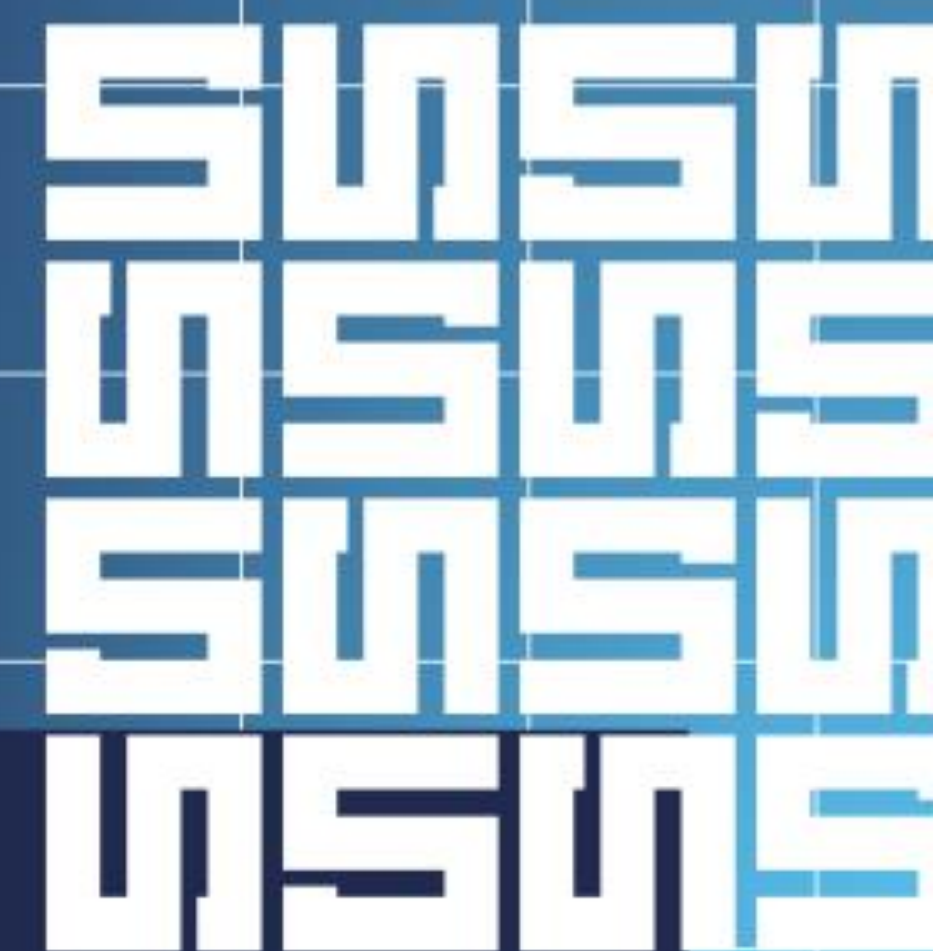




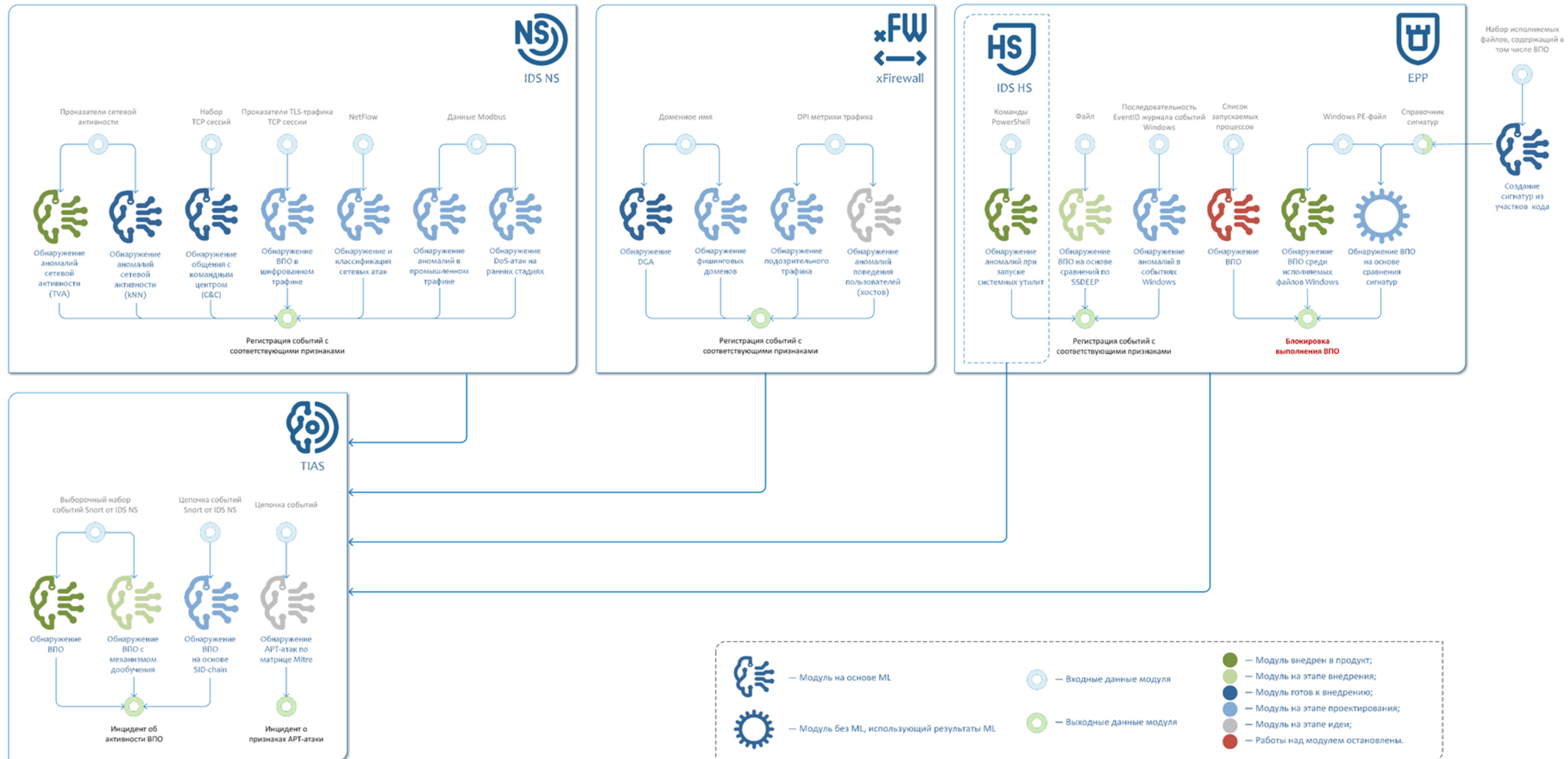
# Технологии SOC



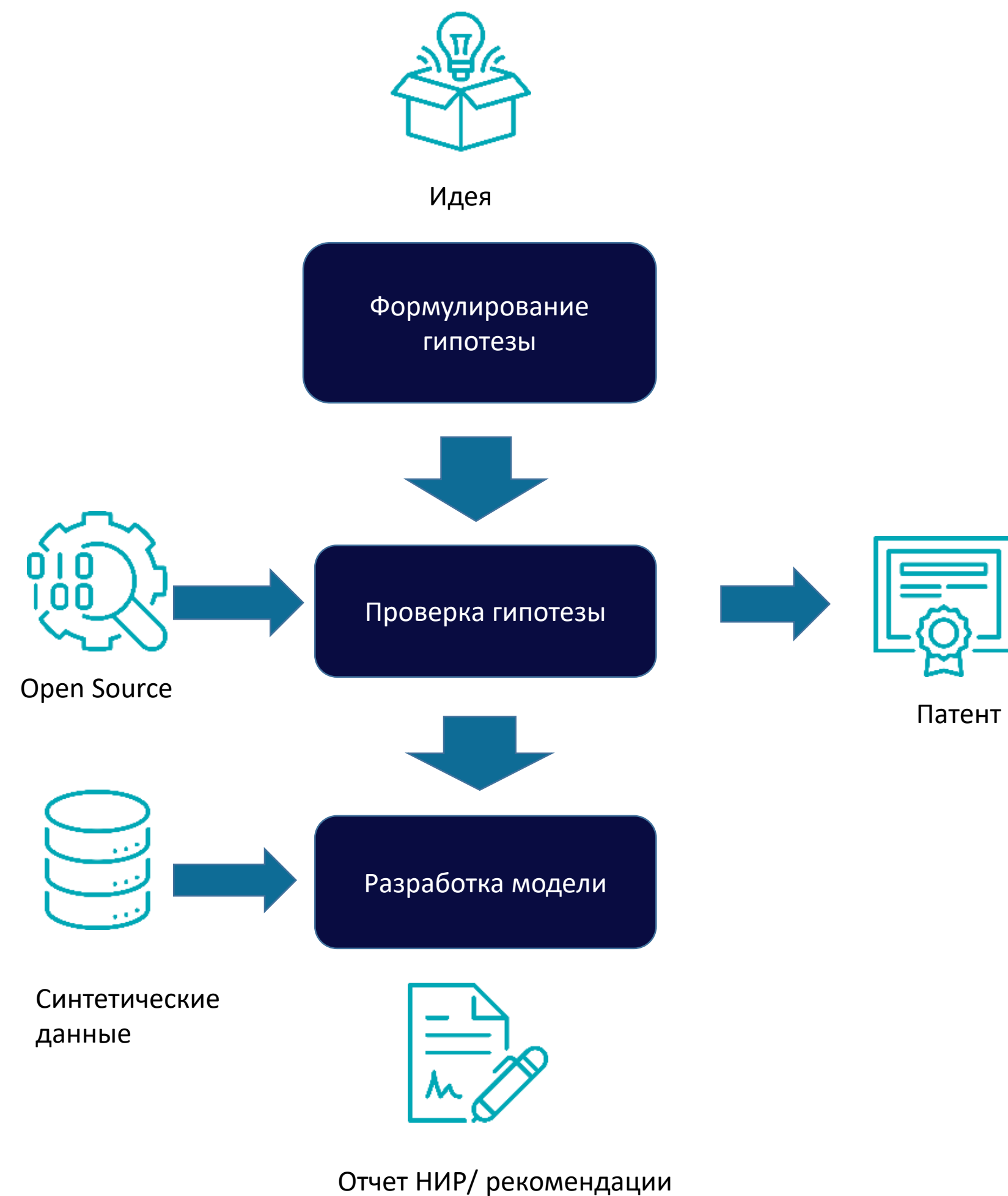
## Проблемы использования технологий машинного обучения в SOC

Светлана Старовойт

# ML-модули в продуктах для SOC



# Процесс разработки

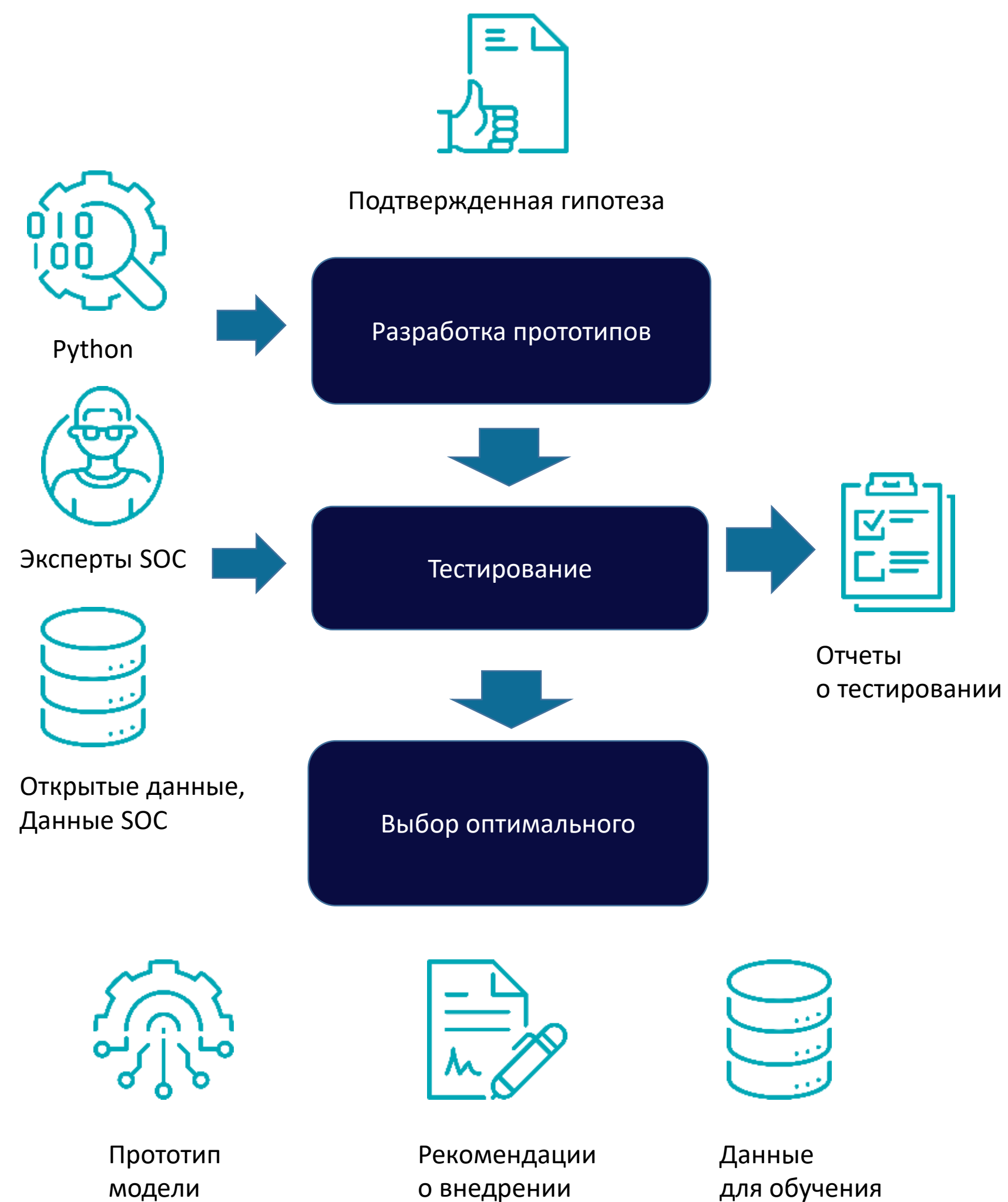


шаг 1:  
проверка гипотезы

- ✓ Привлечение студентов и учёных
- ✓ 2 патента в области ML
- ✓ Использование готовых инструментов и синтетических данных для ускорения процесса



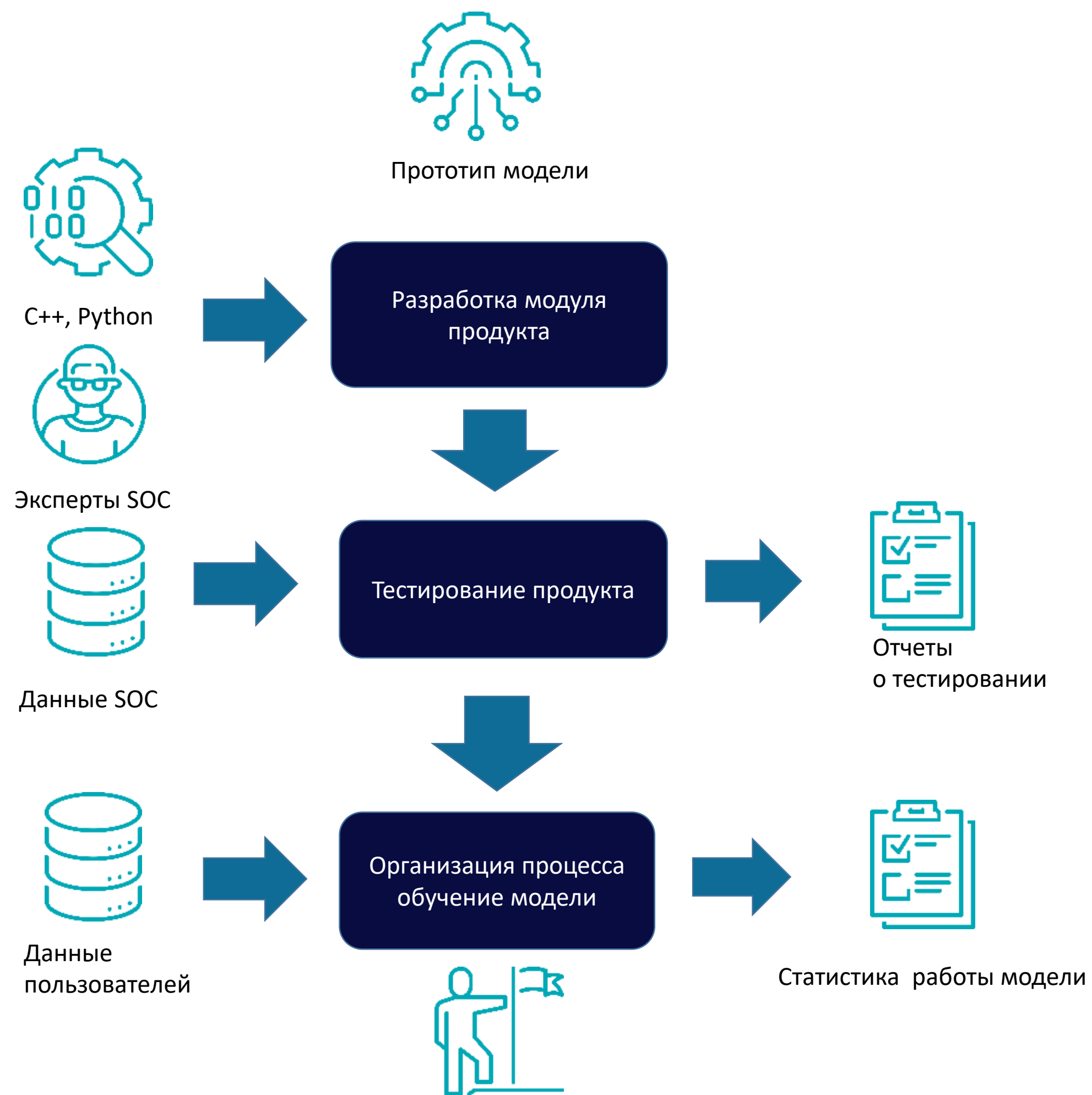
# Процесс разработки



шаг 2:  
прототип

- ✓ Аналитики данных и специалисты по ML
- ✓ Автоматизированные процессы работы с данными
- ✓ Рабочие прототипы для встраивания

# Процесс разработки



**шаг 3:**  
внедрение в продукт

- ✓ Разработчики продуктовых команд и эксперты SOC
- ✓ Только реальные данные
- ✓ Сбора обратной связи о работе модели

# Результаты

00110011  
01011010



- ✓ 40 проверенных гипотез
- ✓ 20 разработанных прототипов
- ✓ 4 внедренных модели



# Проблемы



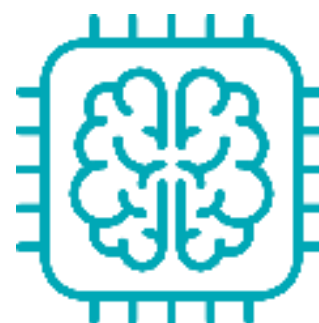
Безопасность



Данные для обучения



Доверие к результатам



Ресурсоемкость (люди и железо)



# Безопасность



- ✓ Атаки на Open Source библиотеки и на данные
- ✓ SDL или MLOps
- ✓ Передача данных третьим лицам
- ✓ Дополнительное регулирование



# Доверие к работе модели



Мне нужна твоя одежда, ботинки и мотоцикл.



✓ Интерпретируемость результата работы модели

✓ Дрейф данных

✓ Доверие при передаче данных



# Данные для обучения и проверки



- ✓ Синтетические данные
- ✓ Открытые источники
- ✓ Данные SOC
- ✓ Данные пользователей
- ✓ Разметка экспертами или пользователями



# Ресурсоёмкость



Специалисты по работе с данными в SOC

Где работает/обучается модель?

- на конечном узле;
- в ЦОДе;
- в облаке.





Благодарю за внимание!

