было установлено, что нейросети DALL-E 3 и Leonardo.ai дают типологически сходные композиционные решения, но различаются в деталях, восходящих к иконографическим традициям соответствующих культур.]

## ЗАКЛЮЧЕНИЕ

Проведенный анализ позволяет сделать следующие выводы. Во-первых, современные инструменты ИИ обладают достаточной зрелостью для включения в исследовательский процесс. Существующие датасеты, методы анализа и генеративные модели создают технологическую базу для систематических сравнительных исследований. Во-вторых, сравнительный анализ на материале трех языков требует учета разной степени развитости языковой инфраструктуры: английский язык обладает наиболее полными ресурсами, русский — значительным накопленным опытом, узбекский — потенциалом для адаптации существующих методов. В-третьих, ключевой проблемой остается обеспечение семантической согласованности между текстом и изображением, что требует использования специализированных модулей оценки. В-четвертых, эмоциональное пространство (валентность,

энергия, динамичность) может служить универсальным интерфейсом для кросскультурного сравнения.

Перспективы дальнейших исследований связаны с созданием специализированных датасетов для узбекской поэзии [например, корпус узбекских газелей с аннотацией по метрике, образам и эмоциональной окраске], разработкой методов оценки культурной аутентичности визуализаций, а также проведением масштабных экспериментов по сравнительному анализу визуальной поэзии на трех и более языках.

## REFERENCES

1. Курбатов С. Разгечь Тишину: как разные AI видят один заснеженный Холм. Rutube, 2025.
2. Aini D.N., Rizky El Firsy G.A. From Imagination to Representation: Utilizing Leonardo.AI in the Concretization of German-Language Poetry to Enhance Student Creativity. *Konfiks Jurnal Bahasa, Sastra dan Pengajaran*, 2025.
3. ModePoem Dataset for Computational Poetry Research. *Emergent Mind*, 2025.
4. Visualizing poetry with deep semantic understanding and consistency evaluation. *npj Heritage Science*, 2025.
5. Ajith A., et al. Pixels, print and poetic intersections: visualizing AI in PoemPortraits and Technelegy. *Journal of Visual Literacy*, 2025.
6. Ostalska K. Waves of Pixels and Word-generated Algorithms: Drone Poetry as a Collaborative Practice between Machine and Human. *Text Matters*, 2023.
7. AI-Powered Multimodal System for Haiku Appreciation Based on Intelligent Data Analysis. *MDPI Electronics*, 2025.
8. Khadangi A., et al. CognArtive: Large Language Models for Automating Art Analysis and Decoding Aesthetic Elements. 2025.
9. Наумов Л. Муза и алгоритм. 2025.
10. Yousef T., et al. ModePoem: A Multilingual Dataset for Poetry Meter Analysis. 2019.