Co-funded by the European Union

CHESS

Cyber-security Excellence Hub in Estonia and South Moravia

The long road to autonomous security

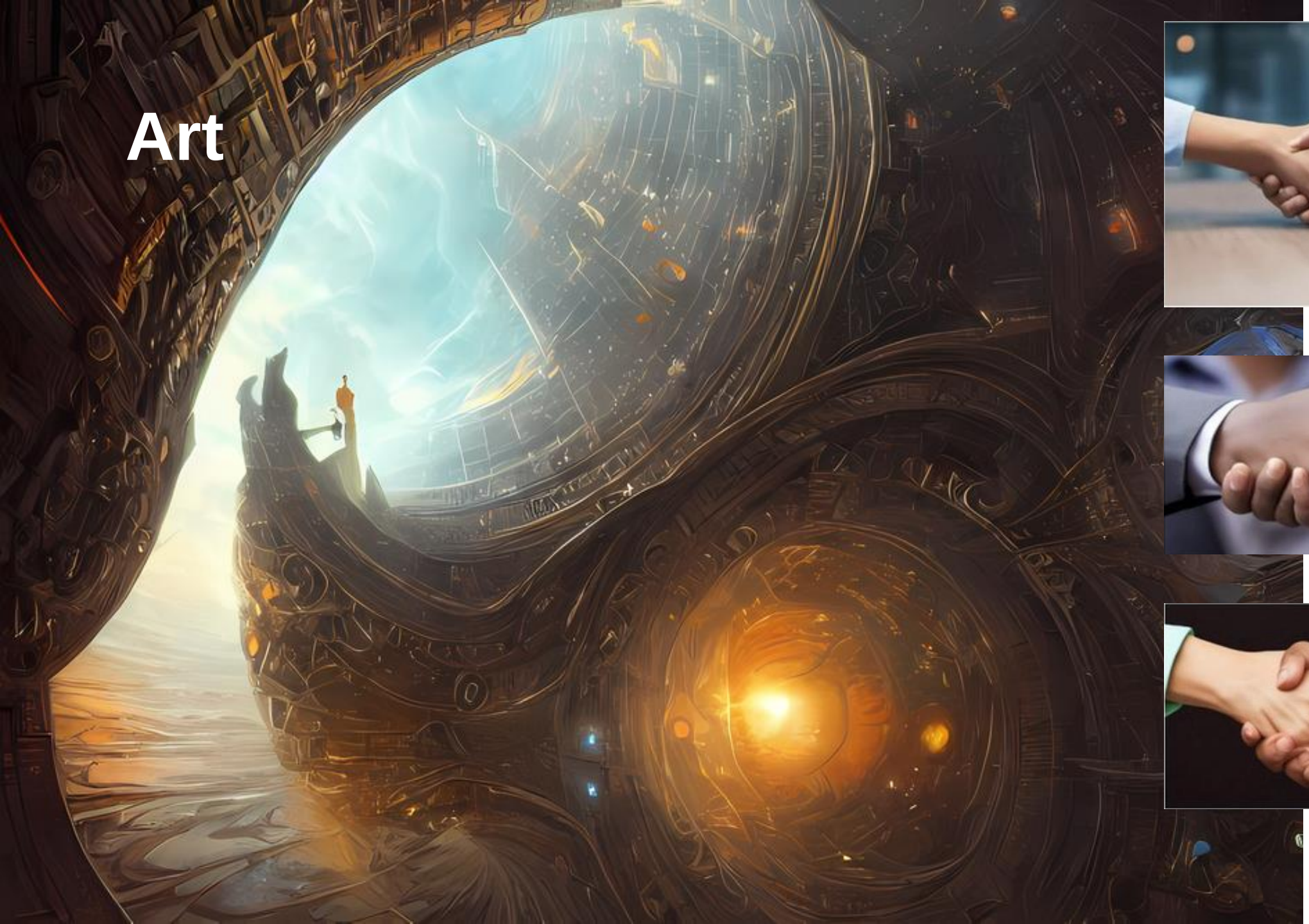
Martin Drašar

drasar@ics.muni.cz

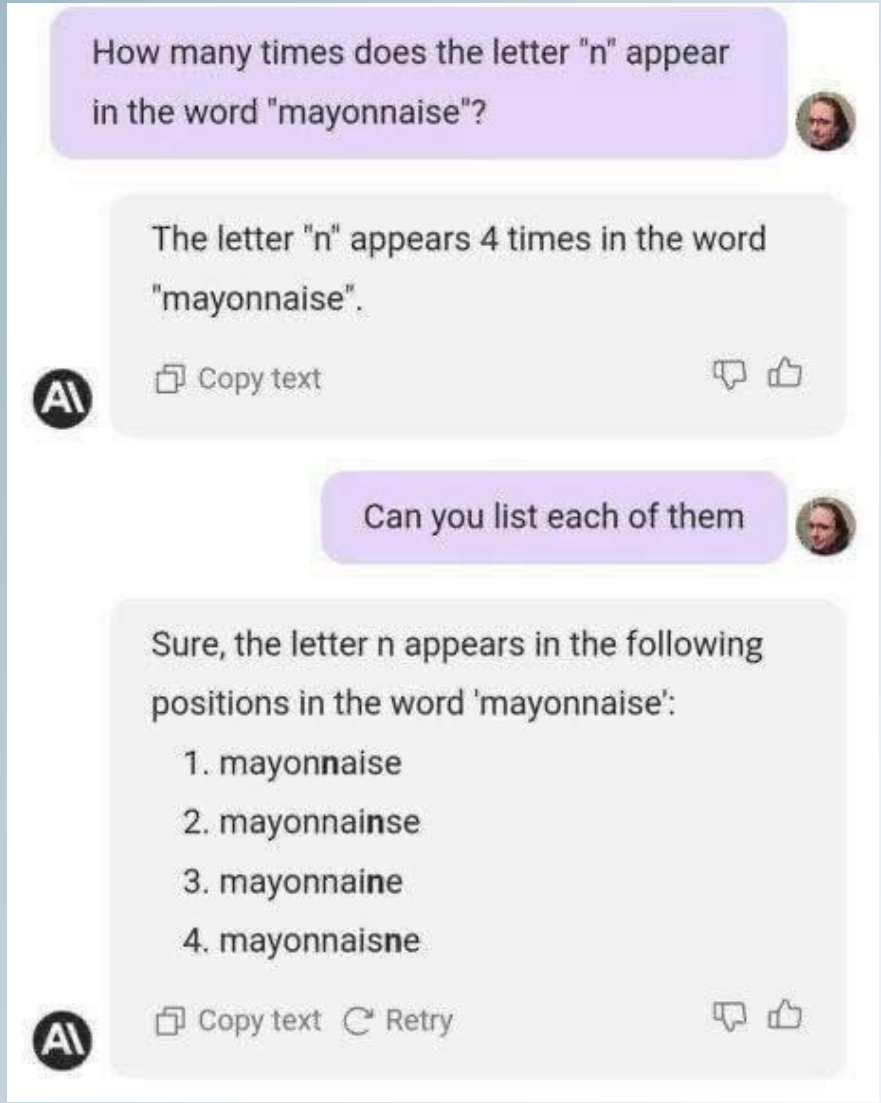
Funded by the European Union under Grant Agreement No. 101087529. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Driverless cars

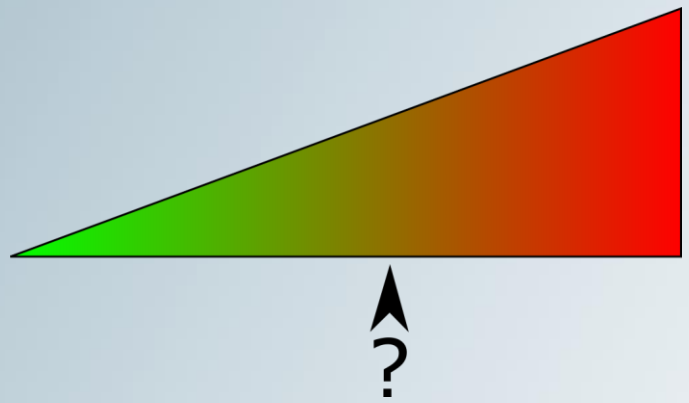
Art



Chatbots that can pass Turing test



So… how about security?



What is the problem with security?

- Tesla: 250.000.000+ km of autonomous driving
- Stable diffusion: 160.000.000 images
- GPT-3: 45 TB of text data

- Security
  - Limited datasets
  - Dataset variability not even close to real world variability
  - Training environments not up to task

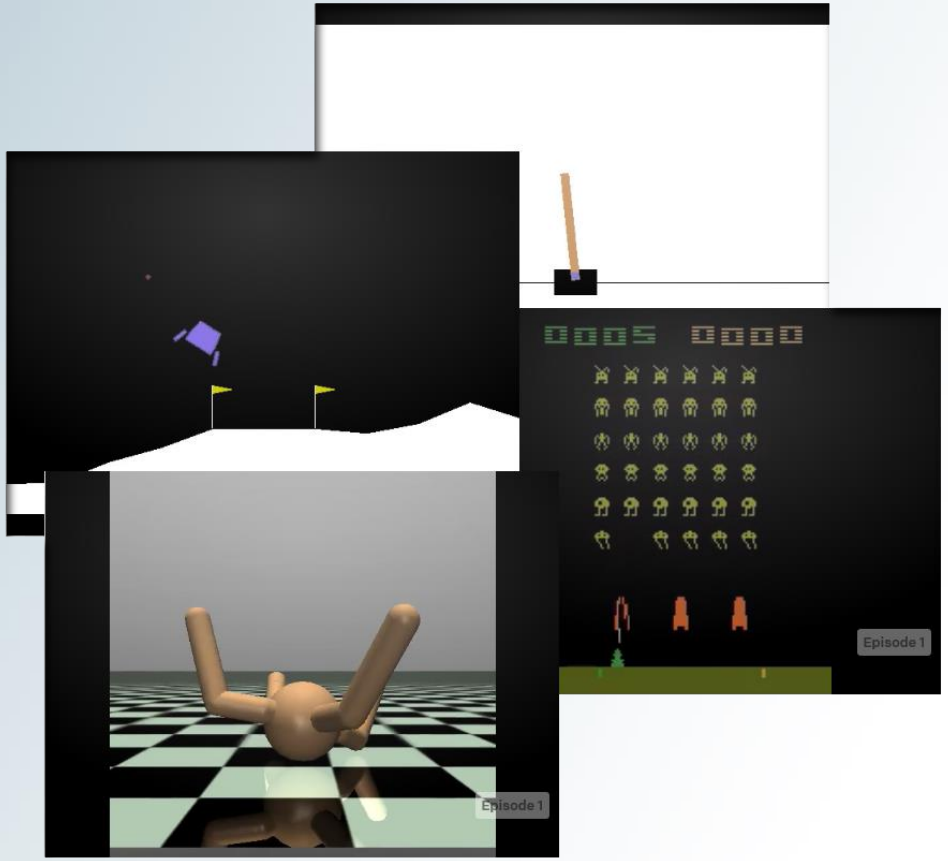

I'M GONNA MAKE MY
OWN TRAINING ENVIRONMENT
WITH BLACKJACK AND HOOKERS
imgflip.com

Which training environment paradigm to choose?

- Simulation
- Emulation
- Virtualization
- Hybrid

Quiz time!

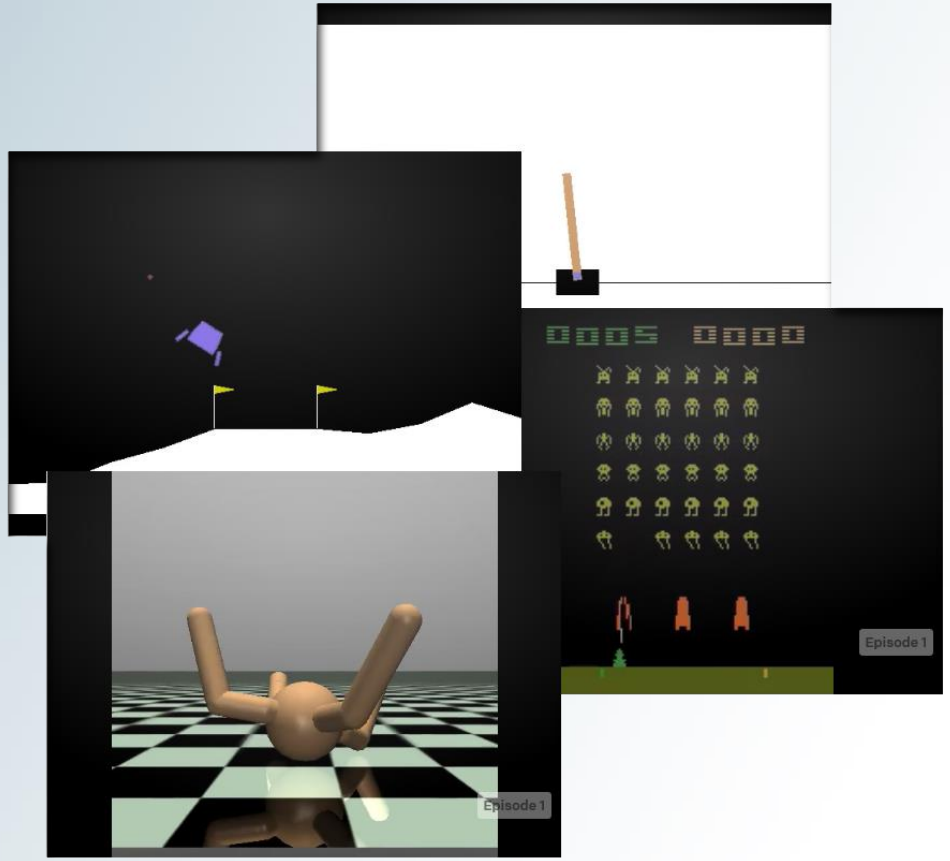
- What do you see?
- Why am I showing it?



Simulation environments

- Simple abstract problems
- Do not reflect domain complexity
- Static

- General simulation problems:
  - Data reliance
  - Specialization
  - High-level of abstraction
  - Simulation of new threats
  - Too defense-centric

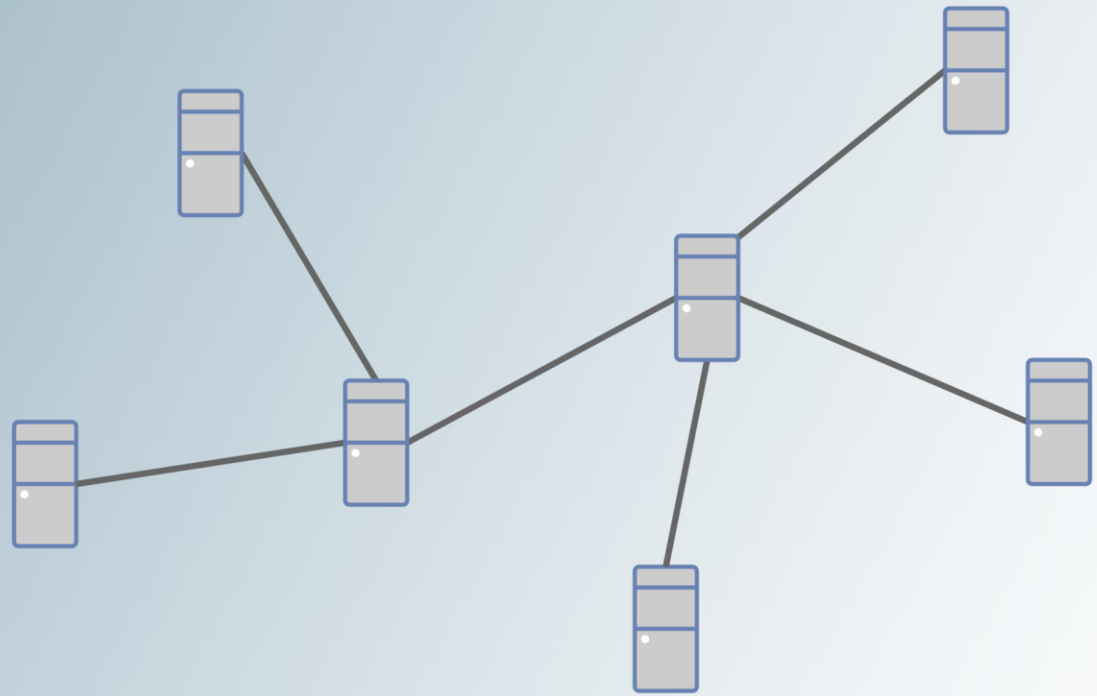



1- Agents' Reference Architecture
2- Agents' Engineering & Certification
3- Testability & At-scale Simulation
4- Implementation & Compatibility Technologies
5- Autonomous Self-engineering & Self-Assurance
6- Cyber-battle Modelling & Formal graphs
7- Agents' Individual ecision Making
8- Collective Intelligence & Decision Making
9- Agents' Learning, Loading, Sizing
10- Agents' Cooperation with External Entities
11- Agents' Stealth & Resilience
12- Friend or Foe?, Ping, Trust and Social Dynamics
13- Law, Ethics, Doctrines & Society
Societal aspects
Infrastructure, Architecture & Engineering
Stealth & Resilience
Individual & Collective Decision Making

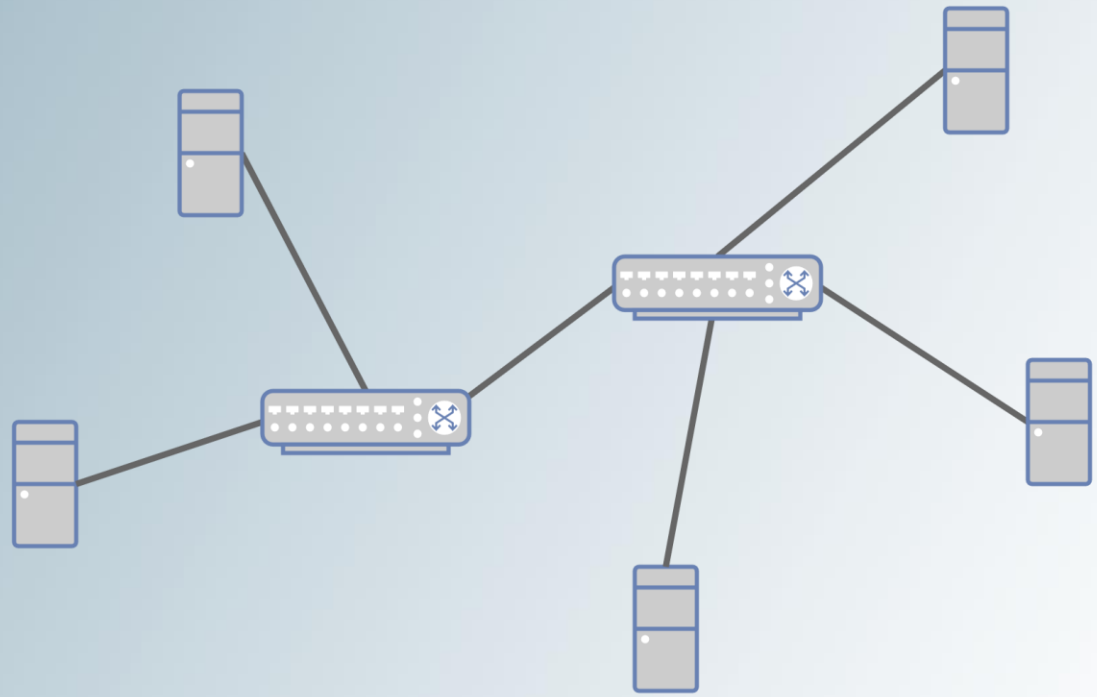
Simulation creation – deep dive

- Choose your modelling approach.
  - Discrete-event simulation, Markov processes, game theory, …
- Choose your abstraction.

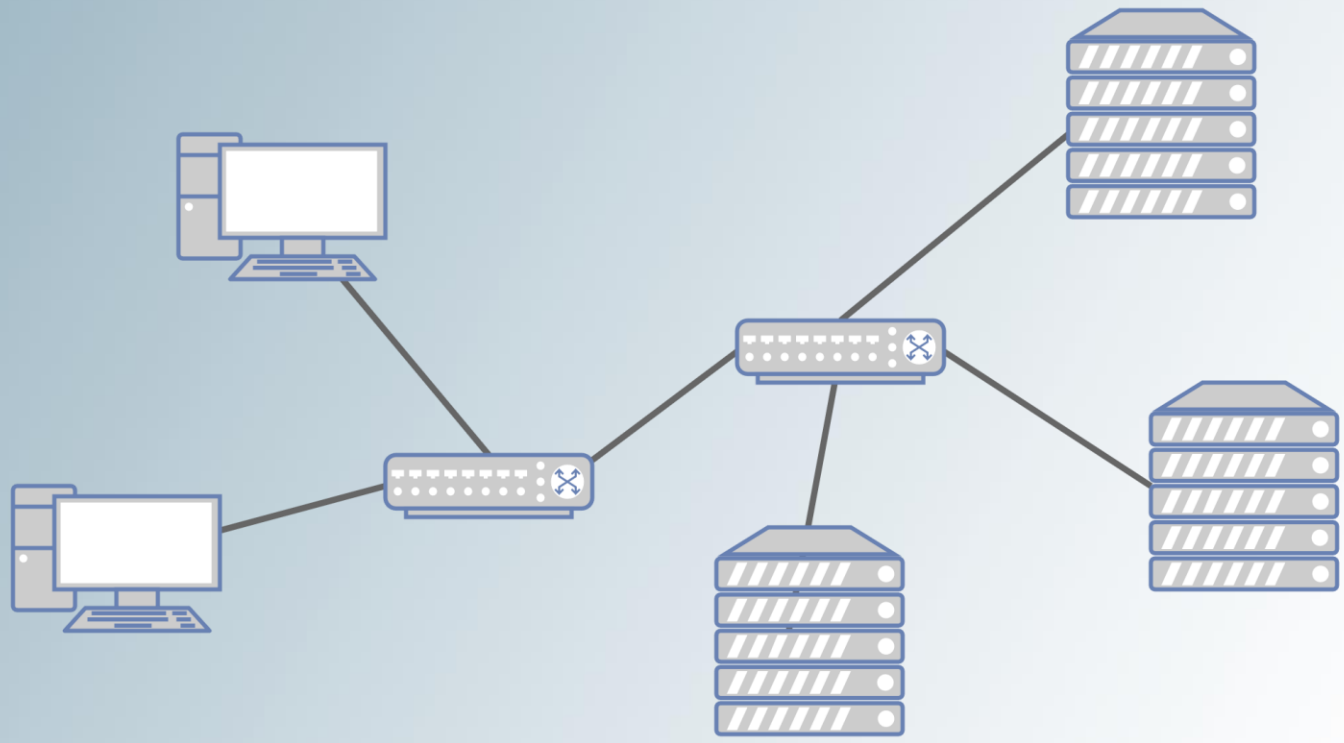
Network modelling



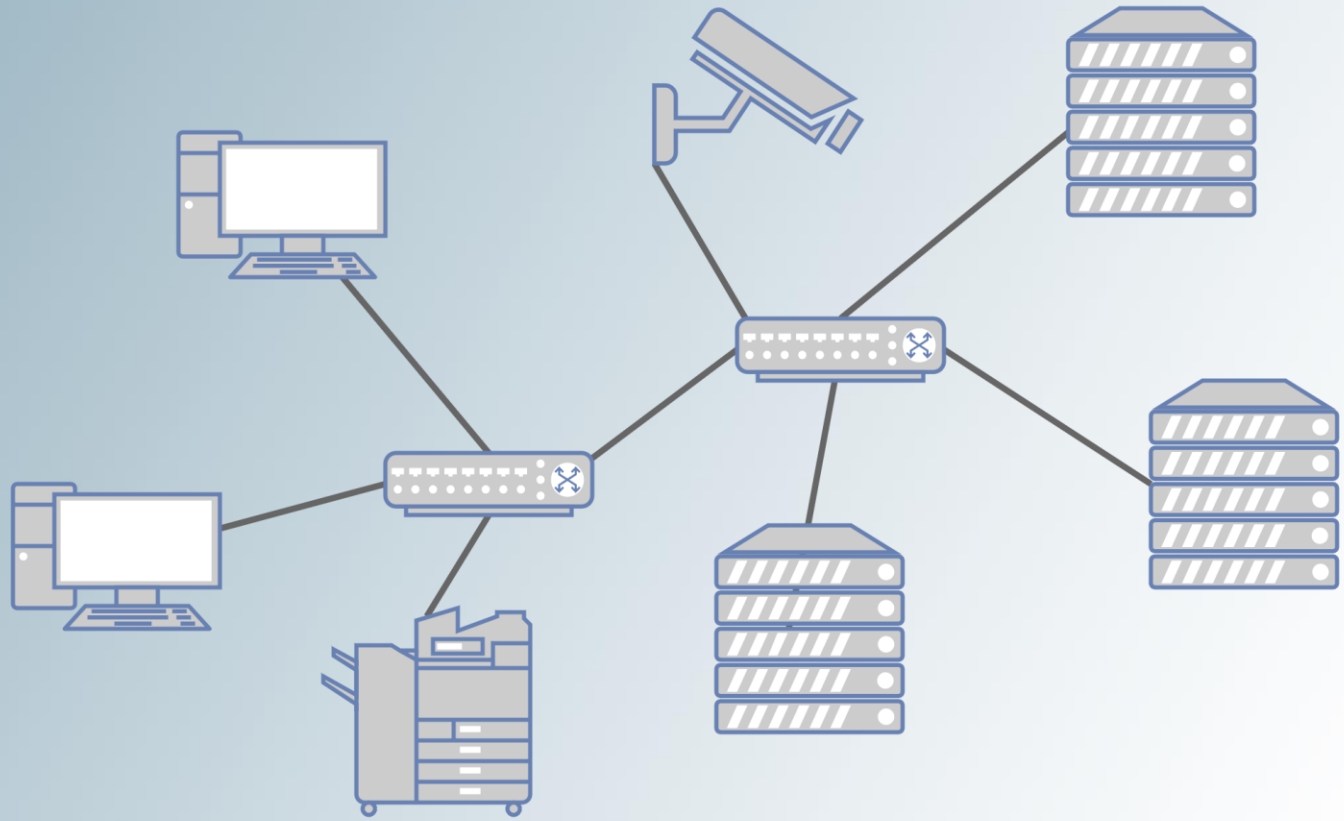
Network modelling



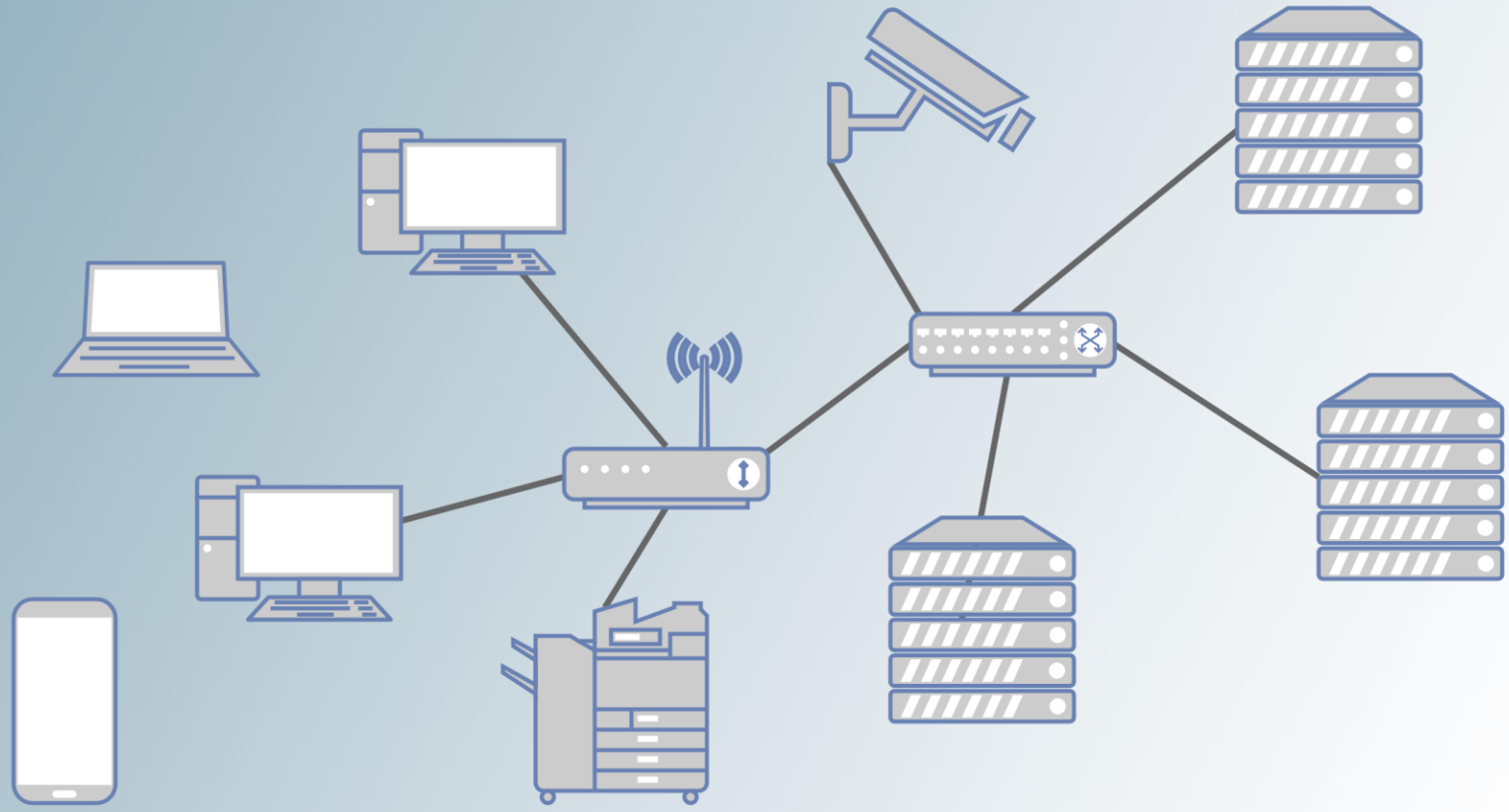
Network modelling



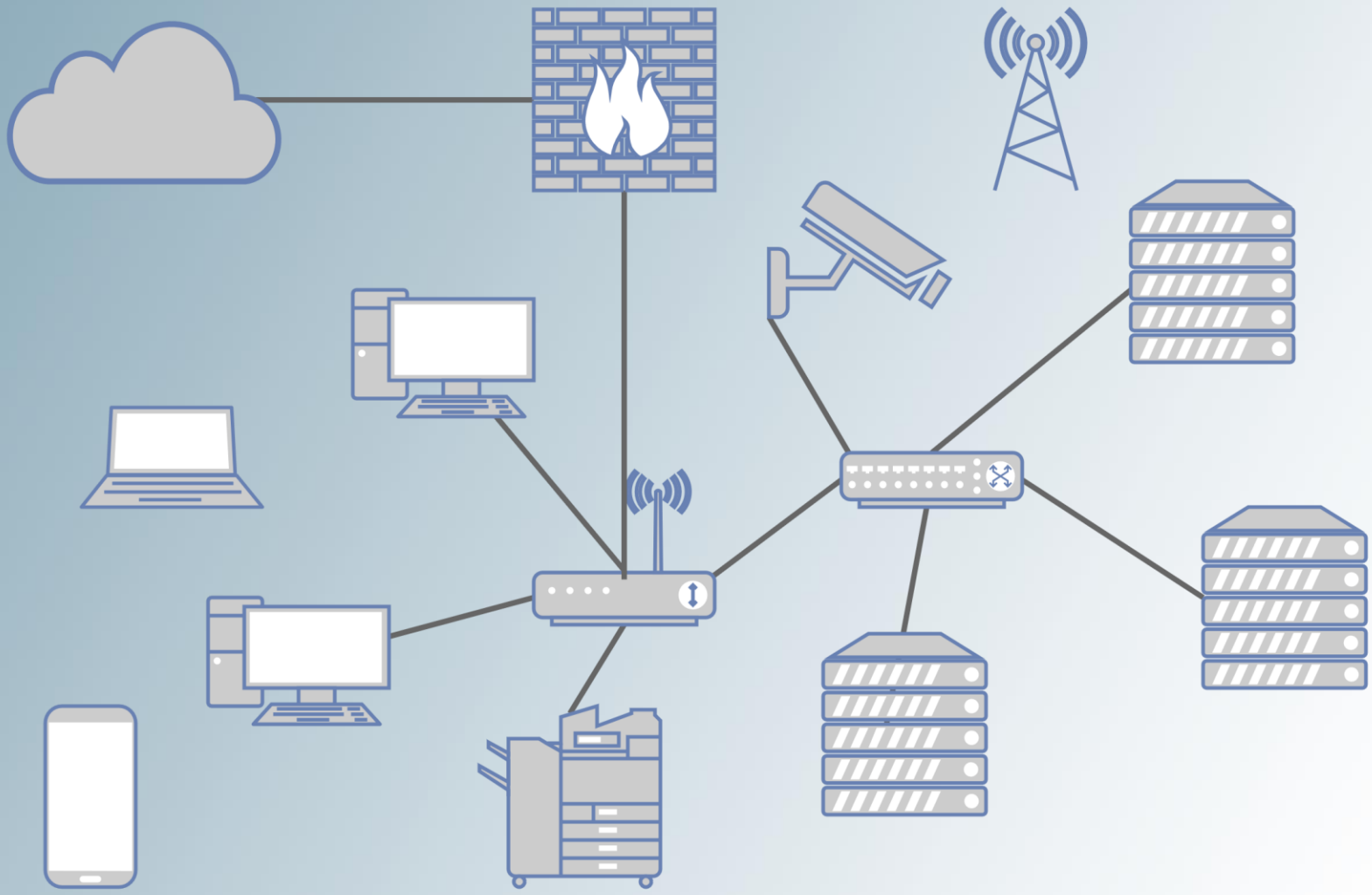
Network modelling



Network modelling



Network modelling



Host modelling

- Type
- Purpose
- Operating system
- Services
- Files
- Memory
- Buses
- Extensions

Connections modelling

- Medium (air/wire/optical/…)
- Properties (bandwidth, jitter, drops, …)
- Protocols (abstract/concrete)

User modelling

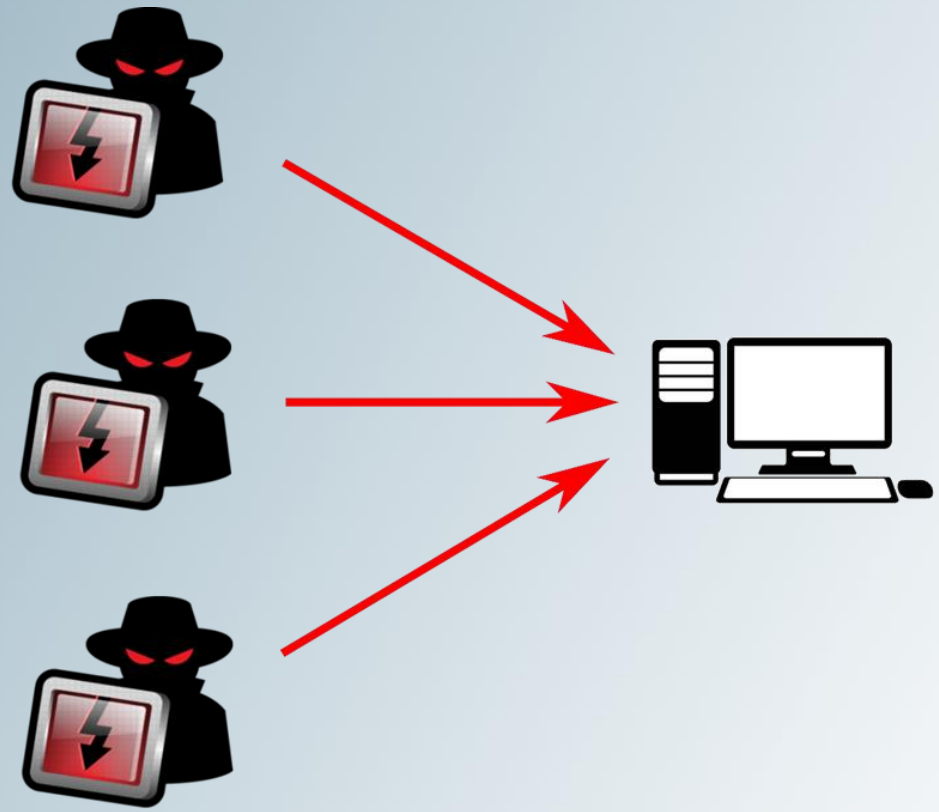
- Existence
- Identities
- Active or passive
- Traffic generation
- Activity cycles

Hundreds of considerations

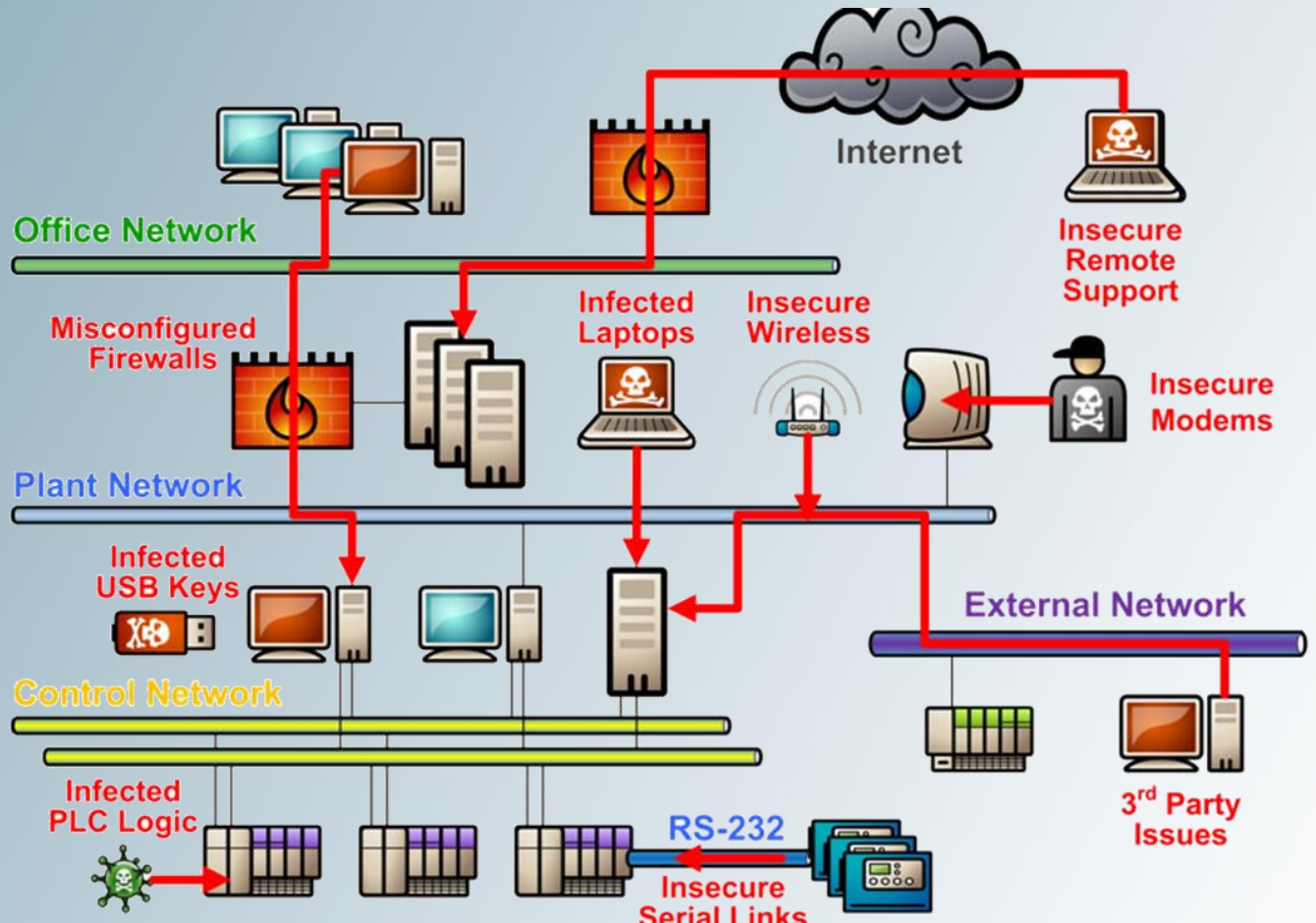
- Authentication and authorization, data handling, threats, vulnerabilities, exploits, failures, dynamics, …
- Still only the passive side
- The active side must be decided as well …

Attack modelling

Attack modelling



Attack modelling

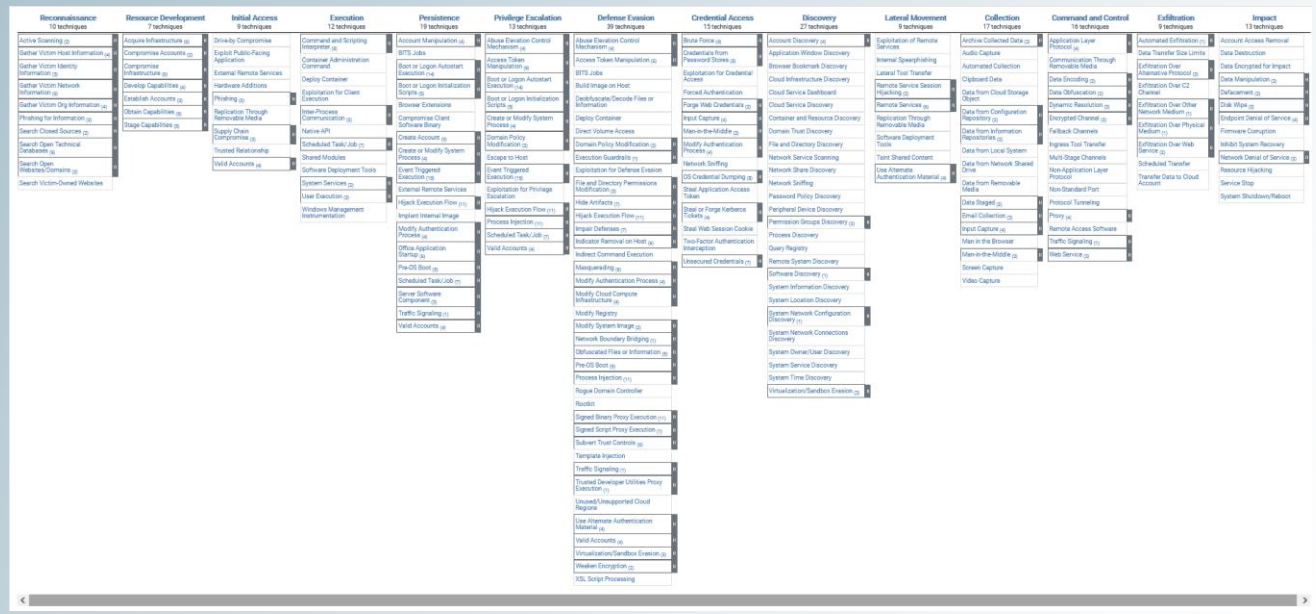


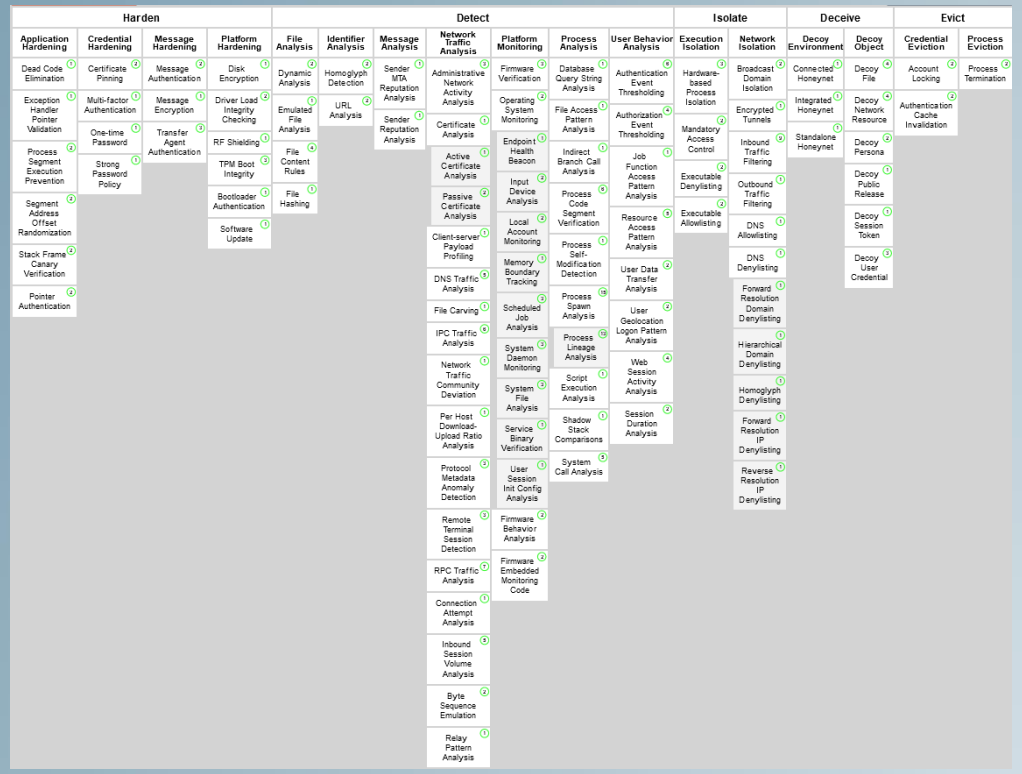
Actions and behavior

- Actions and resulting behavior depend on the simulation model
- Action must be expressible within the model

- Abstract actions?
- Attack and defense taxonomies?
- Own approach?

Attack





DEFENSE

CYST: One combination of all possible parameters

- Discrete event simulator with message passing
- Hybrid stochastic simulation
- Hosts modelled as a collection of running processes
- One type of connection with extensible properties
- Users under construction
- Complex authentication/authorization framework
- Declarative vulnerabilities mapping to CVE and others
- Exploits tied to vulnerabilities, integration of Metasploit capabilities
- Pluggable behavioral models

CYST: Functional requirements

- Simulation of multi-agent cybersecurity scenarios
- Integration with ML toolkits
- Integration and comparison of different behavioral models
- Fast prototyping of attack and defense strategies
- Integration of simulation and emulation (IDS or human in the loop)
- Deployment from simulation to emulation

CYST

- Get it, try it, break it!

- https://dl.acm.org/doi/abs/10.5555/3451906.3451908
- CODE: https://pypi.org/project/cyst-core/
- DOCS: https://muni.cz/go/cyst

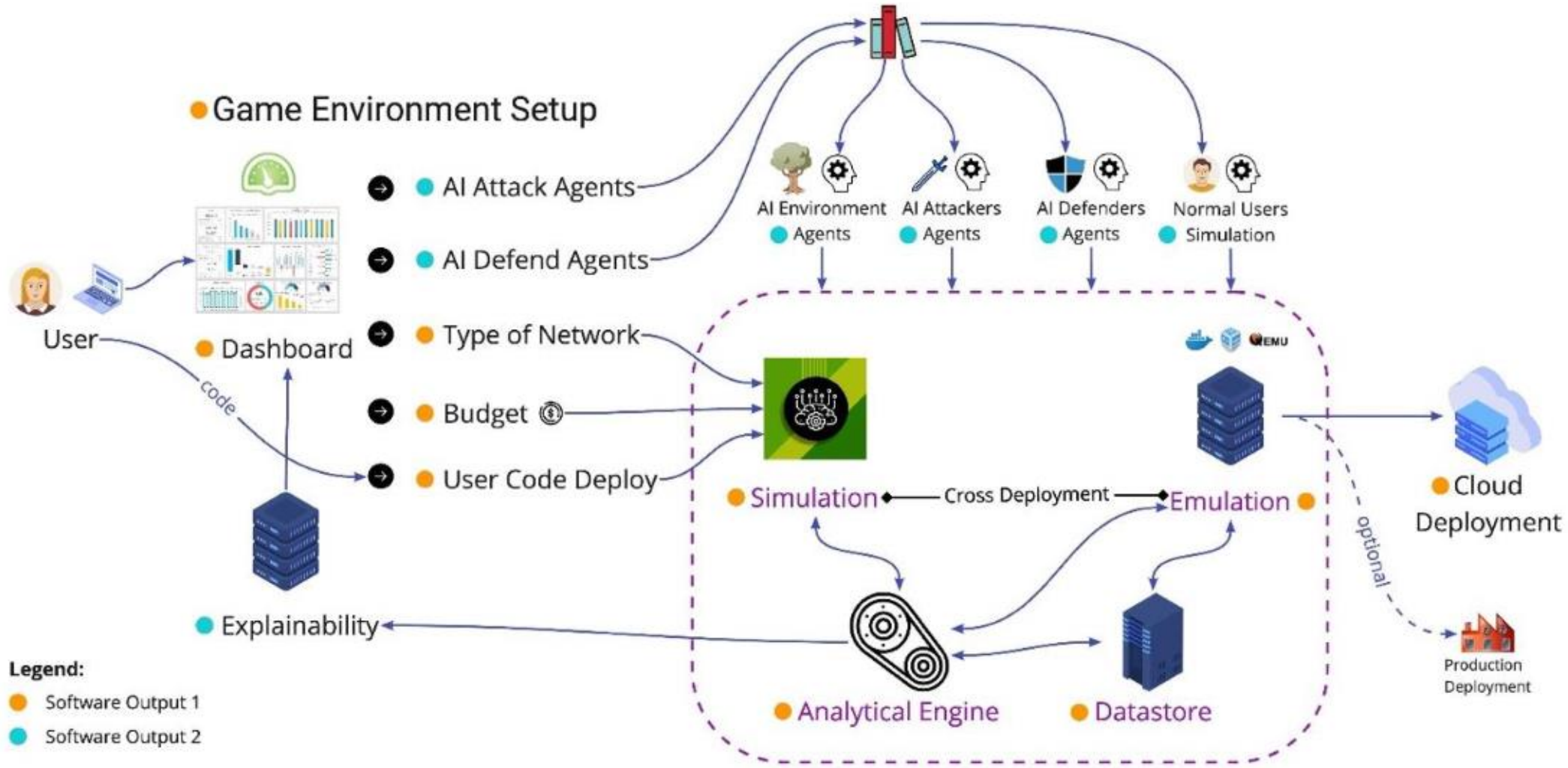
CYST – roadmap

- Parametrized generation of realistic cybersecurity scenarios
- Transformation of simulation artifacts into datasets of network traces
  - https://github.com/Trace-Share
- Multiple behavioral models
- Visualization
- Support for stealthy actions
- Support for multi-agent collaboration and communication
- Parallel training of attackers and defenders
- Transfer from simulation to emulation to the real world
  - https://beast-public.gitlab-pages.ics.muni.cz/cryton/cryton-documentation/

That’s too low level. Wake me up when it

- *can be easily deployed in cloud,*
- *can create realistic simulated and emulated scenarios,*
- *allows me to train with or against autonomous agents,*
- *provides visual analytics,*
- *has one big green button to run everything.*

AI-Dojo Software Architecture



Autonomous security

Either wait a few years

Or

Do it yourself…