



**INTEL<sup>®</sup>  
INNOVATION  
DAY**



# ПРЕИМУЩЕСТВА ГОМОГЕННОЙ ИНФРАСТРУКТУРЫ НА БАЗЕ INTEL ДЛЯ ИИ

Дмитрий Белошицкий

# LEGAL NOTICES & DISCLAIMERS

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at [intel.com](http://intel.com), or from the OEM or retailer. No computer system can be absolutely secure.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit <http://www.intel.com/performance>.

Cost reduction scenarios described are intended as examples of how a given Intel-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

Statements in this document that refer to Intel's plans and expectations for the quarter, the year, and the future, are forward-looking statements that involve a number of risks and uncertainties. A detailed discussion of the factors that could affect Intel's results and plans is included in Intel's SEC filings, including the annual report on Form 10-K.

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

Intel, the Intel logo, Pentium, Celeron, Atom, Core, Xeon, Movidius and others are trademarks of Intel Corporation in the U.S. and/or other countries.

\*Other names and brands may be claimed as the property of others.

© 2018 Intel Corporation.



# ОГРОМНОЕ КОЛИЧЕСТВО ДАННЫХ

В ДЕНЬ К 2020

ИНТЕРНЕТ-ПОЛЬЗОВАТЕЛЬ 1.5 ТБ

АВТОНОМНАЯ МАШИНА 4 ТБ

САМОЛЕТ 5 ТБ

УМНАЯ ФАБРИКА 1 ПБ

ОБЛАЧНЫЙ ВИДЕО-ПРОВАЙДЕР 750 ПБ



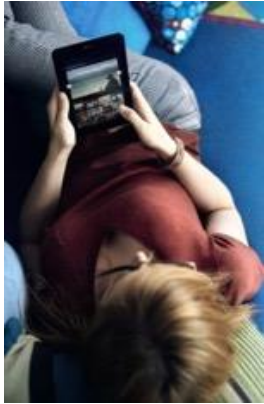
БИЗНЕС-ИНСАЙТЫ

ОПЕРАЦИОННЫЕ  
ИНСАЙТЫ

ИНСАЙТЫ ПО  
БЕЗОПАСНОСТИ

Source: Amalgamation of analyst data and Intel analysis.

# ИИ ТРАНСФОРМИРУЕТ



## В2С

Умные ассистенты  
Чатботы  
Поиск  
Персонализация  
AR  
Роботы



## МЕДИЦИНА

Диагностика  
Открытие лекарств  
Забота о пациентах  
Исследования  
Восприятие/сенсоры



## ФИНАНСЫ

Алго-трейдинг  
Распознавание мошенничества  
Исследования  
Личные финансы  
Снижение рисков



## РОЗНИЦА

Поддержка  
Опыт  
Маркетинг  
Мерчандайзинг  
Повышение лояльности  
Логистика  
Безопасность



## ГОС. СЕКТОР

Оборона  
Безопасность  
Умные города



## ЭНЕРГЕТИКА

Разведка месторождений  
Смарт грид  
Операционная оптимизация



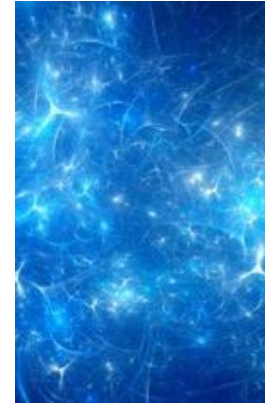
## ТРАНСПОРТ

Автономные машины  
Авиа  
Поиск



## ПРОМЫШЛЕННОСТЬ

Умные фабрики  
Предиктивная аналитика  
Сельское хозяйство



## ДРУГОЕ

Реклама  
Образование  
Гейминг  
Профессиональные & ИТ услуги  
Телеком/Медиа  
Спорт





Портфель Интел

**РЕШЕНИЯ**

Партнерская экосистема в разных вертикальных рынках

**СРЕДСТВА**

Программное обеспечение

**АППАРАТНОЕ  
ОБЕСПЕЧЕНИЕ**

Гомогенная инфраструктура: от облака к устройству

**ДАННЫЕ**

**БУДУЩЕЕ**

# АППАРАТНОЕ ОБЕСПЕЧЕНИЕ

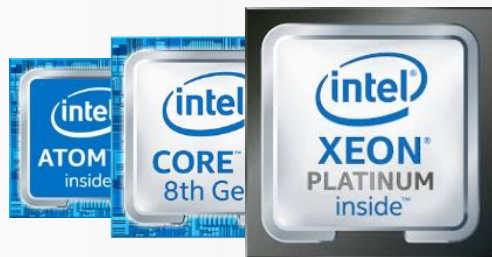
Гомогенная инфраструктура: от облака к устройству



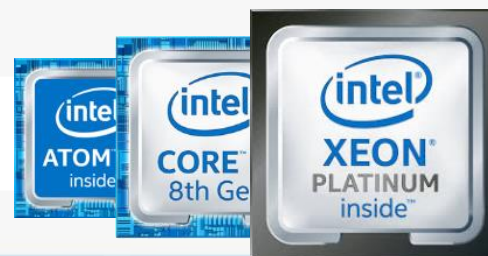
## МЕЙНСТРИМ

## УСКОРЕНИЕ

DEEP  
LEARNING  
TRAINING  
INFERENCE



# ВСЕ ОСТАЛЬНЫЕ ИИ-ЗАДАЧИ



# АППАРАТНОЕ ОБЕСПЕЧЕНИЕ

Гомогенная инфраструктура: от облака к устройству



## ТЕРМИНАЛЬНОЕ ОБОРУДОВАНИЕ



IOT сенсоры



Vision & Inference

Speech



+



Автономные машины

Автономное вождение



Десктопы и моб. устройства



+



SOC



+



M.2 Card

Дисплей, Видео, AR/VR, Жесты, Речь

<1мс

## КРАЙ СЕТИ / EDGE

Серверы и шлюзы



Большинство кейсов

+

Ускорение



Inference в обработке изображений



Inference с низкой задержкой & гибкость



<10-40мс

<5мс

## ЦОД

Серверы



Большинство кейсов

+



NNP-L

Наиболее высокоинтенсивные кейсы

+

Ускорение



Inference с низкой задержкой & гибкость

~100мс



РАЗРАБОТЧИКИ ПРИЛОЖЕНИЙ

ДАТА САЙНТИСТЫ

РАЗРАБОТЧИКИ

СИСТЕМНЫЕ АРХИТЕКТОРЫ

## ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

### Каталог ИИ-решений



## DEEP LEARNING

### OpenVINO™ †

Open Visual Inference & Neural Network Optimization toolkit for inference deployment on CPU, processor graphics, FPGA & VPU using TF, Caffe\* & MXNet\*

### Intel® Movidius™ SDK

Optimized inference deployment for all Intel® Movidius™ VPUs using TensorFlow\* & Caffe\*

## DEEP LEARNING

### Intel® Deep Learning Studio<sup>‡</sup>

Open-source tool to compress deep learning development cycle

СНОРО

## БИБЛИОТЕКИ MACHINE LEARNING

### Python

- Scikit-learn
- Pandas
- NumPy

### R

- Cart
- RandomF
- orest
- e1071

### Distributed

- MLib (on Spark)
- Mahout

## ФРЕЙМВОРКИ DEEP LEARNING

### Now optimized for CPU



TensorFlow\* MXNet\* Caffe\* BigDL/Spark\*

### Optimizations in progress



Caffe2\* PyTorch\* PaddlePaddle\*

СНОРО

## MACHINE И DEEP LEARNING

### Python

Intel distribution optimized for machine learning

### DAAL

Intel® Data Analytics Acceleration Library (for machine learning)

### MKL-DNN

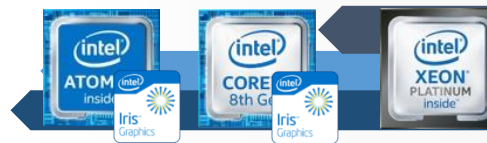
Open-source deep neural network functions for CPU, processor graphics

## КОМПИЛЯТОР DEEP LEARNING

### Intel® nGraph™ Compiler (Alpha)

Open-sourced compiler for deep learning model computations optimized for multiple devices (CPU, GPU, NNP) using multiple frameworks (TF, MXNet, ONNX)

## ОСНОВА



ЦОД  
Край сети/Edge  
Устройство

## УСКОРИТЕЛИ DEEP LEARNING

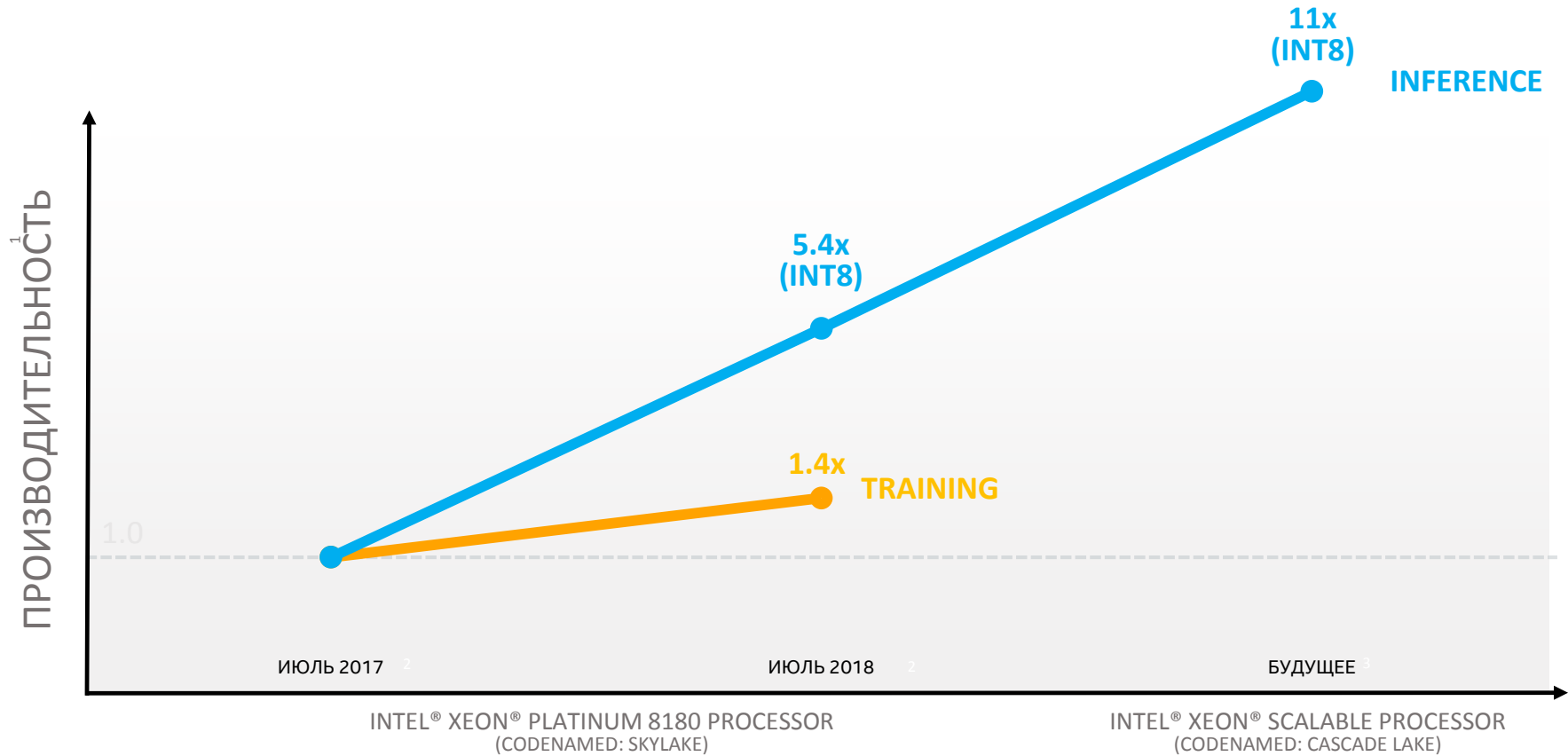


NNP L-1000

Inference

# ПРОЦЕССОРЫ INTEL® XEON® SCALABLE FAMILY

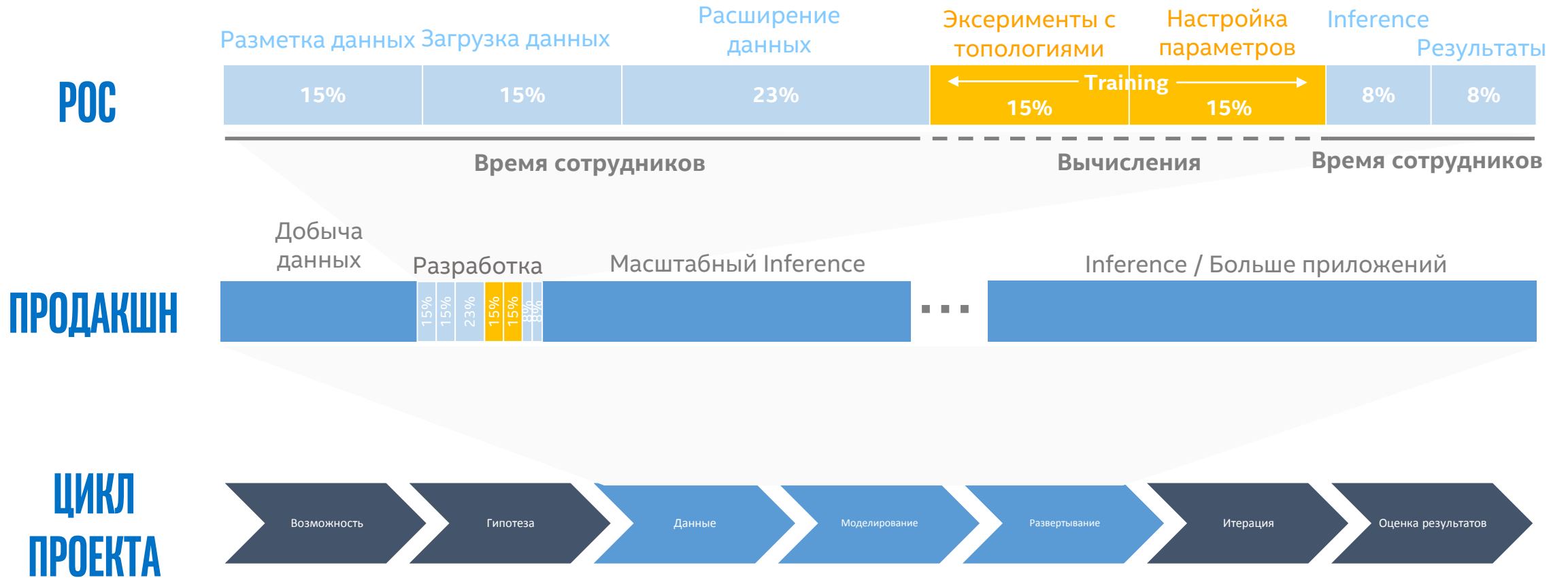
Оптимизация Deep Learning





# ПРИМЕР

Интел взаимодействует с Заказчиком на всех стадиях жизненного цикла ИИ-решения



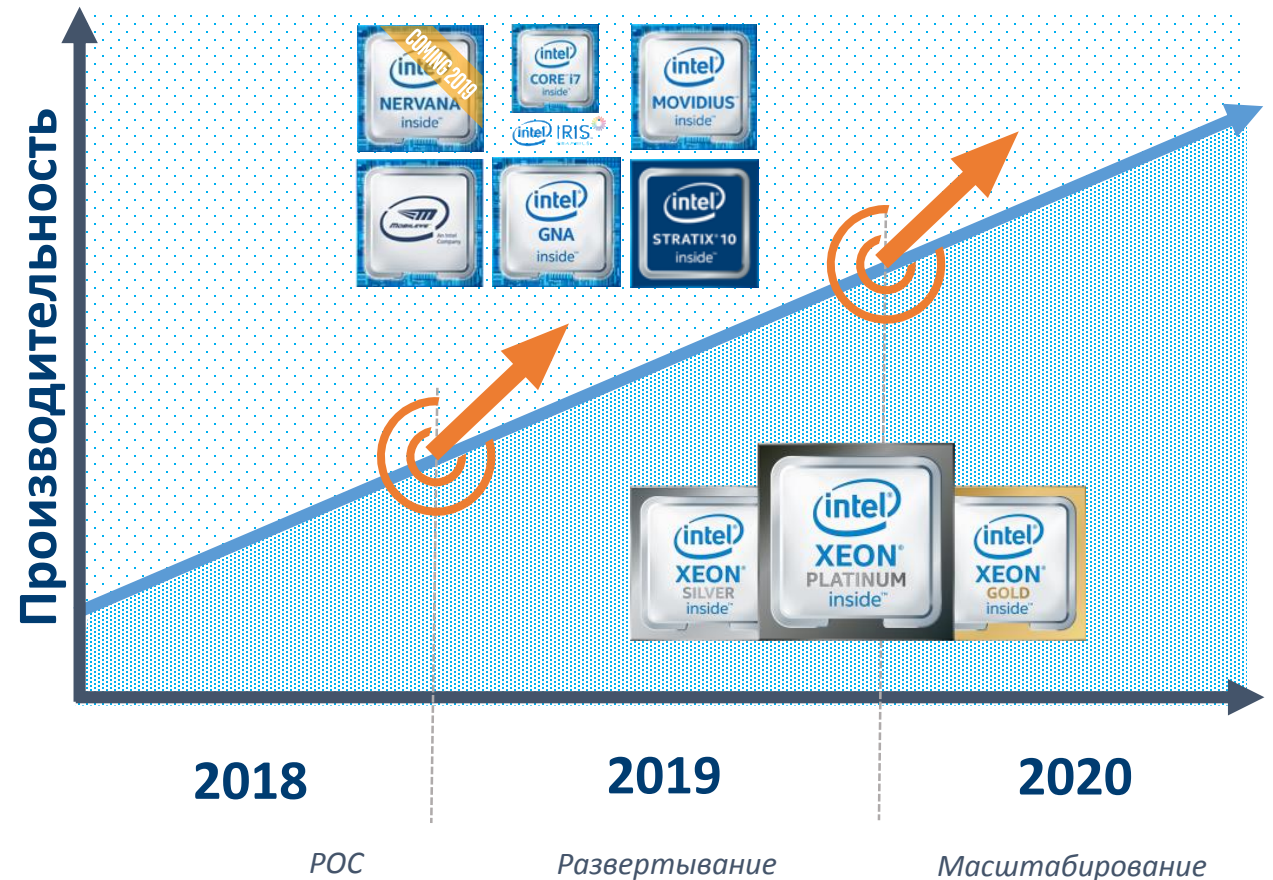


# НУЖНЫ ЛИ УСКОРИТЕЛИ?

Начните с использования/расширения x86 инфраструктуры, в которую вы уже инвестировали

**МИФ:** для ИИ обязательно нужны ускорители.

**РЕАЛЬНОСТЬ:** большинство компаний используют инфраструктуру на базе x86 для Deep Learning, и только **некоторые** из них **смогут** достичь тех требований по производительности, когда использование ускорителей экономически обосновано



# РЕСУРСЫ

Intel® AI academy

Для разработчиков,  
студентов,  
преподавателей

[software.intel.com/ai](https://software.intel.com/ai)

AI builders: Библиотека  
решений

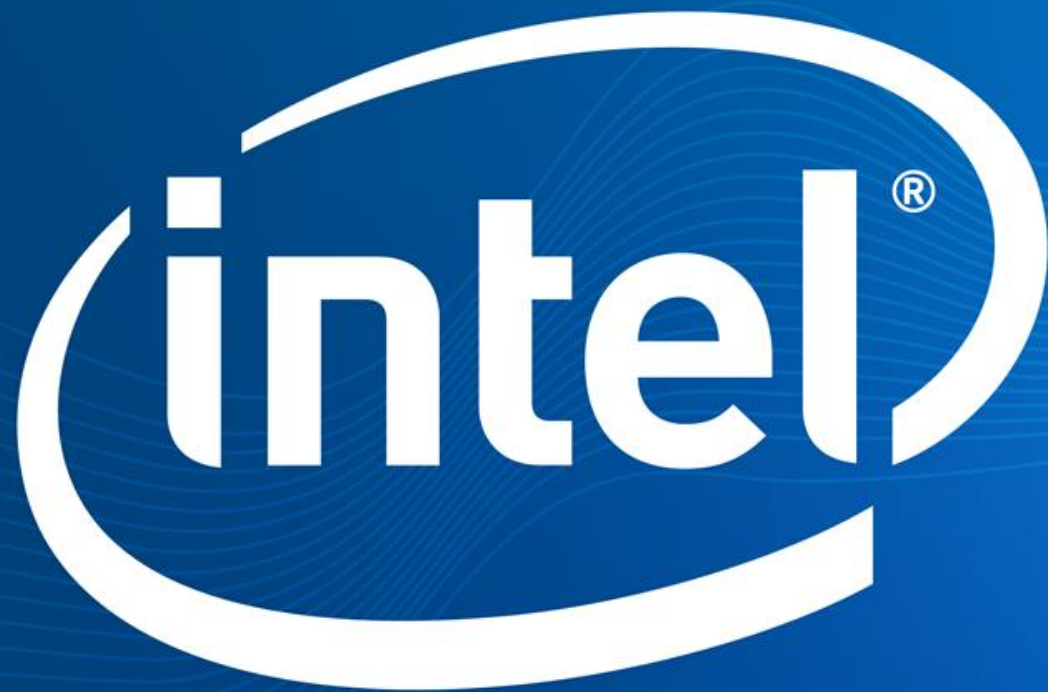
30+ публикаций о  
решениях

[Builders.intel.com/ai/solutionslibrary](https://builders.intel.com/ai/solutionslibrary)


AI builders: Экосистема

100+ партнеров

[Builders.intel.com/ai](https://builders.intel.com/ai)





The background is a solid blue color. Overlaid on this are several abstract, wavy patterns of thin lines. On the left, there are yellow-green lines that curve and flow across the frame. On the right, there are orange lines that also curve and flow. A bright, multi-colored lens flare (yellow, orange, red, purple) is positioned on the right side, partially overlapping the text.

**INTEL<sup>®</sup> INNOVATION DAY**  
**В РИТМЕ ТЕХНОЛОГИИ**