

# Автоматизированное описание содержания изображений

## *Общие принципы функционирования системы*

На вход система получает изображение, результатом выдачи является предложение текста, кратко описывающее содержание полученного изображения с указанием основных объектов и действий. К примеру (изображение и выдача):



*Самолет летит в пасмурном небе с выпущенными шасси*



*Два взрослых слона идут по полю с выжженной травой*

Может также применяться для улучшенного поиска по массиву изображений. К примеру, система может подбирать изображения по запросу “Мужчина в черной рубашке играет на белой гитаре на сцене”, если таковые есть в базе.

### *Технология и взаимодействие*

В основе системы лежит искусственная нейронная сеть, построенная по принципам, описанным в научной работе [Show, Attend and Tell: Neural Image Caption Generation with Visual Attention](#). Для надлежащего функционирования нейросеть обучена на объемном массиве данных (миллионы изображений). При “просмотре” нового изображения, таким образом, нейросеть выделяет значимые объекты, которые ей удастся узнать, а затем связывает их в единую модель и формирует связное текстовое описание.

NB! Нейронная сеть на текущем уровне технического развития может ошибаться и принимать одни объекты за другие, средний процент верного результата составляет 86%.  
К примеру, система может подумать, что “Мужчина держит в руке часы”, когда как мужчина держит значок.

Система может работать как удаленно, получая на вход изображения и выдавая результат через API, так и может быть установлена локально на соответствующем требуемым характеристикам серверном оборудовании с поддержкой GPU.

Срок изготовления 1 функционирующей копии системы – 1 месяц.

–

mira / arb.digital

<http://ai.arb.digital>  
[hello@arbdigital.ru](mailto:hello@arbdigital.ru)

+7 (495) 132-61-12  
Москва, Садовая-Триумфальная, 4-10