

Open Source Licensing Course

ODS Course Fest #1, 2021

ods.ai | profiscope.io
Code Mining track



whoami & howeare



Алексей Смирнов

- ~20 лет в коммерческой разработке
- ~15 лет с питоном и >13 лет его преподаю
- основал profiscope.io
 - технический аудит
 - CodeScoring.com
 - Code Mining @ods.ai, tg: [@codemining](https://t.me/codemining)



Представление курса

О чем этот курс?

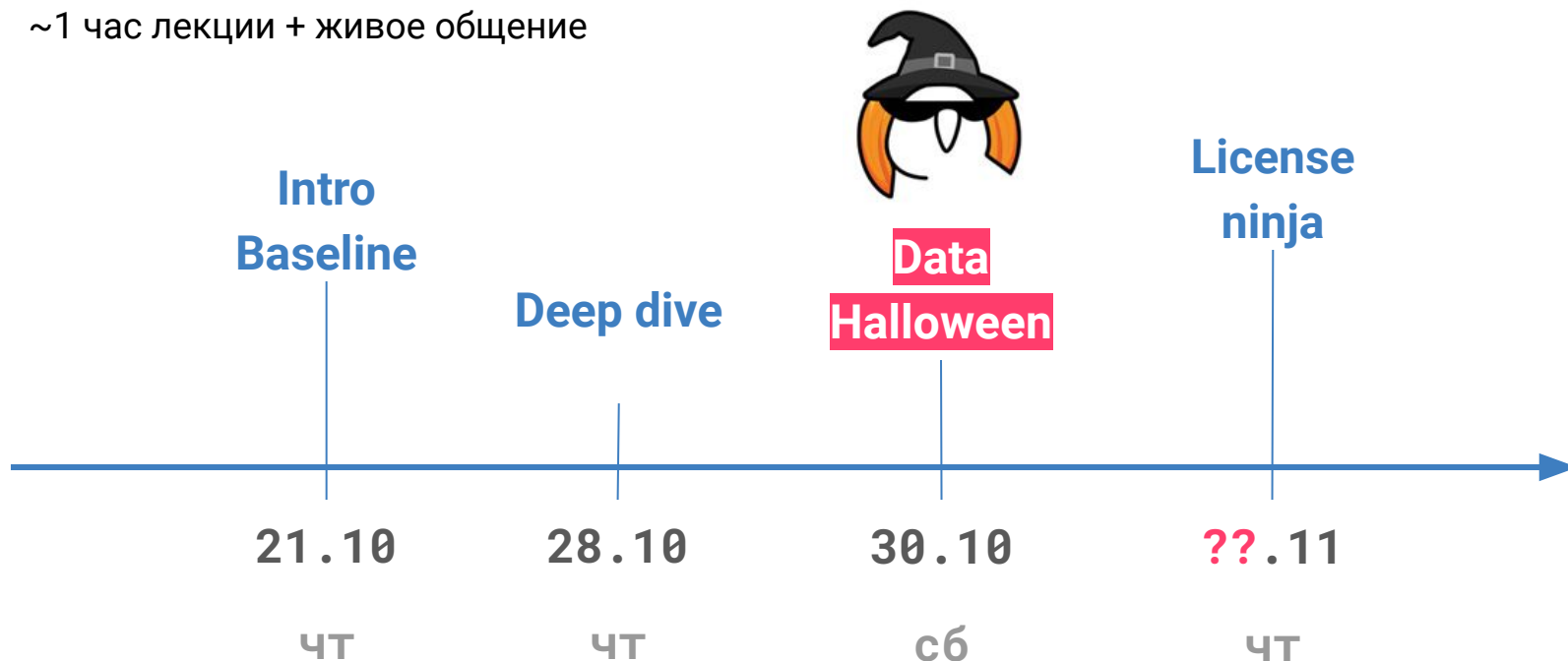
- Что это вообще такое, ~~открытые~~ свободные лицензии?
- Какие особенности у лицензий в Data Science сообществе?
- Как их учитывать в своей работе?
- Как правильно выбрать лицензию для своего проекта?
- К чему может привести некорректный выбор лицензии?

Кому полезен этот курс?

- Data science engineer
- Data engineer
- Developer
- Lawyer
- DevSecOps
- PM/PO/Lead

Timeline

~1 час лекции + живое общение



Роль лицензий в мире безопасной разработки

Встреча — 21.10.2021.

Вводная часть:

- введение в лицензирование СПО
- лицензии: группировка и типизация
- как воспринимать пункты лицензий
- выбор лицензии для DS/ML-проекта
- лицензионная совместимость
- безопасная разработка и лицензирование
- домашнее задание

Тонкости лицензионной совместимости на примерах

Встреча — 28.10.2021.

- тонкости лицензионной совместимости
- двойное лицензирование, как частный случай множественного лицензирования
- двойное лицензирование как способ коммерциализации?
- тонкости и кейсы совместимости лицензий
- разбираем примеры несовместимости лицензий в популярных библиотеках
- совместимость с российским законодательством

Лицензионный ландшафт разработки в Data Science

Встреча — ?? .11.2021.

- Картина лицензирования инструментов в Data Science: для разных языков и направлений деятельности.
- Какие свободные и несвободные инструменты могут помочь с отслеживанием лицензионной чистоты проектов и как с ними работать в рамках CI/CD.
- Подведение итогов по домашним заданиям.



Введение в лицензирование СПО

Что такое лицензия свободного ПО?

Лицензионное соглашение, согласно которому,
автор предоставляет пользователям **свободу** выполнять,
копировать, распространять, изменять и улучшать программу

Центральные свободы по Столлману

0. Свобода запускать программу в любых целях
1. Свобода изучения работы программы и адаптация её к вашим нуждам. Доступ к исходным текстам является необходимым условием
2. Свобода распространять копии, так что вы можете помочь вашему товарищу
3. Свобода улучшать программу и публиковать ваши улучшения, так что всё общество выиграет от этого. Доступ к исходным текстам является необходимым условием

Как может выглядеть лицензия СПО?

- В виде текстового файла в главной директории проекта и/или его поддиректориях.

Например: LICENSE, LICENSE.txt, license.txt, etc.

- В виде текста сопровождающего файл с исходным кодом.

Пример текста лицензии

```
1 Copyright (c) Django Software Foundation and individual contributors.
2 All rights reserved.
3
4 Redistribution and use in source and binary forms, with or without modification,
5 are permitted provided that the following conditions are met:
6
7     1. Redistributions of source code must retain the above copyright notice,
8       this list of conditions and the following disclaimer.
9
10    2. Redistributions in binary form must reproduce the above copyright
11       notice, this list of conditions and the following disclaimer in the
12       documentation and/or other materials provided with the distribution.
13
14    3. Neither the name of Django nor the names of its contributors may be used
15       to endorse or promote products derived from this software without
16       specific prior written permission.
17
18 THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND
19 ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED
20 WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE
21 DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR
22 ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES
23 (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES;
24 LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON
25 ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT
26 (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS
27 SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.
```



Лицензии:
группировка
и типизация

Коротко о главном: группы лицензий

Очень упрощенное объяснение основных групп лицензий:

- Permissive (разрешительная) — практически не ограничивает свободу действий пользователей
- Copyleft — произведение должно лицензироваться под той же лицензией
- Public domain — общественное достояние, без копирайта

Лицензия влияет на «многое».

Топ-5 лицензий PyPI

№	Название	Группа	% в PyPI
1.	MIT	Permissive	55%
2.	BSD-3-Clause	Permissive	13%
3.	Apache 2.0	Permissive	12%
4.	GPL 3.0	Copyleft	7%
5.	AGPL 3.0	Copyleft	5%

BSD-3 (Berkeley Software Distribution)

Разрешает:

- коммерческое использование
- распространение
- изменение
- личное использование

Требует:

- Упоминания авторства и лицензии в работе

Ограничивает:

- Отказ от ответственности
- Никакой гарантии

~80 вариаций

MIT (Massachusetts Institute of Technology)

Разрешает:

- коммерческое использование
- распространение
- изменение
- личное использование

Требует:

- Упоминания авторства и лицензии в работе

Ограничивает:

- Отказ от ответственности
- Никакой гарантии

~30 вариаций

Apache 2.0

Разрешает:

- коммерческое использование
- распространение
- изменение
- личное использование
- предоставление патентных прав

~8 вариаций

Требует:

- Упоминания авторства и лицензии в работе
- Указывать изменения, внесённые в работу

Ограничивает:

- Никаких обязательств
- Никакой гарантии
- Не передаются права на торговые марки

(A)GPL-3 (General Public License)

Разрешает:

- коммерческое использование
- распространение
- изменение
- личное использование
- предоставление патентных прав

Требует:

- Распространять исходный код вместе с продуктом
- Упоминания авторства и лицензии в работе
- Указывать изменения, внесённые в работу
- Производные продукты необходимо выпускать под той же лицензией
- Распространение “копий” по сети (AGPL)

Ограничивает:

- Отказ от ответственности
- Никакой гарантии

~100 вариаций GPL, AGPL — 7 штук



Разберем
лицензии
по пунктам

Сокращенное изложение лицензии для Tensorflow



tensorflow/tensorflow is licensed under the

Apache License 2.0

A permissive license whose main conditions require preservation of copyright and license notices. Contributors provide an express grant of patent rights. Licensed works, modifications, and larger works may be distributed under different terms and without source code.

Permissions

- ✓ Commercial use
- ✓ Modification
- ✓ Distribution
- ✓ Patent use
- ✓ Private use

Limitations

- ✗ Trademark use
- ✗ Liability
- ✗ Warranty

Conditions

- ⓘ License and copyright notice
- ⓘ State changes

Permissions

Commercial use	Разрешение на использование в коммерческих целях.
Distribution	Разрешение на распространение лицензированных компонент.
Modification	Разрешение на распространение лицензированных компонент.
Patent use	Разрешение на предоставление патентных прав от соавторов.
Private use	Лицензионные материалы могут быть использованы и изменены в частном порядке.

Conditions / начало

Disclose source	Исходный код должен быть доступен при распространении лицензионного материала.
License and copyright notice	Копия лицензии и уведомления об авторских правах должна быть приложена к лицензируемому материалу.
License and copyright notice for source	Копия лицензии и уведомление об авторских правах должны быть включены в лицензируемый материал в исходной форме, но не требуются для двоичных файлов.

Conditions /продолжение

Same license	Модификации должны выпускаться под той же лицензией при распространении лицензионного материала. В некоторых случаях может быть использована аналогичная или смежная лицензия.
Same license (file)	Модификации существующих файлов должны выпускаться под той же лицензией при распространении лицензионного материала. В некоторых случаях может быть использована аналогичная или смежная лицензия.
Same license (library)	Модификации должны выпускаться под той же лицензией при распространении лицензионного материала. В некоторых случаях может использоваться аналогичная или смежная лицензия, или это условие может не применяться к произведениям, использующим лицензионный материал в качестве библиотеки.

Conditions /продолжение

Network use is distribution	Пользователи, взаимодействующие с лицензионным материалом через сеть, получают право на получение копии исходного кода.
State changes	Изменения, внесенные в лицензионный материал, должны быть задокументированы.

Limitations

Liability	Данная лицензия включает ограничение ответственности.
Trademark use	В этой лицензии прямо указано, что она НЕ предоставляет права на товарный знак, хотя лицензии без такого указания, вероятно, не предоставляют никаких неявных прав на товарный знак.
Warranty	В данной лицензии прямо указано, что она НЕ предоставляет никаких гарантий.



Какую лицензию
выбрать для DS/ML
проекта

Apache 2.0

TL;DR

Смена лицензий: Catalyst

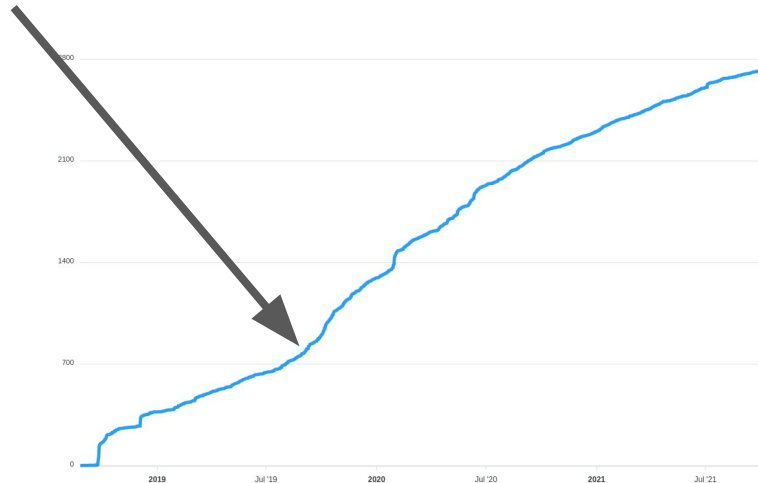
PyTorch-фреймворк для глубокого обучения.

Лицензии: MIT → Apache 2.0, 18.08.2019

Смена лицензии не повлияла
на популярность проекта.

2720 ★★

Catalyst



<https://github.com/catalyst-team/catalyst>

Комментарий от автора

В отличие от MIT, Apache 2.0 требует «указывать внесённые изменения» при использовании доработанного исходного кода.

Кроме этого Apache 2.0 чуть удобнее с точки зрения патентного права и торговых марок.

Поэтому — рекомендую Apache 2.0 для своих проектов 😊



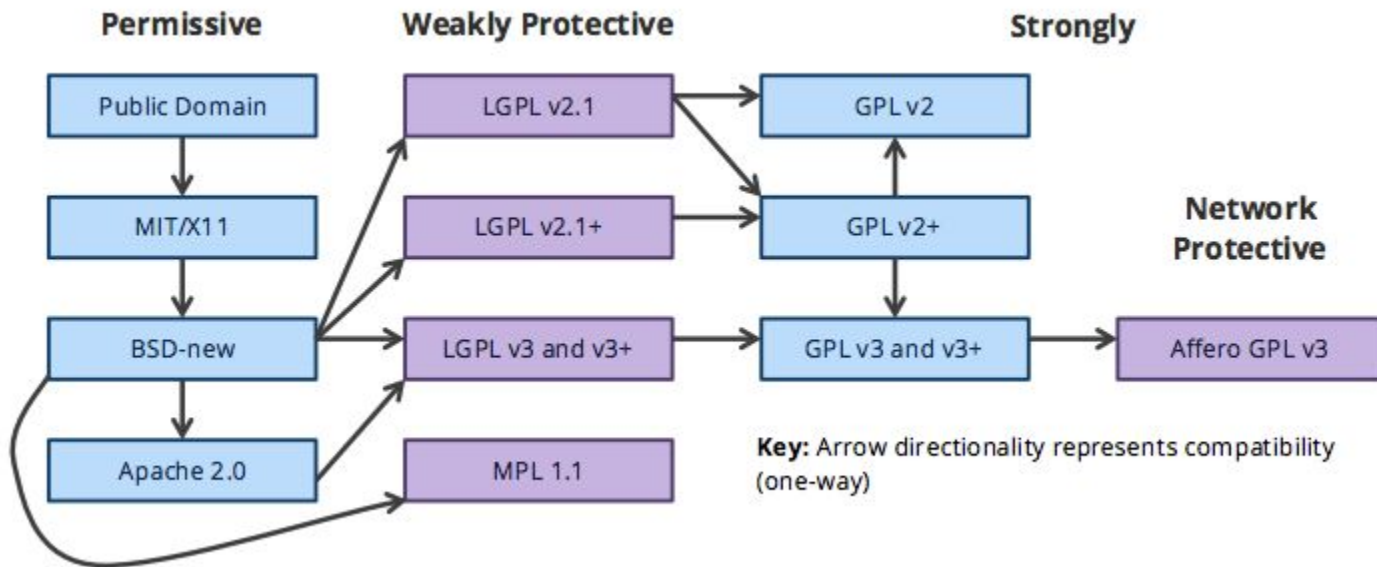
Сергей Колесников

руководитель отдела
исследований Tinkoff,
создатель Catalyst



Лицензионная
совместимость
(license compliance)

Проблема лицензионной совместимости



Совместимость лицензий: ожидание

Is Compatible With:	License-A	License-B	License-C	License-D	License-E	License-F	License-G
License-A	X				X	X	
License-B		X					
License-C			X				
License-D		X		X			X
License-E					X		
License-F			X			X	
License-G	X						X

<https://www.linux.com/news/5-practical-ways-legal-counsel-advise-developers-open-source/>

Совместимость лицензий: реальность

[illegible]

Разобраться по верхам — можно

License	Commercial use	Distribution	Modification	Patent use	Private use	Disclose source	License and copyright notice	Network use is distribution	Same license	State changes	Liability	Trademark use	Warranty
BSD Zero Clause License	●	●	●		●						●		●
Academic Free License v3.0	●	●	●	●	●		●			●	●	●	●
GNU Affero General Public License v3.0	●	●	●	●	●	●	●	●	●	●	●		●
Apache License 2.0	●	●	●	●	●		●			●	●	●	●
Artistic License 2.0	●	●	●	●	●		●			●	●	●	●
BSD 2-Clause "Simplified" License	●	●	●		●		●				●		●
BSD 3-Clause Clear License	●	●	●	●	●		●				●		●
BSD 3-Clause "New" or "Revised" License	●	●	●		●		●				●		●
BSD 4-Clause "Original" or "Old" License	●	●	●		●		●				●		●
Boost Software License 1.0	●	●	●		●		●				●		●
Creative Commons Attribution 4.0 International	●	●	●	●	●		●			●	●	●	●
Creative Commons Attribution Share Alike 4.0 International	●	●	●	●	●		●		●	●	●	●	●
Creative Commons Zero v1.0 Universal	●	●	●	●	●						●	●	●

<https://choosealicense.com/appendix/>



Домашнее задание #1

Ручной анализ совместимости зависимостей

1. Выбрать любой проект
2. Изучить лицензии его зависимостей
3. Подтвердить или опровергнуть лицензионную совместимость

Чтобы не искать, предлагаем изучить:

- <https://github.com/django/django>
- <https://github.com/psf/requests>
- <https://github.com/apache/ignite>

Домашнее задание отправить

alexey@profiscope.io



Безопасная
разработка

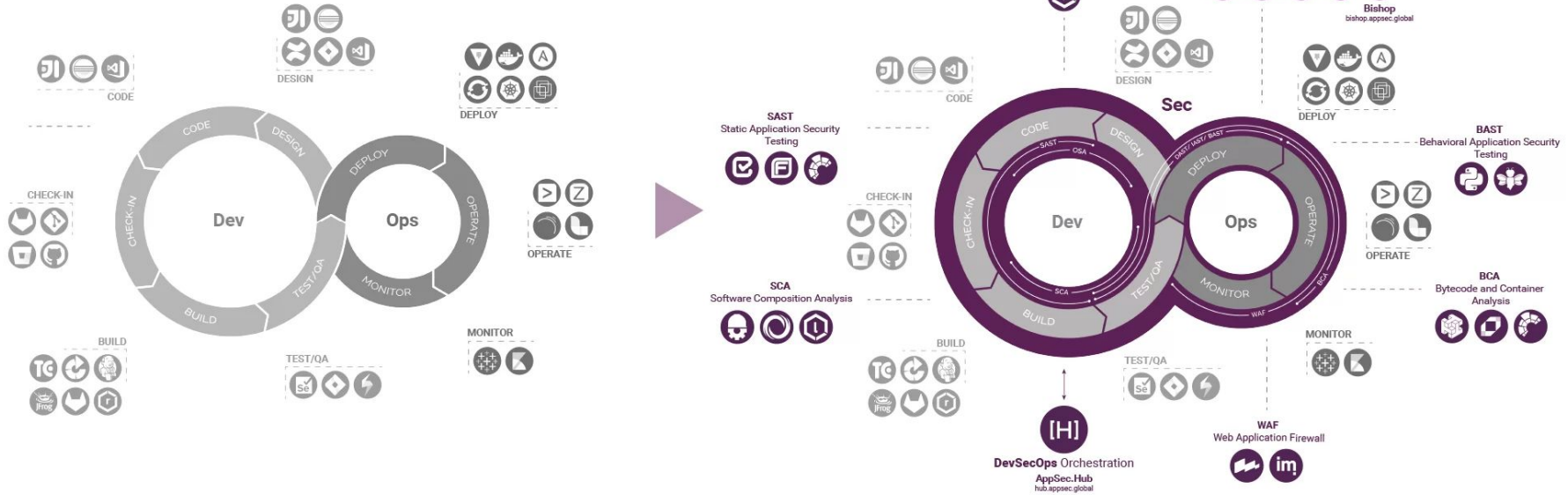
DevSecOps



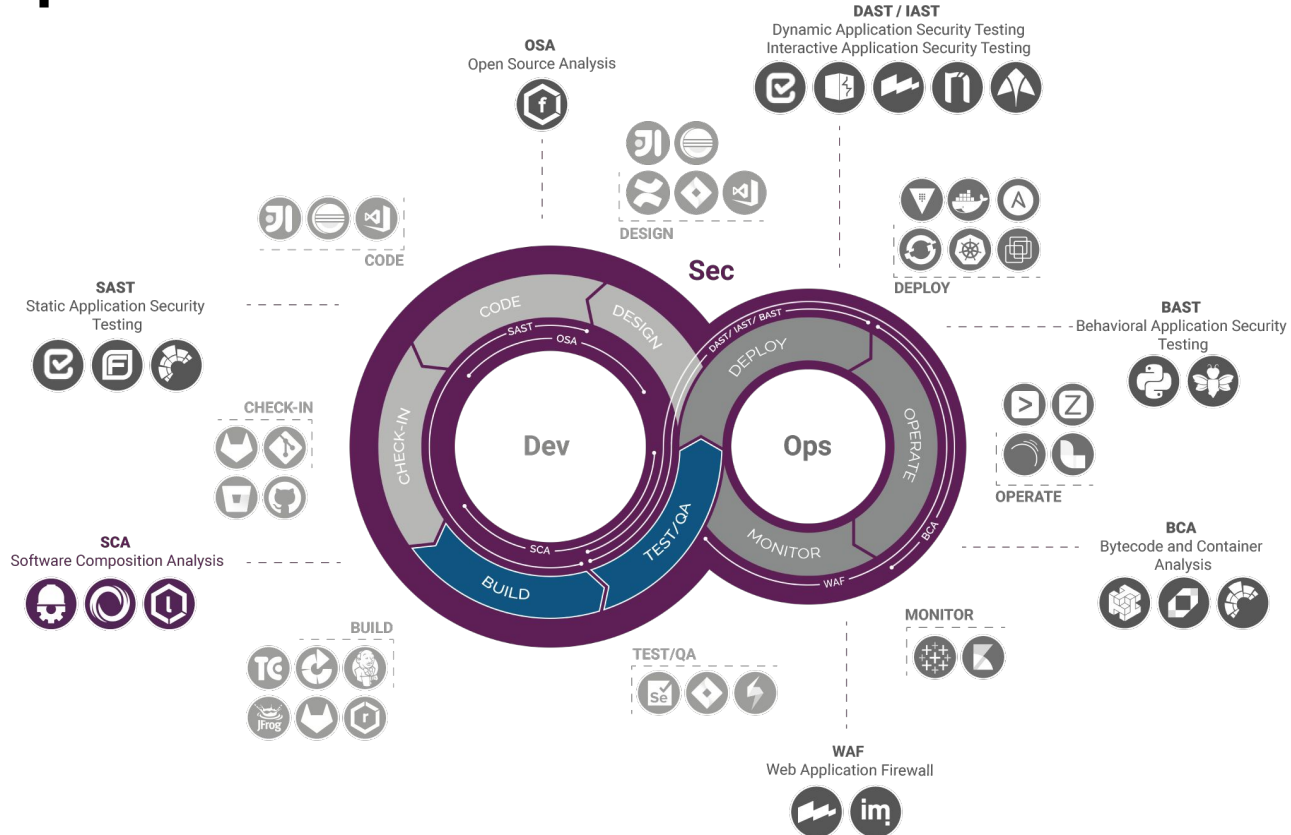
DevSecOps — automates the integration of security at every phase of the software development lifecycle, from initial design through integration, testing, deployment, and software delivery.



DevOps → DevSecOps



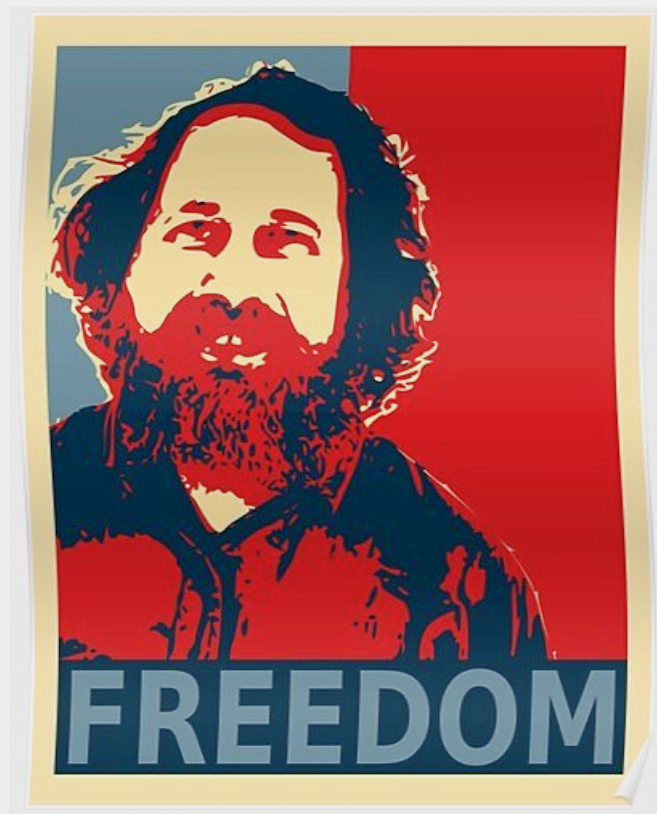
DevSecOps & SCA



И помните

GNU GPL
was not designed to be
«open source»

© Richard Matthew Stallman



Спасибо за внимание!

alexey@profiscope.io

Канал про анализ кода и всего такого:

t.me/codemining

