Open Source: инструменты российских команд для задач ИИ-разработки

80 решений

1. Библиотеки и фреймворки (25)

Наборы готовых классов объектов, решений и функций, необходимых для выполнения определенной задачи. Элементы фреймворков по сравнению с аналогичными у обычных библиотек гораздо более взаимосвязаны между собой и направляют разработчиков при решении задачи в строго определенном направлении.

- Работа с данными — **AggMe** — SBER DEVICES
- Работа с данными — **DataProcessing Framework** — СБЕР
- Оптимизация качества обучения нейросетей — **CoolGraph** — МТС
- Оптимизация качества обучения нейросетей — **Py-Boost** — СБЕР
- Оптимизация качества обучения нейросетей — **Stalactite** — СБЕР
- Оптимизация качества обучения нейросетей — **Turbo Alignment** — Т
- Оптимизация качества обучения нейросетей — **XLand-MiniGrid** — Т
- Оптимизация качества обучения нейросетей — **XNumPy** — КРИПТОПРО
- Оптимизация качества обучения нейросетей — **YaFSDP** — YaFSDP
- Сжатие и оптимизация инфраструктуры — **Eco4cast** — AIRI
- Сжатие и оптимизация инфраструктуры — **NNTile** — AIRI
- Сжатие и оптимизация инфраструктуры — **SaluteVision Mobile SDK** — СБЕР
- Текстовые задачи — **RURAGE** — МТС
- Текстовые задачи — **Yandex Cloud ML SDK** — Yandex Cloud
- Работа с рекомендательными системами — **MRGSRec** — AIRI
- Работа с рекомендательными системами — **RePlay** — СБЕР
- Работа с рекомендательными системами — **Sim4Rec** — СБЕР
- Определение эмоций — **Aniemore** — Aniemore
- Другие задачи — **RIDE** — СБЕР
- Другие задачи — **StyleFeatureEditor** — AIRI

Проекты, которые создаются командами в регионах: **EyeFeatures** **ChatAI** **SimpleLLMServer** **SharpGPT** **ProtoLLM**

2. Модели искусственного интеллекта (20)

Алгоритмы, которые могут обучаться определенным действиям на основе предоставляемых данных.

- Большая мультимодальная модель — **OmniFusion** — AIRI
- Большая мультимодальная модель — **Vikhr Salt**
- Большая языковая модель — **mGPT** — SBER DEVICES
- Большая языковая модель — **ruGPT-3.5** — СБЕР
- Большая языковая модель — **T-Lite** — Т
- Большая языковая модель — **T-Pro** — Т
- Модель для обработки и генерации звука — **GigaAM**
- Модель для обработки и генерации изображений — **Kandinsky** — СБЕР
- Модель для обработки и генерации изображений — **MiVOLO** — SBER DEVICES
- Модель для обработки и генерации изображений — **Switti** — Yandex Research
- Модель для обработки и генерации изображений — **UniDet3D** — AIRI
- Модель-классификатор — **ESGify** — СБЕР
- Модель-классификатор — **GigaEmbeddings** — SBER DEVICES
- Модель контекстного обучения с подкреплением — **Headless-AD** — Т
- Модель для научных задач — **AIpom**
- Модель для научных задач — **GENA** — AIRI
- Модель для научных задач — **LegNet**
- Модель для научных задач — **ProBAN**
- Модель для научных задач — **SciRus-tiny**

Проекты, которые создаются командами в регионах: **«Менон» (Meno)**

3. Бенчмарки (18)

Наборы данных, тестов и заданий для оценки производительности моделей искусственного интеллекта по различным метрикам.

- Общего профиля — **MERA**
- Для сравнения производительности — **LLM Arena** — LLMARENA
- Для сравнения производительности — **RuArenaGeneral**
- Для сравнения производительности — **rulm-sbs2** — RUS LAB
- Для сравнения производительности — **ruMT-Bench** — SBER DEVICES
- Для длинного контекста — **BABILong** — AIRI
- Для длинного контекста — **LIBRA**
- Для обработки и понимания естественного языка — **RuCoLa**
- Для обработки и понимания естественного языка — **ruIFEval** — SBER DEVICES
- Для обработки и понимания естественного языка — **RussianSuperGLUE** — SBER DEVICES
- Для обработки и понимания естественного языка — **TAPE** — SBER DEVICES
- Для русскоязычного контекста — **Шлепа**
- Для русскоязычного контекста — **SLAVA**
- Для медицинских данных — **MedBench** — СБЕР
- Для эмбеддингов и эмбеддеров — **ruMTEB** — SBER DEVICES
- Для эмбеддингов и эмбеддеров — **ruSciBench**
- Для эмбеддингов и эмбеддеров — **RuSentEval**
- Для программного кода — **RealCode_eval** — SBER DEVICES

4. Датасеты (8)

Обработанные и структурированные наборы данных для обучения и/или тестирования моделей искусственного интеллекта.

- Визуальные задачи — **EasyPortrait** — SBER DEVICES
- Визуальные задачи — **Skoltech3D** — Skoltech
- Научные задачи — **∇²DFT**
- Научные задачи — **XLand-100B** — Т
- Распознавание языка и эмоций — **Bukva** — SBER DEVICES
- Распознавание языка и эмоций — **Dusha** — SBER DEVICES
- Распознавание языка и эмоций — **Slovo** — SBER DEVICES
- Другие задачи — **Geo Reviews Dataset 2023**

5. Методы (6)

Наборы действий, процедуры, необходимые для выполнения какой-либо задачи.

- Оптимизация обучения моделей — **ReBRAC** — Т
- Оптимизация обучения моделей — **SAC-RND** — Т
- Оптимизация обучения моделей — **SDDE** — Т
- Сжатие моделей — **AQLM и PV-tuning** — Yandex Research
- Сжатие моделей — **LLM Microscope** — AIRI
- Борьба с мошенничеством — **PROWN** — AIRI

6. Архитектуры (2)

План, по которому строится та или иная нейросеть: как выстроены различные слои, как они взаимосвязаны и так далее.

- **AASIST3** — AIRI
- **ReBased** — Т

7. Платформы (1)

В этой категории представлена общедоступная платформа обработки и хранения данных.

- **YTsaurus**

Подборка на основе данных ICT.Moscow включает решения, публично представленные с декабря 2022 года по январь 2025 года.

Для разработок, создаваемых в столице, предусмотрена ссылка на карточку в профильном разделе портала. В ней можно изучить возможности решения, условия использования, а также найти дополнительную информацию и контакты разработчиков.

По всем вопросам: hello@ict.moscow

Версия 1.0 от 27 января 2025 года