MLSECOPS 101

Ahmet Akan

$whoami

- development (java, go, c#)
- application security research, code review
- xsecops
- social media
  - ahmetakan.com
  - x.com/ahmetak4n
  - ahmetak4n@gmail.com
  - github.com/ahmetak4n

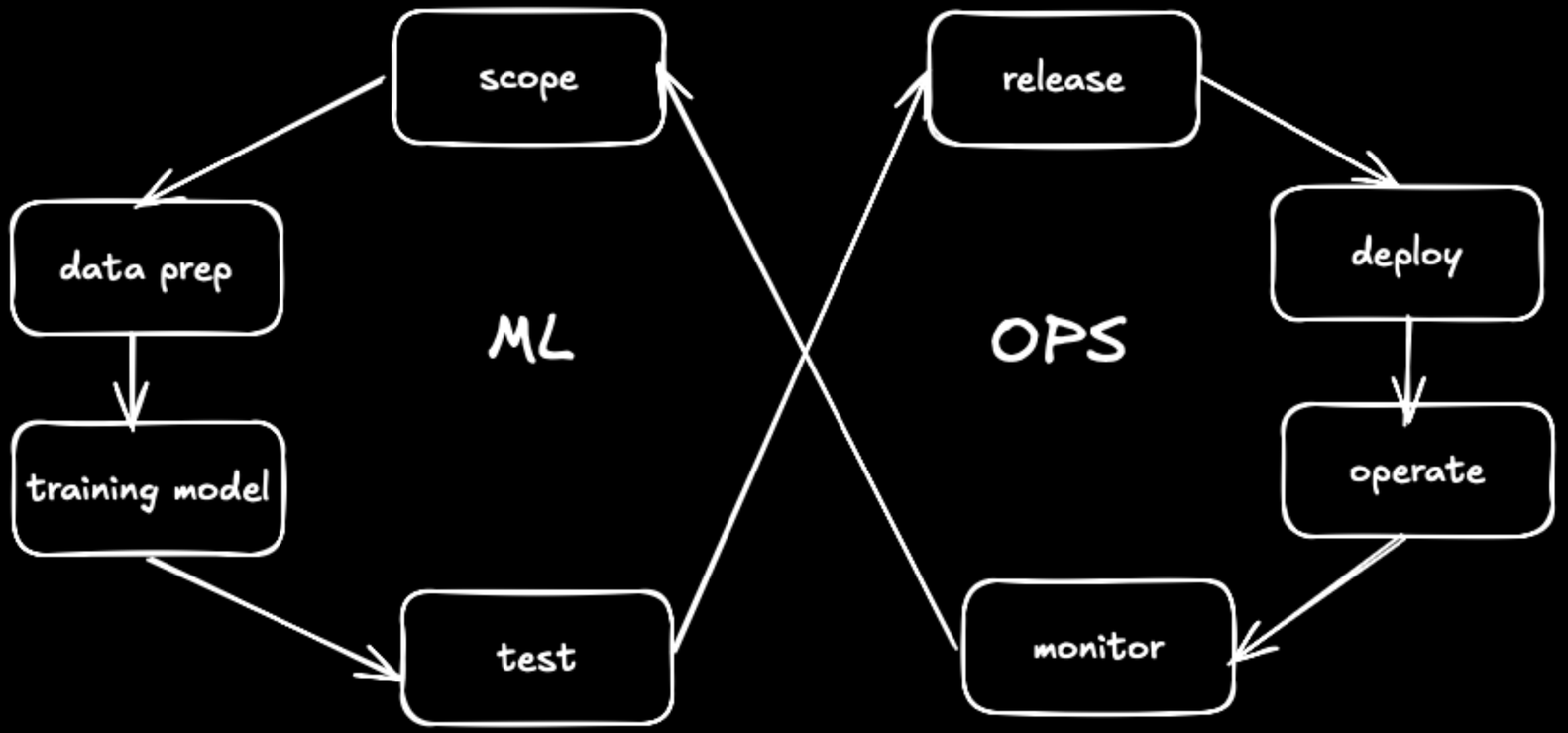
$agenda

- ml / ai
- mlops
- mlsecops
  - supply chain vulnerabilities
  - model provenance
  - governance, risk & compliance
  - trusted ai: bias, fairness & explainability
  - adversarial ml
- references

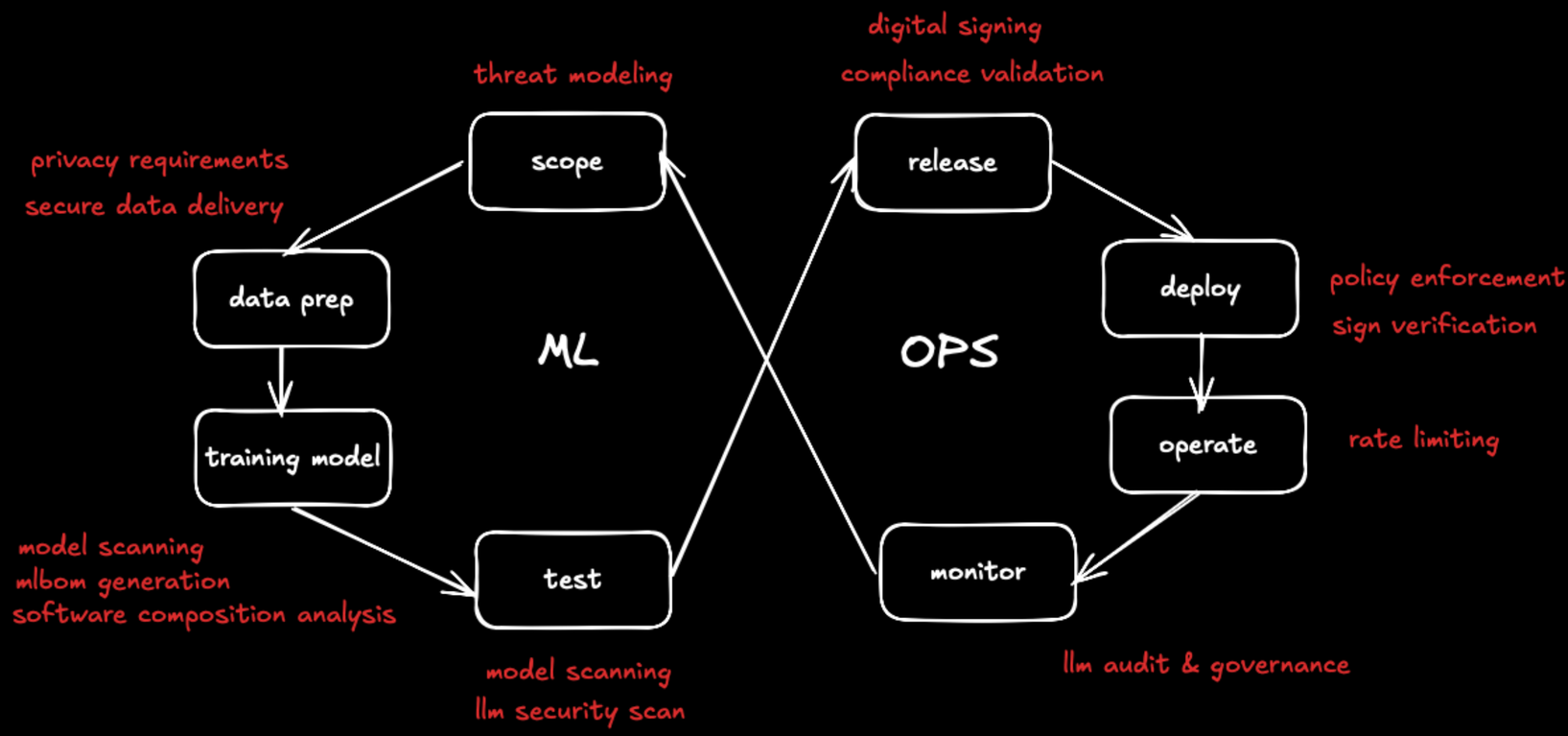
$ml / ai

- ml
  - supervised model
  - unsupervised model
  - reinforcement
- ai
  - speak, listen, writing like a human
- ai <superset> ml

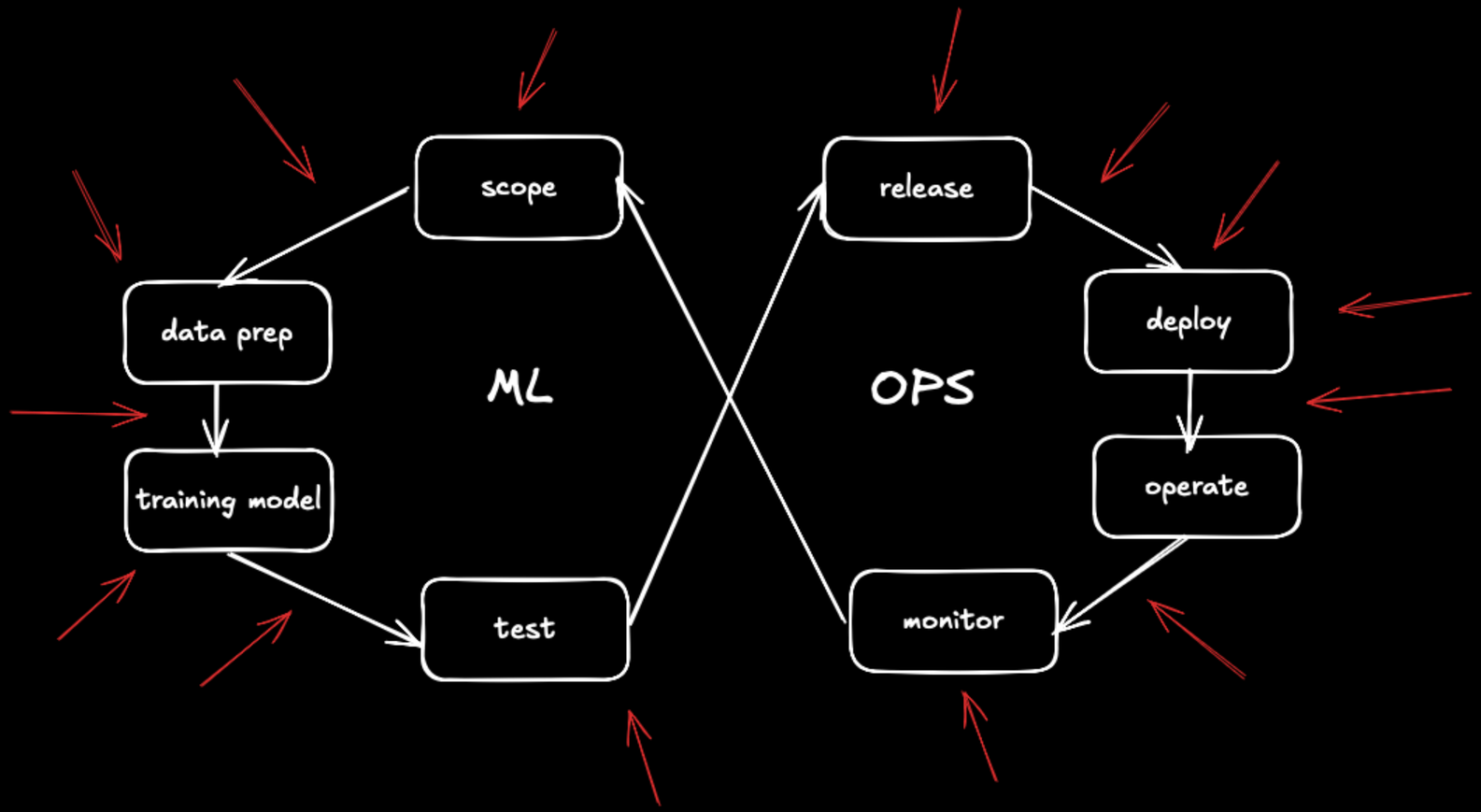
$mlops


scope
data prep
training model
test
ML
release
deploy
operate
monitor
OPS

$mlsecops


digital signing
compliance validation
threat modeling
privacy requirements
secure data delivery
scope
release
data prep
ML
OPS
deploy
policy enforcement
sign verification
training model
operate
rate limiting
model scanning
mlbom generation
software composition analysis
test
monitor
model scanning
llm security scan
llm audit & governance

$supplychain



$provenance

- history of model
  - change trackings
  - who cause changes
  - why / when changed
- gdpr, hippa

$grc

- governance
- risk & compliance
- mlbom
  - algorithms
  - data sets
  - frameworks

$trustedai

- bias
- fairness
- explainability

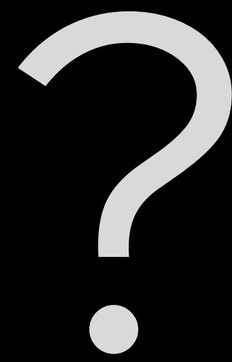
$adversarialml

- manipulating input data
- manipulation of the model
- architecture & access control attacks
- .....

$references

- mlsecops.com
- neptune.ai/blog
- protectai.com/blog
- github.com/RiccardoBiosas/awesome-MLSecOps
- huntr.com

$qa




thanks for listening