

# Лисичка і Журавель

Жили-були у лісі Лисичка та Журавель. Вони товаришували і часто проводили багато часу разом.

Запросила Лисичка Журавля до себе в гості та пригостила його манною кашою. Але каша була розмазана тоненьким шаром по плоскій тарілці

Як не старався Журавель поїсти - ніяк у нього не вдавалося. Його дзоб був дуже незручним для того, щоб їсти манну кашу з плоскої тарілки



Тоді Лисичка взяла, та з'їла всю кашу сама, а Журавель голодний повернувся з гостювання

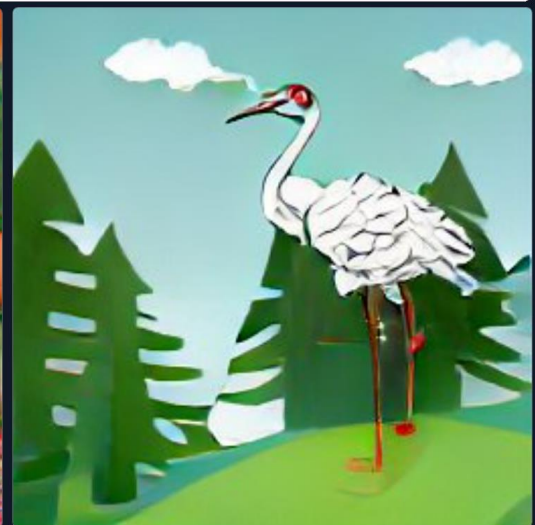
Згодом Журавель вирішив відплатити тією ж монетою і запросив Лисичку до себе в гості

Він приготував смачне запечене м'ясо і помістив його у високий глиняний глечик з вузьким горлечком.



Як би Лисичка не старалася, але ніяк не могла скуштувати смачної їжі, бо лапки були занадто короткі, а голова не пролазила у горлечко глечика.

Так і посварилися Лисичка і Журавель. Вони більше не товаришували і не проводили час разом.



Query list:

1. one fox and one crane bird next to to each other friends cartoon
2. one fox and one crane bird inside the room cartoon
3. crane bird eating porridge from flat plate on the table cartoon
4. fox sitting near the flat plate on the table cartoon
5. one crane bird and one fox inside the room cartoon
6. fox sitting next to clay jug inside room cartoon
7. crane bird sitting near clay jug on the table inside room cartoon
8. one fox standing in the forest cartoon
9. one crane bird standing in the forest cartoon

## Звіт використання генеративної моделі Dall-E mini

Для генерації картинок для казки «Лисиця та Журавель» я обрав генеративну модель Dall-E mini на веб-сервісі <https://www.craiyon.com> Ця модель швидко будує спрощені зображення на основі моделі Dall-E 2 від компанії OpenAI. На відміну від сервісів типу NightCafe, Dream Studio чи MidJourney, тут можна задавати стиль зображення в стрічці запиту. NightCafe має декілька моделей заточених під генерування зображень в певному стилі: CGI, PopART, realistic, тощо. MidJourney створює більш абстрактні, естетичні зображення часто з помітною симетрією. Проте за рахунок їх специфікації зображення мають більшу чіткість і «точність».

При роботі з Dall-E mini можна прослідкувати принцип роботи цієї моделі. Вона знаходить схожі картинки до запиту і запам'ятовує ключові артефакти. У випадку з лисицею, це помаранчевий колір, великий хвіст. При генерації журавля часто присутні довгі лапи, білий силует крил та тонка закручена шия. Неточності виникають при спробах моделі об'єднати ці частинки до купи. Модель часто генерувала помаранчеву лисицю на довгих тонких лапах журавля. У 6 і 7 запиті модель часто поєднувала образи тварин та глечика у один контур. Потрібно було уточнювати кількість тварин, адже інколи при запиті «fox», модель могла згенерувати більше ніж одну лисицю, хоча слово написано в однині. Інколи приходилося регенерувати зображення тим самим запитом, аби отримати кращі зображення.

Я вважаю, що складно стверджувати розуміння світу моделлю. Я б радше сказав, що модель просто розбиває запит на окремі слова, генерує ключові артефакти кожної частини запиту і потім об'єднує їх в одне зображення. Помітна наявність розмежування



об'єктів на «головні» та «фонові», адже задавши «in the room» чи «in the forest» модель генерує відповідне оточення позаду персонажів, проте на картинці помітно, що модель не згенерувала ноги лисиці під столом. Також тут помітно, що на столі знаходяться 2 тарілки, проте у запиті була вказана одна, до того ж плоска.

Враховуючи факт відсутності глибокого розуміння речей на зображеннях, чи можна вважати їх мистецтвом? Я вважаю, що кожен може дати відповідь собі сам, враховуючи важливість закладеного сенсу автором. Для когось мистецтво там, де автор заклав у нього свій глибокий сенс і потрібно добре проіннятися, аби його зрозуміти. Проте для мене мистецтво – це все, що нас оточує і кожен рефлексує індивідуально. Глибокі сенси можна побачити у осінньому листі, що падає на землю, безглуздих написах на стінах і згенерованих ШІ картинках. Прикладом таких рефлексій можуть виступити обговорення згенерованих зображень за запитом «the last war» чи «civilization in 3022».

Попри слабе розуміння запитів користувача, модель все ж навчається з кожним запитом все більше і точно можна сказати, що одного дня модель зможе перевершити професійного художника в генеруванні дуже чітких і естетичних моделей. Цей фактор навчання і свідчить про певну інтелектуальність цих моделей. Куди складніше буде отримати від такої моделі зображення, яке при запиті уявляє користувач. Тобто попрохавши свого друга художника намалювати певну картину, ви можете багато уточнювати найменші деталі, експозицію, кольорову гаму та емоції, які передає зображення, доки не отримаєте картину максимально близьку до своєї уявної. Це і є найбільшою проблемою у використанні моделей ШІ. Люди можуть комунікувати про емоції, які викликаються різними речами, проте відчуття схожі. Модель емоцій не відчуває. Іронічно, проте це і породжує основну проблему раціональності – люди здатні раціонально донести інформацію за допомогою емоцій, що дозволяє максимально точно передати все до найменших деталей.