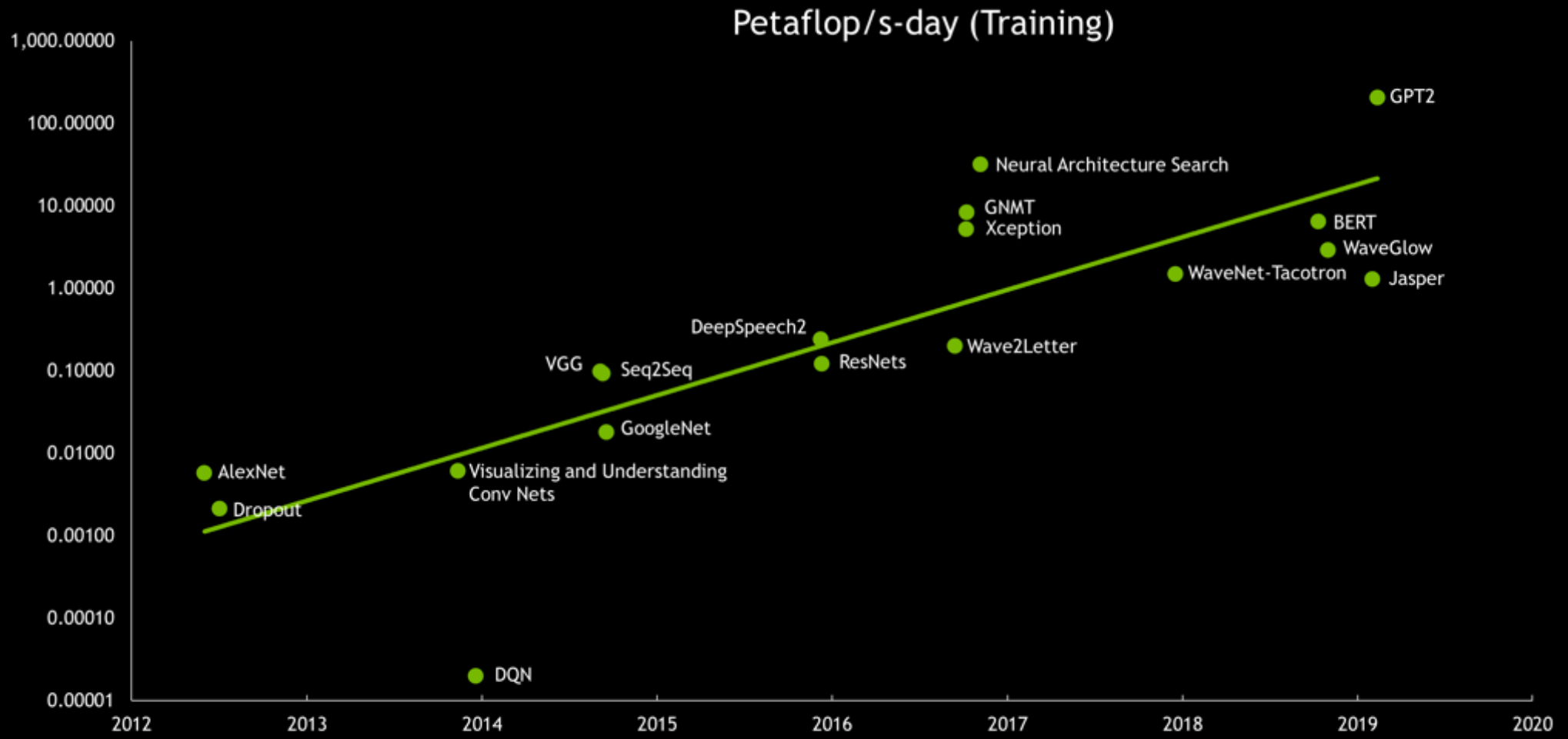
ПРИКЛАДНОЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Антон Джораев

ВЗРЫВНОЙ РОСТ СЛОЖНОСТИ НЕЙРОСЕТЕЙ



Source: OpenAI and NVIDIA

ЧТО ТАКОЕ РАЗГОВОРНЫЙ ИИ?

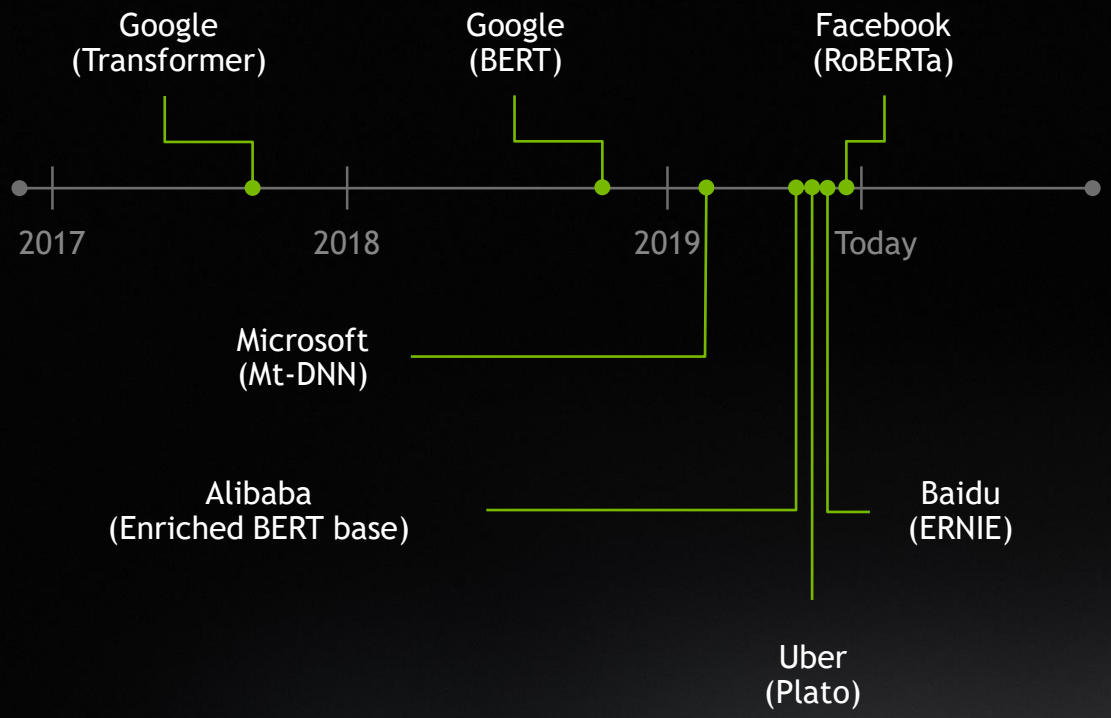
Диалоговый, не транзакционый

Используется контекст

Ответы должны быть моментальными

Модели становятся очень большими

ГОНКА ЗА РАЗГОВОРНЫМ ИИ



Лидеры рейтинга GLUE

Rank	Name	Model	Score
1	Google	ALBERT (ensemble)	89.4
2	Microsoft	ADV-RoBERTa (ensemble)	88.8
3	Facebook AI	RoBERTa	88.5
6	GLUE Human Baselines	GLUE Human Baselines	87.1

https://gluebenchmark.com/leaderboard

NVIDIA DGX-2

Самая мощная в мире DL-система для самых сложных задач глубокого обучения

- Первая в мире 2 PFLOPS система
- 16 Tesla V100 32GB GPUs с общим интерконнектом
- NVSwitch: 2.4 TB/s пропускная способность
- 0.5 TB унифицированной памяти GPU

www.nvidia.ru/dgx

NVIDIA DGX SUPERPOD

Лидерство в ИИ невозможно без лидерства в инфраструктуре

Автопилот | Речевой ИИ | Здравоохранение | Графика | HPC

Испытательный стенда для масштабируемых HPC систем
- 9.4 PF Linpack | ~200 PF AI | #22 в списке Top500
- <2 минут на обучение RN-50

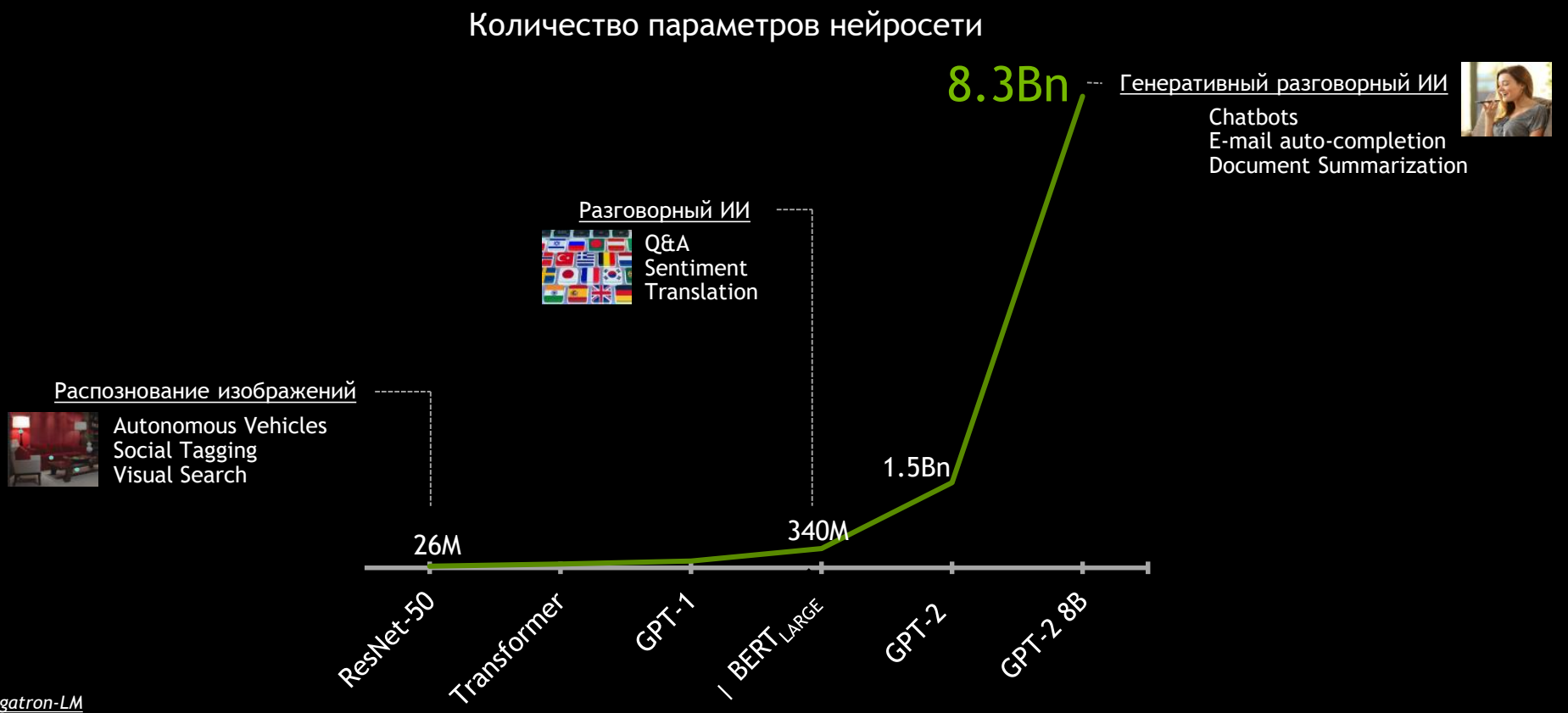
Модульная архитектура DGX SuperPOD
- Собран и запущен за 3 недели
- Оптимизации в Compute, Networking, Storage & Software

Интегрированный софтверный стек
- Свободно доступен через NGC

- 96 DGX-2H
- 10 Mellanox EDR IB per node
- 1,536 V100 Tensor Core GPUs
- 1 megawatt of power

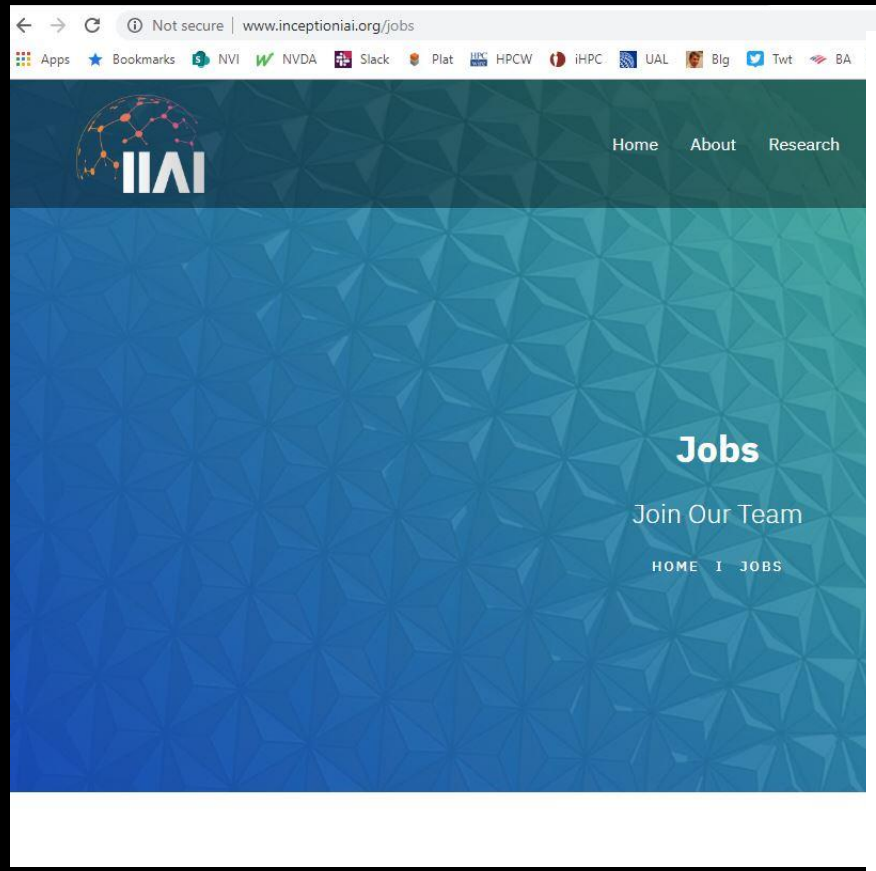
САМЫЕ СОВРЕМЕННЫЕ РЕЧЕВЫЕ МОДЕЛИ

Разработка разговорного ИИ требует гигантских объемов вычислений



https://github.com/NVIDIA/Megatron-LM

НАЙМ DATA SCIENTIST'ОВ – ЧТО ВАЖНО?



in the country.

Location

Our office is situated in the central business area of Abu Dhabi, within walking distance of shopping malls, various restaurants, the Cleveland Clinic that offers the world's best healthcare and multiple sea view luxury residential blocks. Being the home to 175 nationalities, the UAE provides an open-minded and hospitable living and working environment to all expats. From clothes to cuisines, in the UAE, you can always find the humblest and the most extravagant.

Minimum requirements

- PhD degree in Computer Science, Data Science, Mathematics, Statistics or related streams.
- Proven experience in Computer Vision, Machine/Deep Learning, Natural Language Processing, Data Mining, Multimedia, Medical Imaging, Bioinformatics or related fields.
- Strong publication record at conferences and in journals, such as TPAMI, IJCV, NIPS, ICML, ICCV, CVPR, ECCV, ICLR, KDD, ACL and MICCAI.

Benefits

- A highly motivating position that will drive you to broaden your horizon and learn from senior peers.
- Freedom of academic research without the pressure of grants and teaching in academia.
- Huge amounts of data across the country that can help you fulfill and fast-track your research into real-world scenarios.
- A GPU farm of **100 NVIDIA DGX-1** servers and other GPU machines (**1000 Tesla V100 GPUs**).
- An enjoyable modern working environment with a spectacular sea view.
- World-leading standard of healthcare, education and hospitality industry.
- All year round sunshine and amazing beaches.
- Extremely competitive package (tax-free salary).

APPLY NOW

NVIDIA DGX-1

Самый эффективный инструмент для ИИ

960 Tensor TFLOPS | 8x Tesla V100 | NVLink Hybrid Cube

Заменяет сотню традиционных серверов на задачах ИИ

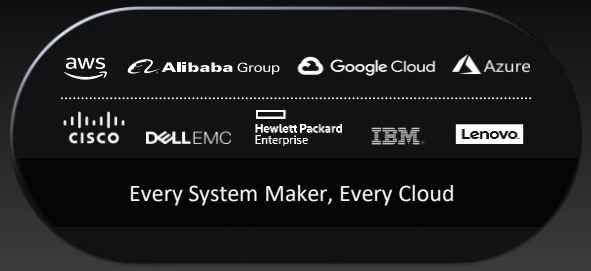
60 ТФлопс FP64

3U в стойке, 3.2 КВт

Интерконнект 4 EDR IB

www.nvidia.ru/dgx

Партнерство с NVIDIA для развития компетенции в ИИ

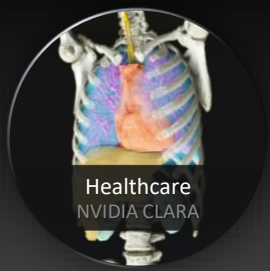


Технологии и экосистема

Оборудование и ПО компании NVIDIA является стандартом де-факто для построения ИИ-систем во всем мире. NVIDIA GPU доступны от любого производителя серверов, компьютеров и из любого облака.

Экспертиза и инвестиции

Ведущие специалисты мира в области ИИ работают на NVIDIA, ведущие ВУЗы сотрудничают с NVIDIA, программа NVIDIA Inception поддерживает ИИ-стартапы по всему миру, а институт NVIDIA DLI обучает специалистов по всему миру.



Отраслевые платформы

Четыре наши индустриальные платформы объединяют технологии и экосистемы с тем чтобы способствовать развитию фундаментальных отраслей.

ВЫВОДЫ

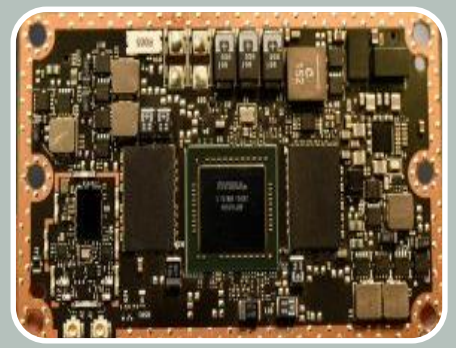
1. Ни один ВУЗ или институт больше не будет восприниматься всерьез без экспертизы в области ИИ

2. Организации без вычислительного ресурса для ИИ - неконкурентоспособны

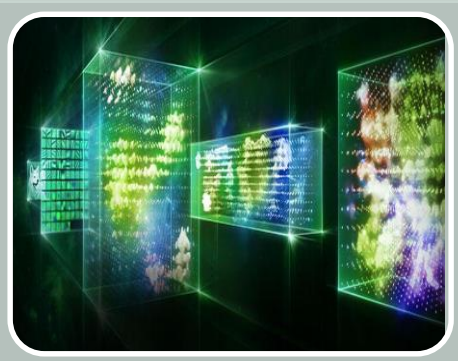
3. NVIDIA DGX и партнерство с NVIDIA – самый быстрый способ получить действующую систему и начать исследования, разработку, курсы

NVIDIA Deep Learning Accelerator

NVDLA — процессор с открытым исходным кодом



Разработан как часть Xavier – платформы NVIDIA для автономных автомобилей

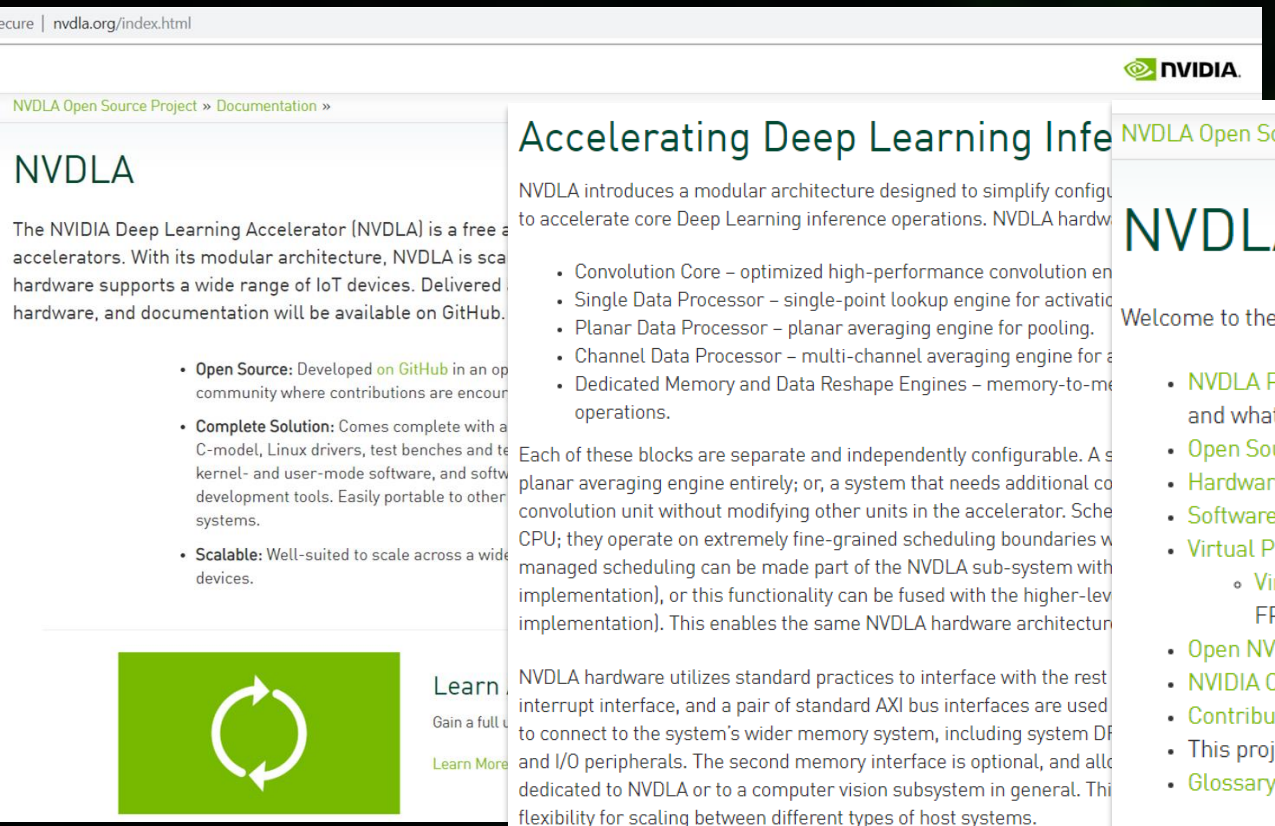


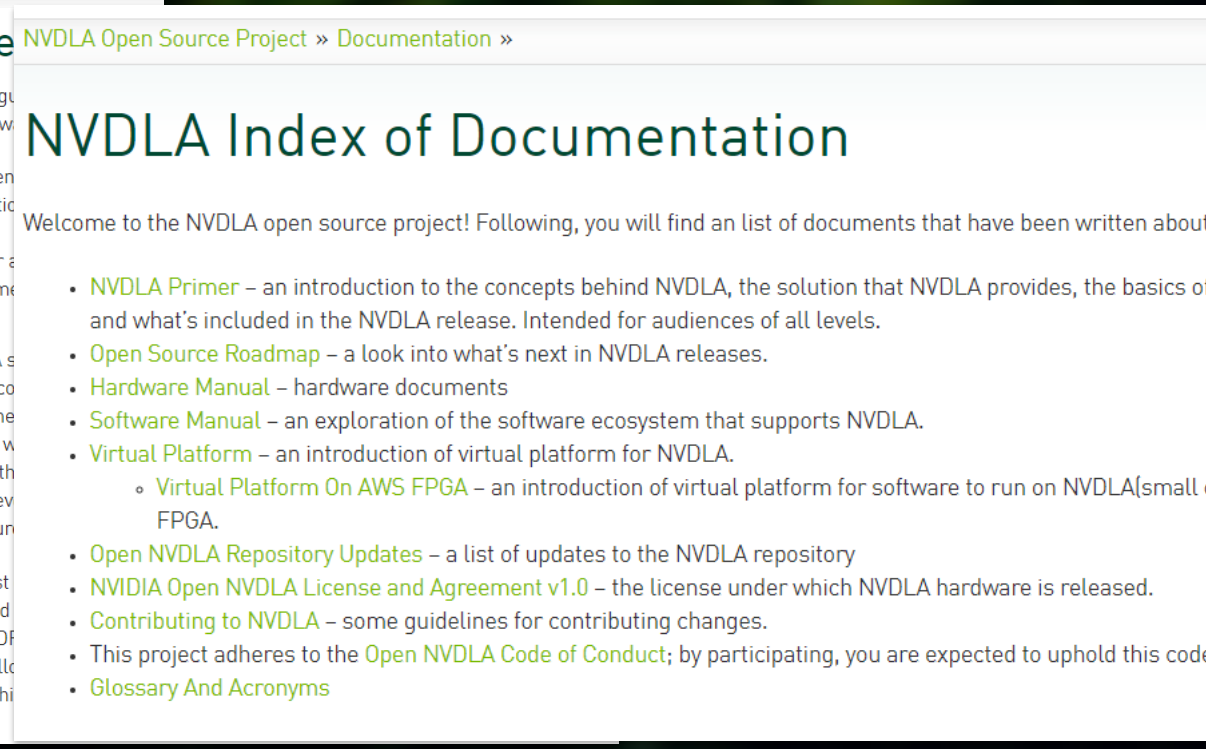
Оптимизирован для инференса сверточных нейросетей, машинного зрения

Открытая архитектура и RTL. Открытый исходный код компилятора

Законченное решение с Verilog и C-моделью, Linux драйвером, тестами, kernel и user-mode ПО и виртуальной платформой

Детали на WWW.NVDLA.ORG





Антон Джораев, adzhoraev@nvidia.com