



Финансирано от
Европейския съюз
Следващо Поколение ЕС



План за възстановяване и
устойчивост



Република България

ВЪТРЕШЕН ДОКЛАД

от гл. ас. д-р Йордан Кралев
до проф. д-р Светла Коева, ръководител на проекта,

за текуща самооценка на техническата готовност на софтуер

Разработване на софтуер за генериране на текст от изображения чрез фина настройка на големи езикови модели

планиран резултат от **Работен пакет 5. Фина настройка на големи езикови модели** по проекта „Инфраструктура за фина настройка на предварително обучени големи езикови модели“, договор No ПБУ – 55 от 12.12.2024 г. /BG-RRP-2.017-0030-C01/ от Механизма за възстановяване и устойчивост за изпълнение на инвестиция по C2I2 „Повишаване на иновационния капацитет на Българската академия на науките (БАН) в сферата на зелените и цифровите технологии“ от Плана за възстановяване и устойчивост

Вътрешен доклад ID: IFGPT-2025-REPORT-006

Име и версия на продукта: **Софтуер за генериране на текст от изображения version 0.1**

Тип на продукта: **Софтуер**

Текущо ниво на техническа готовност: **TRL 4**

Дата: **15.09.2025 г.**

✓ **TRL 3: Начални скриптове и функции (Септември 2025)**

Изисквания:

- Разработени са базови алгоритми, модули или скриптове, които решават ядрото на поставения проблем.
- Не се изисква цялостен интерфейс или архитектура.
- Липсва интеграция между отделни модули.

Доказателства:

- Прототипен код на инфраструктура за финна настройка на визуални модели от типа на Detectron2 (<https://github.com/DCL-IBL/MIC21-framework>)
 - Среда за анализиране на изображения FiftyOne и за сравнение на резултатите от семантичната анотация с различни модели
 - Разширение на възможностите на базовата система FiftyOne с бутон за качване на изображения от потребител
 - Backend услуга и REST API за автоматична анотация на качените изображения посредством обучени модели по тематични области
 - Jupyter Notebook съдържащ скриптове за настройка на визуални модели към изображения от конкретна тематична област
 - Достъп до база данни с предварително анотирани изображения от 130 тематични области (https://dcl.bas.bg/MIC-21/model_images/)
 - Достъп до база данни с предварително обучени модели за анотиране и етикетиране на изображенията от всяка тематична област (<https://dcl.bas.bg/MIC-21/models/>).
 - Онлайн версия на системата (<https://mic21.dcl.bas.bg>)
- Обогаляване на базата данни от анотирани изображения с лингвистични описания към всяко изображение
 - Прототипен код за автоматично генериране на заглавия и описания на изображения с използване на големи езикови модели (https://github.com/DCL-IBL/MIC21-Summarizer/blob/main/notebooks/prepare_dataset.ipynb)
 - Ръчно обработени и подобрени описания за всички изображения по категории
 - Подготвена и качена в HuggingFace база данни във формат подходящ за извършване на финна настройка на мултимодални езикови модели (<https://live.european-language-grid.eu/catalogue/>)
- Прототипен код за финна настройка на интегрирана структура включваща визуален модел и езиков модел (<https://github.com/DCL-IBL/MIC21-Summarizer>)
 - Обектен клас за реализация на структурата за директно проектиране на изходит на конволюционния слой на визуалния модел във входното векторно пространство на езиковия модел (<https://github.com/DCL-IBL/MIC21-Summarizer/blob/main/mic21s.py>)
 - Скрипт за финна разпределяне на модела върху няколко графични карти и финна настройка чрез векторизация на примерите от базата данни

(https://github.com/DCL-IBL/MIC21-Summarizer/blob/main/train_mic21s_base.py)

- Скрипт за финна настройка на модела по тематични области
(https://github.com/DCL-IBL/MIC21-Summarizer/blob/main/train_mic21s.py)
- Документация на алгоритмите:
 - Публикация на тема “Fusion of Object-Centric and Linguistic Features for Domain-Adapted Multimodal Learning”
(<https://acl-bg.org/proceedings/2025/RANLP%202025/pdf/2025.ranlp-1.69.pdf>)

✓ TRL 4: Алфа версия – вътрешно тестване (Ноември 2025)

Изисквания:

- Алфа версия със заданието на основната архитектура и функционалност.
- Тествана вътрешно от разработчиците или QA екипа.
- Има базов интерфейс и ограничена стабилност.

Примерни доказателства:

- Вътрешни тестови доклади
- Скриншоти/видео от интерфейса
- Release notes на Alpha версия
- Документирана архитектура

✓ TRL 5: Алфа версия – външно тестване (Предстои: Януари/Февруари 2026)

Изисквания:

- Същата Alpha версия е предоставена на ограничен брой външни потребители (напр. партньори, тестери).
- Събират се обратна връзка и бъгове от реална употреба.
- Тестовата среда е частично контролирана.

Примерни доказателства:

- Резултати от потребителски тестове (UX reports)
- Доклади от външни бета тестери
- Журнал на грешки (issue tracker)
- Споразумения с външни тестови партньори

