A close-up profile of a person's face, looking towards the right. The image is overlaid with a semi-transparent digital interface consisting of white lines, circles, and geometric shapes, suggesting artificial intelligence or data processing. In the top left corner, there is a white graphic element consisting of a vertical line ending in a circle.

AutoML платформа ENOT  
для оптимизации и ускорения  
нейронных сетей  
от компании Exrasoft

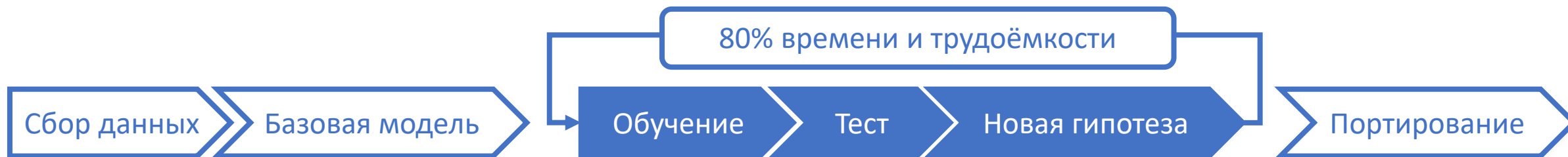
The logo for Exrasoft, featuring a white circle with a dot inside, connected to a horizontal line.

exrasoft\*

# Назначение продукта

AutoML платформа ENOT позволяет автоматизировать процесс разработки AI-решений, а именно **автоматически подбирать и оптимизировать структуру нейронных сетей** для достижения оптимальных значений качества и скорости работы AI-решений. Результаты внедрения:

- стандартизация процесса разработки – сокращение трудоёмкости разработки **в 2 раза**, снижение производственных рисков
- существенное снижения потребления аппаратных ресурсов (**от 3 до 25 раз**).



ENOT предназначен для компаний, ведущих **разработку ML-решений** с помощью внутренних команд или с привлечением партнёров

# Проблема продуктивизации нейронных сетей

В последние годы **глубокие нейронные сети** зарекомендовали себя в области автоматизации когнитивных задач

Выросла доступность инструментов для ML-разработки, при этом **стоимость владения ML-решениями** остаётся высокой

Перед запуском решения в промышленную эксплуатацию требуется **оптимизация нейросети под аппаратную платформу**

Разработчики тратят **4-6 месяцев** на проведение сотен экспериментов в рамках оптимизации, полученные результаты зачастую не являются оптимальными.

# Оптимизация (тюнинг) нейросетей с платформой ENOT

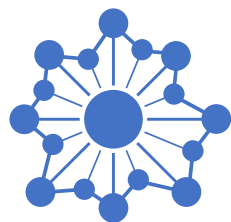
Применение платформы ENOT **позволяет ускорить базовую нейросеть в 3-25 раз** в зависимости от типа решаемой задачи

Платформа в автоматическом режиме поддерживает **все основные методы оптимизации глубоких нейронных сетей** такие как поиск оптимальной архитектуры, дистилляция, структурный прунинг, квантование

ENOT поставляется в виде программной библиотеки для PyTorch и применим **для любых типов процессоров** – x86/x64, GPU, ARM, FPGA

Используемый инновационный подход позволяет **сократить процесс оптимизации с нескольких месяцев до двух недель** и при этом значительно снизить использование вычислительных ресурсов в рамках оптимизации.

# Партнёры и клиенты



ENOT интегрирован с ведущими российскими облачными платформами для ML-разработки **Yandex DataSphere** и **SberCloud ML Space** и доступен в базовых версиях для всех пользователей

Во 2-ом полугодии 2020 года ENOT стал доступен для установки в локальной инфраструктуре заказчиков. Среди клиентов использующих пробную версию платформы технологические компании **LG Electronics** и **PicsArt**



**Yandex** Cloud



**SberCloud**



**LG**



**SONY**



**Himax**



**HUAWEI**

# Примеры оптимизации популярных открытых нейросетей

Imagenette	Accuracy	MMAC	Acceleration
MN_v2 (baseline)	0.89	299	1
ENOT_MN_v2_1	0.89	100	3
ENOT_MN_v2_2	0.88	75	4

CIFAR-100	Accuracy	MMAC	Acceleration
MN_v2 (baseline)	0.68	54.4	1
ENOT_MN_v2_1a	0.68	27.5	2
ENOT_MN_v2_2a	0.67	17.3	3

Pascal VOC	mAP	MMAC	Acceleration
MN_v2_SSD-lite (baseline)	0.65	1340	1
ENOT_MN_v2_SSD-lite_1	0.64	664	2
ENOT_MN_v2_SSD-lite_2	0.63	527	2.5

Mseg	mIoU	MMAC	Acceleration
HR-net (baseline)	0.44	46256	1
ENOT_HR-net	0.43	17731	2.6

# Кейс: сокращение затрат на оборудование для ИИ

## Задача

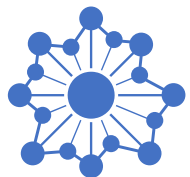
- 8 производственных объектов, каждый из которых включает 7 распределённых площадок
- На каждой площадке 25 видеокамер
- Необходимо контролировать соблюдение правил промышленной безопасности.

## Решение

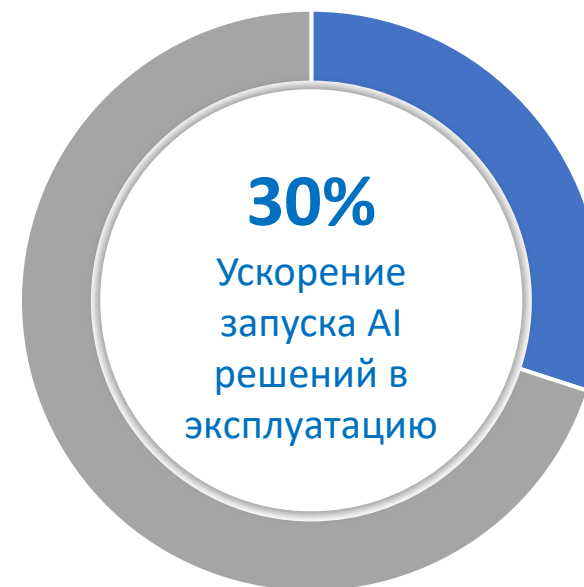
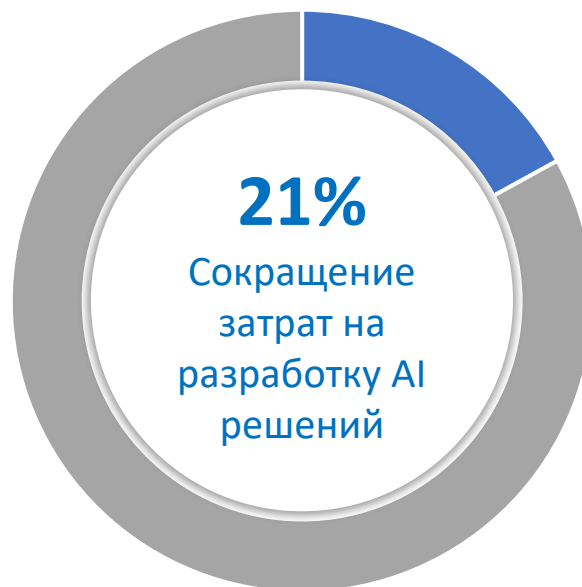
- Система видеоаналитики на основе 2 нейросетей (детектирование человека в кадре, определение позы человека), 5 кадров в сек.
- NVIDIA Tesla V100 (время работы одной нейросети 0,01 секунды, цена 750 000 руб.).

	Без оптимизации	С оптимизацией (x2.6)
Время обработки одного кадра	0,020	0,008
Требуемое число Tesla V100 на одну площадку	3	1
Требуемое число Tesla V100 на все объекты	168	56
Бюджет на оборудование	126 млн. руб.	42 млн. руб.

# Выгоды от внедрения ENOT



ENOT позволяет в 2 раза **сократить трудоёмкость** процесса разработки AI-решения и, как следствие, снизить затраты на разработку и сократить время запуска AI-решений в эксплуатацию (time to market)\*



\* На диаграммах приведена оценка выгод для команды из 9 ML-разработчиков.



# О компании Exrasoft

Компания с 2010 года успешно работает на рынке решений искусственного интеллекта, имеет более 100 реализованных проектов для более чем 10 отраслей

## Кто мы?

---

Компания основана в **2010** году группой молодых учёных из Института математики им. С.Л.Соболева СО РАН

Головной офис компании находится в **г. Новосибирск**. Ещё два офиса расположены в Москве и Риге (Латвия)

В штате компании **3** кандидата ф.-м. наук, **80%** сотрудников – специалисты в области **data science** (с высшим математическим образованием)

## Что мы делаем?

---

Разрабатываем и поставляем **платформенные продукты** на основе технологий компьютерного зрения (CV) и обработки естественного языка (NLP)

Оказываем **услуги заказной разработки и консалтинга** в области машинного обучения

**Обучаем** специалистов анализу данных и прикладному машинному обучению

## Почему мы?

---

Уникальность наших разработок в области нейронных сетей, подтверждена независимыми тестами и **победами в международных конкурсах\***

За годы работы мы сформировали **широкий портфель решений и технологий**

Мы находим оптимальные форматы сотрудничества с компаниями различного профиля – от предоставления готовых **SaaS** сервисов до **ускорения нейронных сетей**

\* Команда Exrasoft заняла два первых места на **IEEE International Low-Power Image Recognition Challenge 2018** (<https://rebootingcomputing.ieee.org/lpirc/2018>). Модель распознавания лиц от Exrasoft прошла тестирование в рамках инициативы Национального института стандартов и технологий США **NIST Ongoing Face Recognition Vendor Test** ([https://pages.nist.gov/frvt/reports/11/frvt\\_report\\_2020\\_01\\_21.pdf](https://pages.nist.gov/frvt/reports/11/frvt_report_2020_01_21.pdf))

# Контакты

Дмитрий Ефремов, директор по развитию бизнеса

- моб. +7 962 987 11 40
- [d.efremov@expasoft.com](mailto:d.efremov@expasoft.com)
- [linkedin.com/in/efremovdmitry](https://www.linkedin.com/in/efremovdmitry)

Сергей Алямкин, технический директор

- моб. +7 913 954 63 66
- [s.alyamkin@expasoft.com](mailto:s.alyamkin@expasoft.com)
- [linkedin.com/in/sergey-alyamkin-5489777a](https://www.linkedin.com/in/sergey-alyamkin-5489777a)