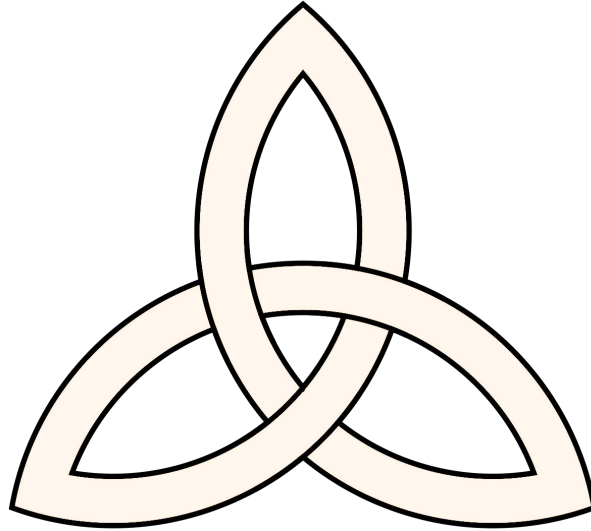




MLOps

Герард Костин

DevOps



Software Engineering

Machine Learning

Software Engineering

От перфокарт... до
многопользовательских UNIX-систем...

К персональным компьютерам...

К сетевым компьютерам...

Отправка файлов исправлений по
электронной почте...

Централизованный контроль версий...

Распределенный контроль версий... Git!
Гитхаб!



DevOps

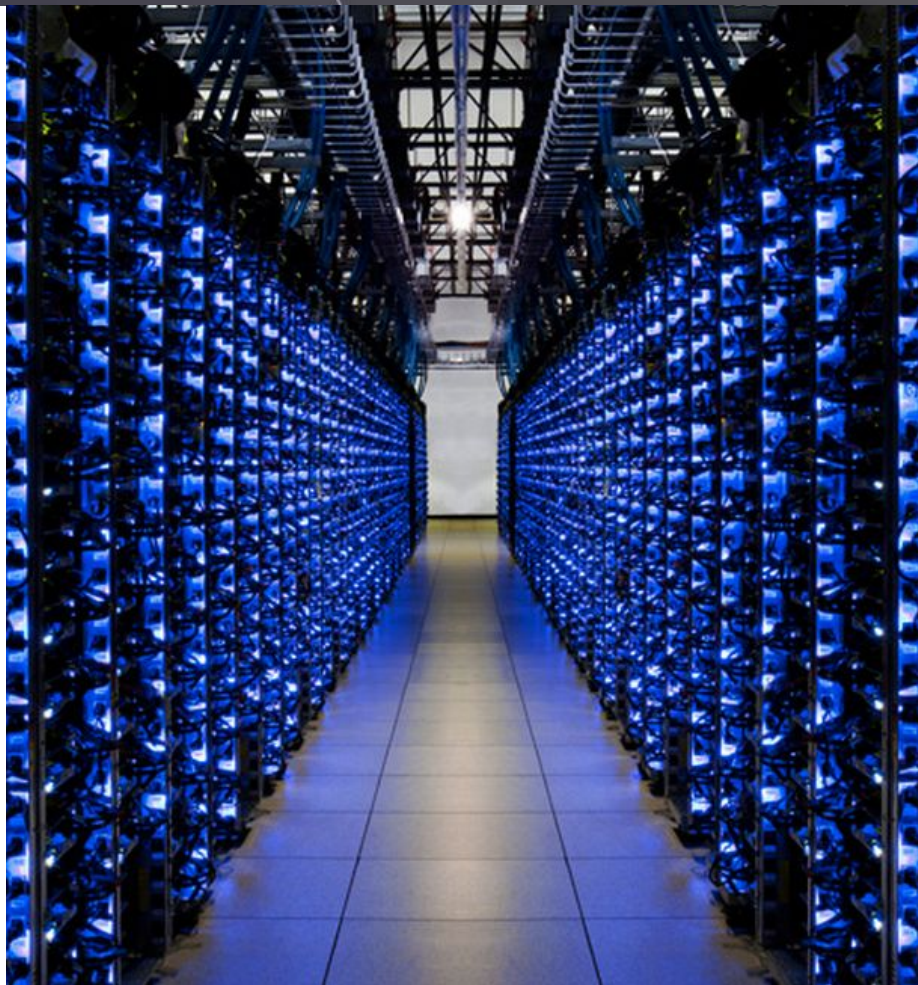
От редактирования кода на сервере...

Для создания двоичных файлов и отправки их по электронной почте...

К непрерывной интеграции...

В общедоступное облако...

К неизменяемой инфраструктуре...
Docker! Kubernetes! GitOps!



Machine Learning (AI)

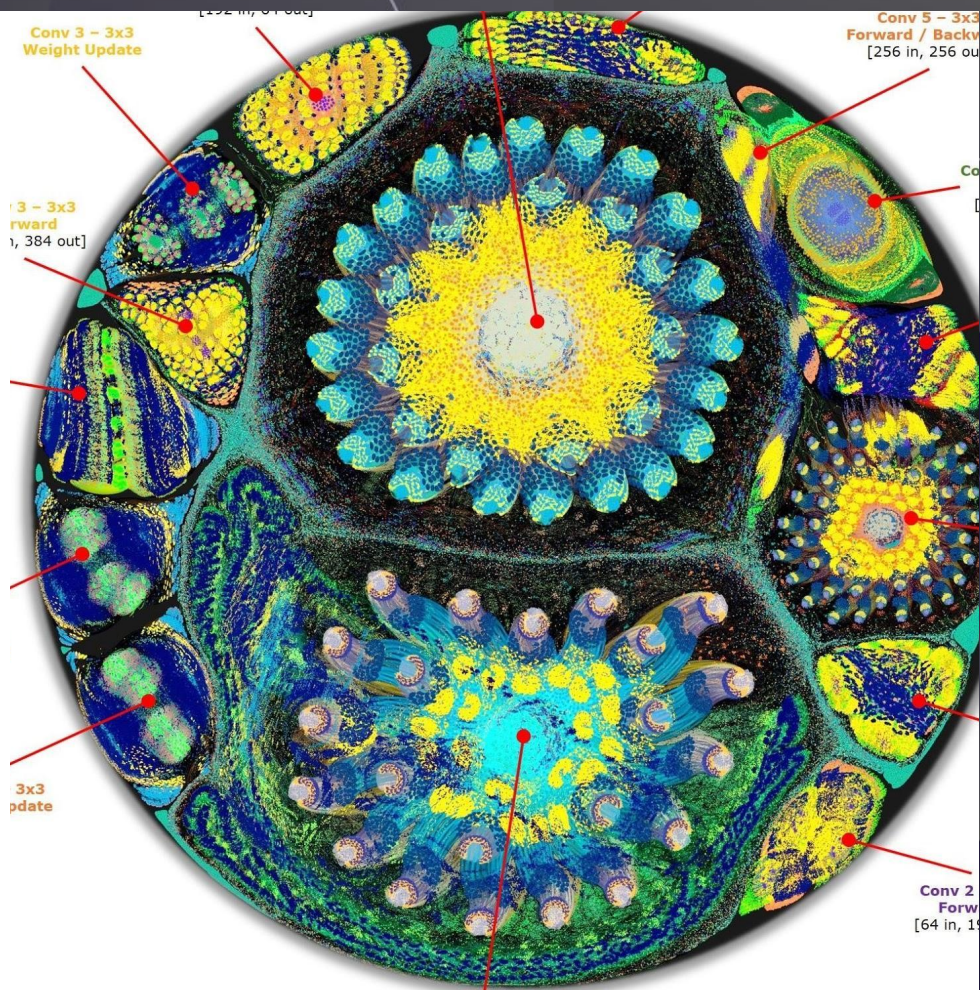
От подгонки кривой до линейной регрессии...

К цепям Маркова... к нейронным сетям...

Повторное открытие обратного распространения ошибки

Глубокое обучение становится вычислительно возможным...

К экспоненциальному росту объемов данных



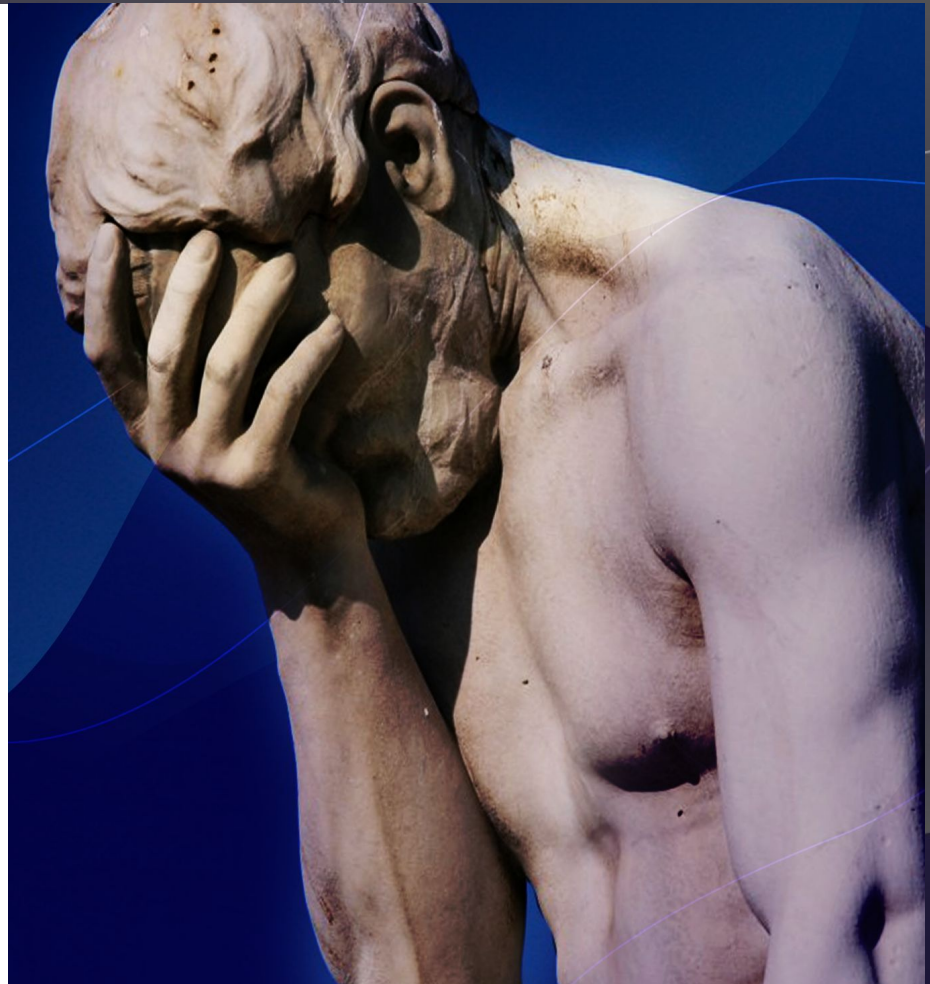
По мере того, как машинное обучение появляется в результате исследований, эти дисциплины должны сойтись воедино...

Конвергенция программной инженерии, DevOps и ML = MLOps.

Ключевые проблемы, влияющие на работу ИИ сегодня

1. Модели заблокированы перед развертыванием
2. Трата времени
3. Неэффективное сотрудничество
4. Ручное отслеживание
5. Нет воспроизводимости или происхождения
6. Неконтролируемые модели

...



Мы были здесь раньше!

В 90-х разработка программного обеспечения была изолирована.

Не все было под контролем версий, непрерывной доставки не было.

На доставку программного обеспечения ушли месяцы; теперь он доставляется за считанные минуты.

То же самое возможно и для ИИ...



Requirements MLOps

1. Reproducible

Должно быть возможно переобучить 9-месячную модель с точностью до нескольких процентов.

3. Collaborative

Ассинхронная коллаборация

2. Accountable

Должен быть в состоянии проследить от модели в проде до ее происхождения

4. Continuous

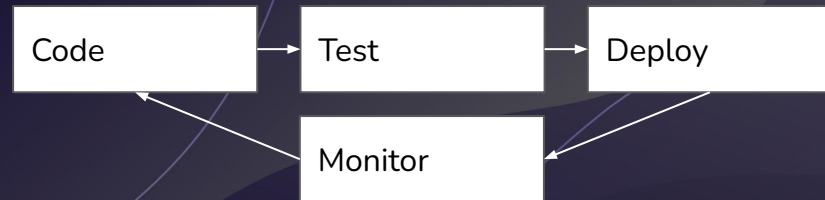
Должна быть возможность автоматического развертывания и статистического мониторинга

Инжиниринг данных и машинное обучение отличаются от разработки программного обеспечения

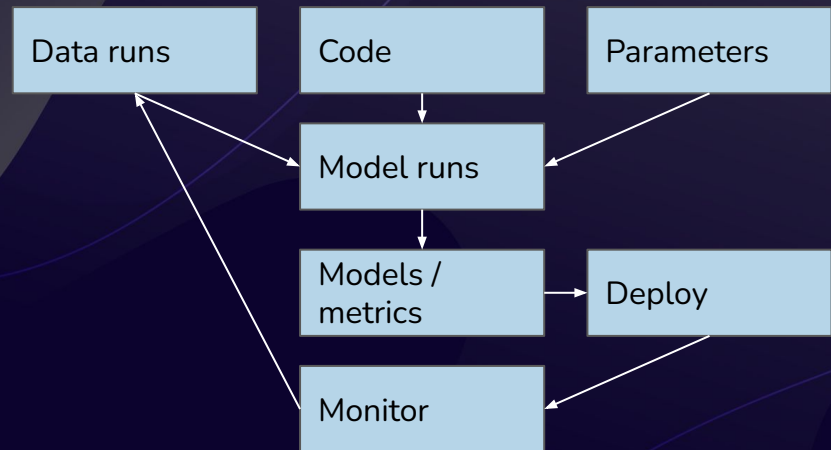
Существуют уникальные способы, которыми работа с данными, кодом и моделями отличается от работы только с кодом.

Это делает существующие программные инструменты недостаточными...

Software Development (DevOps)



Machine Learning (MLOps)

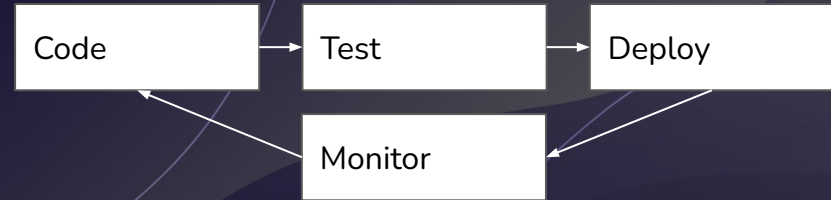


Отслеживая прогоны, вы можете достигать MLOps

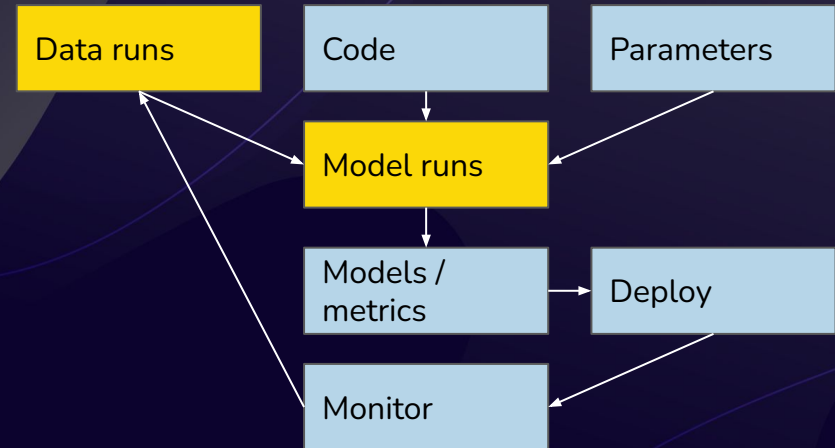
Обычные инструменты разработки программного обеспечения сосредоточены на коммитах кода. Но у машинного обучения гораздо больше движущихся частей: данные, модели и метрики.

Вам необходимо отслеживать прогоны, объединение данных, версии кода и параметров, которые использовались при создании промежуточного набора данных или модели. Это обеспечивает полный контекст для воспроизводимости и происхождения, чтобы связать инженерию данных с обучением модели для отслеживания для обеспечения подотчетности.

Software Development (DevOps)

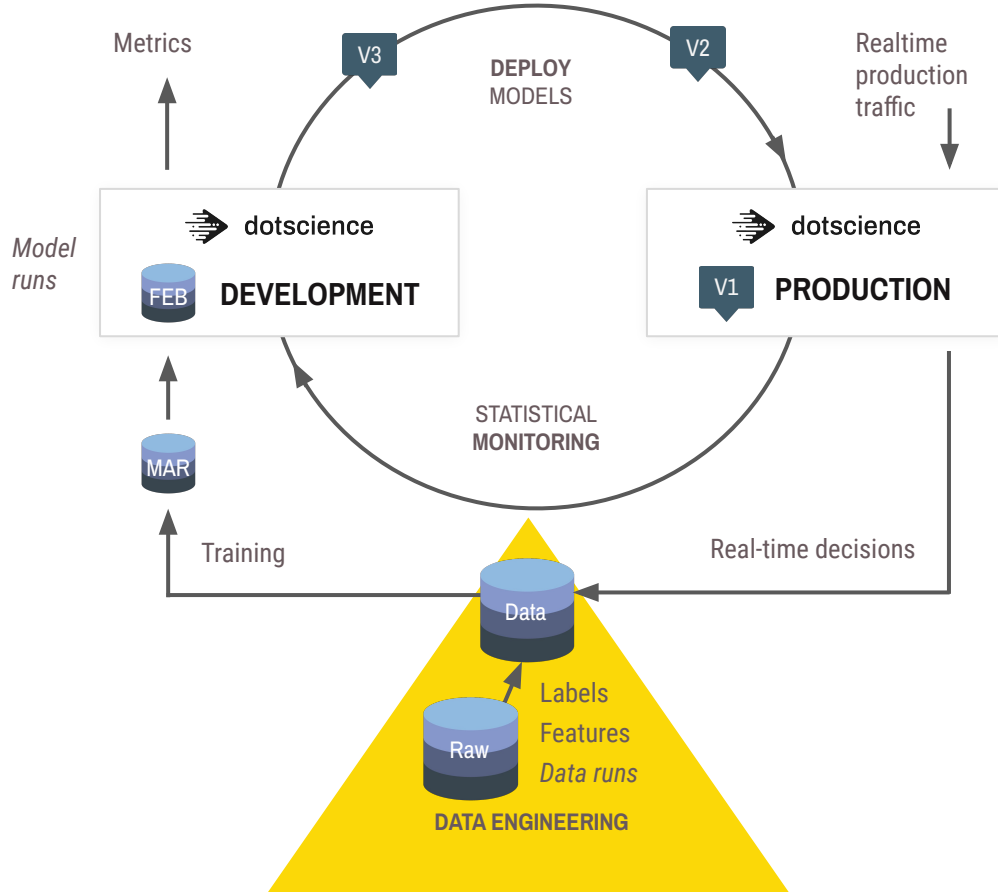


Machine Learning (MLOps)

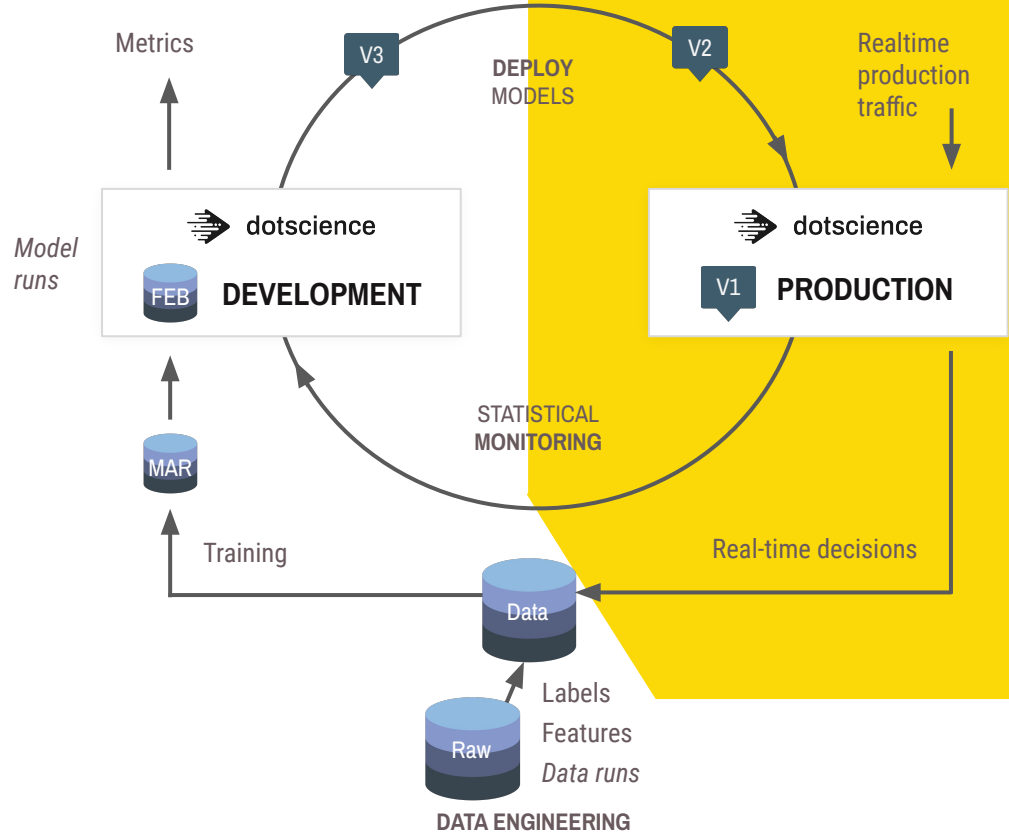




MLOps model lifecycle.

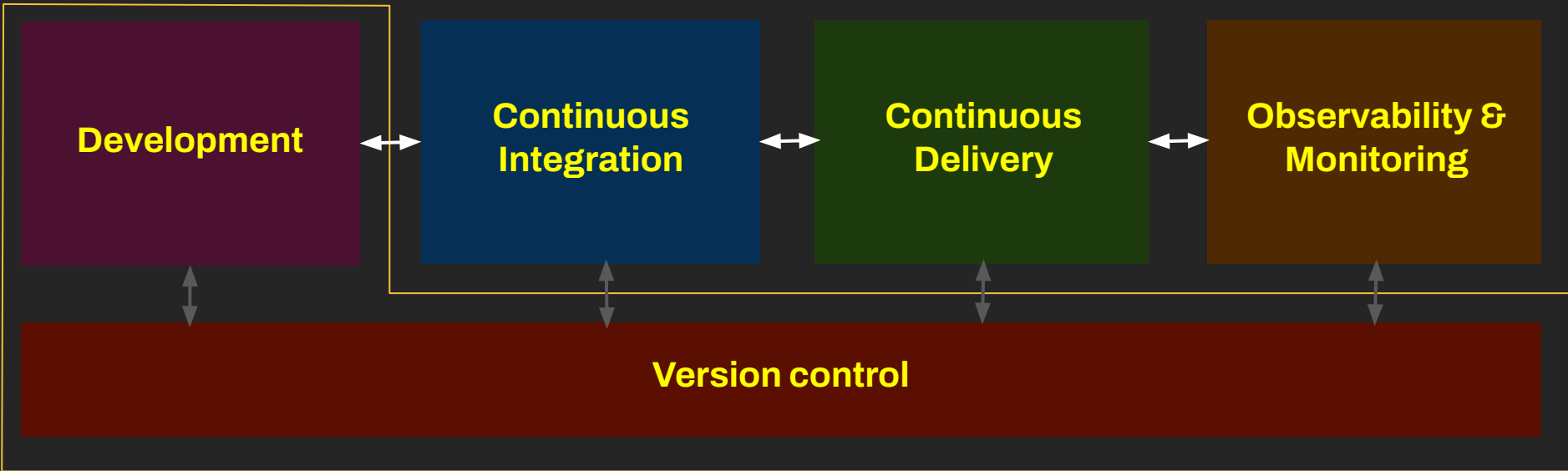


Data Engineering

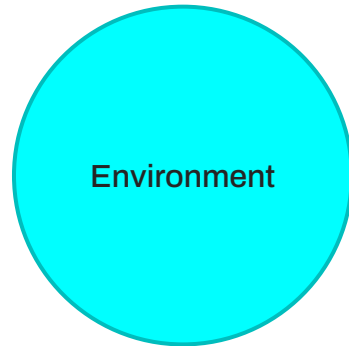


Models in Production

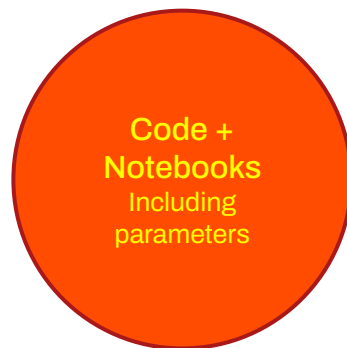
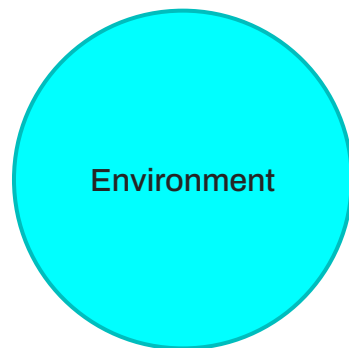
Version control



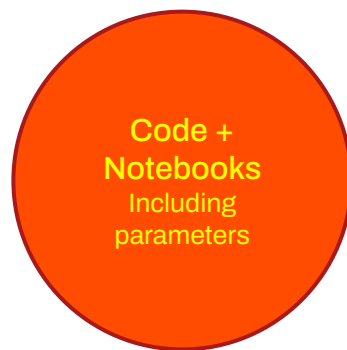
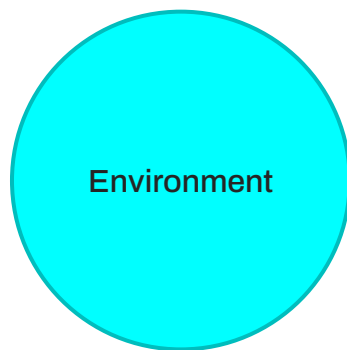
Воспроизводимость!



Воспроизводимость!



Воспроизводимость!



Как?

flow

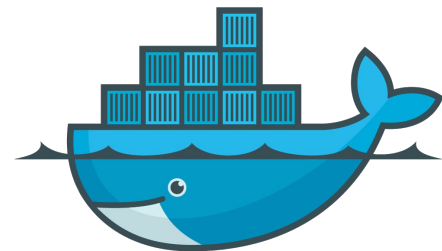
ICEBERG 

 SELDON



 Kubeflow

 **dremio**
The Data Lake Engine



docker

 BENTOML