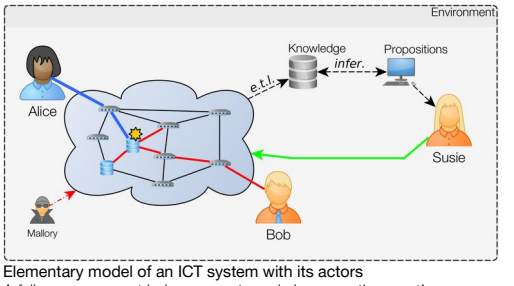
NORIA

machine learNing, Ontology and Reasoning for the Identification of Anomalies

Motivations & context

Infrastructure impairments and cyber security issues are hard to detect on large Information and Communication Technology (ICT) systems

- ✗ Events are distributed over time and locations
  cascading failures, requests for changes, stealthy attack campaigns
- ✗ High volume of heterogeneous data
  logs, alarms, measures, reports
- ✗ Partially observable states
  dropped alarms, absence of metrology, non cross-referenced data

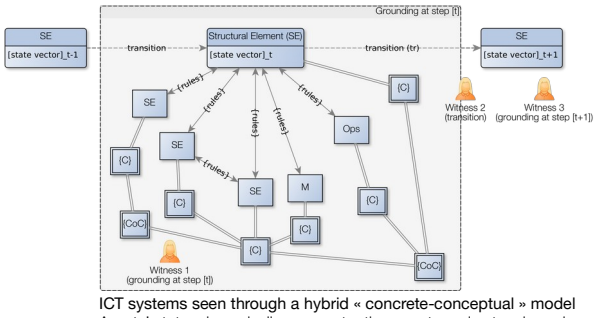


Elementary model of an ICT system with its actors
A failure on an asset induces events and alarms on the asset's neighbourhood. Susie, a network/security administrator, needs to distinguish primary events from secondary events. These events and alarms should also be contextualized w.r.t. « in policy » or « out of policy » activities.

Problem definition

How can we efficiently detect anomalies and provide explainable Root Cause Analysis (RCA)?

- ➔ Solving data heterogeneity issues w.r.t. efficient data representation and mining techniques?
- ➔ AI architectures for causal rules inference and exploitation on temporal and structural data?



ICT systems seen through a hybrid « concrete-conceptual » model
Assets' states dynamically vary w.r.t. other assets and actors based on behavioral rules. Sets of states are interpreted through higher level (composite) concepts. Predicting a next set of states/concepts is a sequential & uncertain decision problem. States and transitions are two different ways of representing the system's dynamics.

Initial exploration space

Anomaly Detection (AD)

Rule-based systems, machine learning

Knowledge Representation (KR)

Knowledge graphs, semantic graphs

CompleXity (CX)

Stream reasoning, data sketching

eXPlainability (XP)

Subsumption, graph neural networks, neuro-symbolic computing

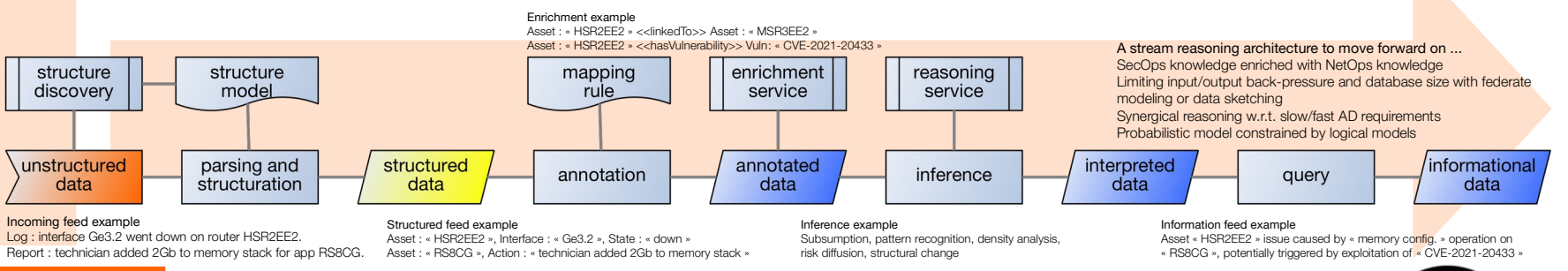
State of the Art (SotA) observables	Exploration space impact			
	KR	CX	AD	XP
Classical control-loop model do not scale.		X		
Infrastructure technical-behavioral-operational characteristics embed implicit or explicit logical systems.	X			X
NetOps and SecOps share operational and functional characteristics.		X	X	
Available AD techniques are mostly « narrow AI ».		X	X	
Graph representation (un)directly applicable.	X		X	
« good » KR means common sense and efficient data handling.	X	X		
Discourse domain can be covered by a combination of existing taxonomies, thesaurus and ontologies.	X	X		

Observables influence research axis in an intricate way
The 4 research axis will take advantage of potential findings on these observables.

Evaluation

Functional capabilities for NetOps/SecOps analyst

Learn understandable model of system behavior
Map time-location events to potential deleterious system states
Contextualize (sequence of) events w.r.t. operational process
Reduce root cause search space in near-real-time over events



orange™

EURECOM
Sophia Antipolis

Lionel TAILHARDAT (Orange INNOV)
Yoan CHABOT (Orange INNOV)
Raphaël TRONCY (EURECOM)

Sept. 2021