Security Forum

How to Measure Anything in Cybersecurity Risk II: Measure Harder

Ben Goldsworthy, rapidly losing that ‘new consultant’ smell

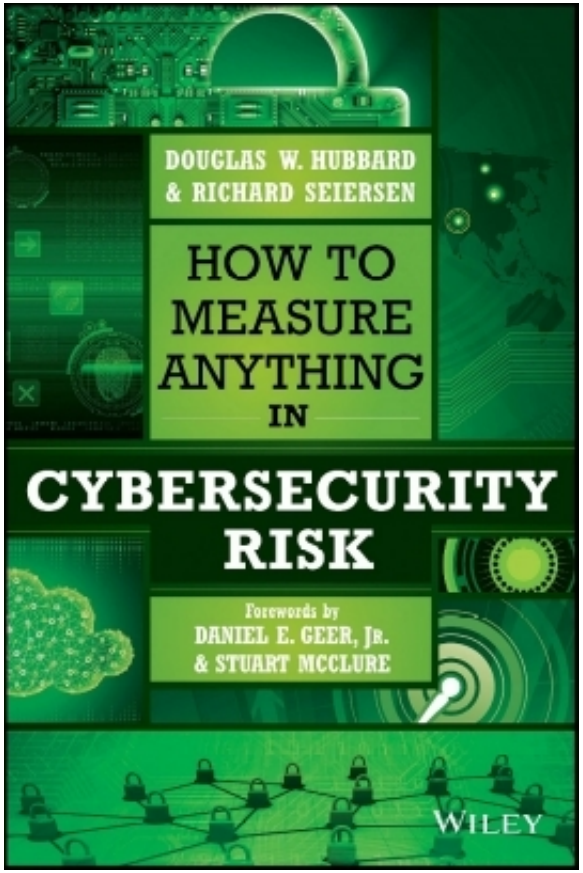
Delivered 2021-11-17

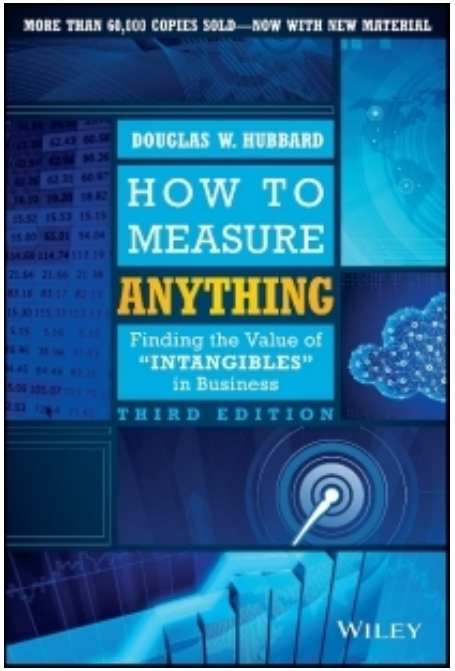
Contents

1. Previously on “How to Measure Anything in Cybersecurity”
2. Decomposition
3. Measuring control efficacy
4. Tooling

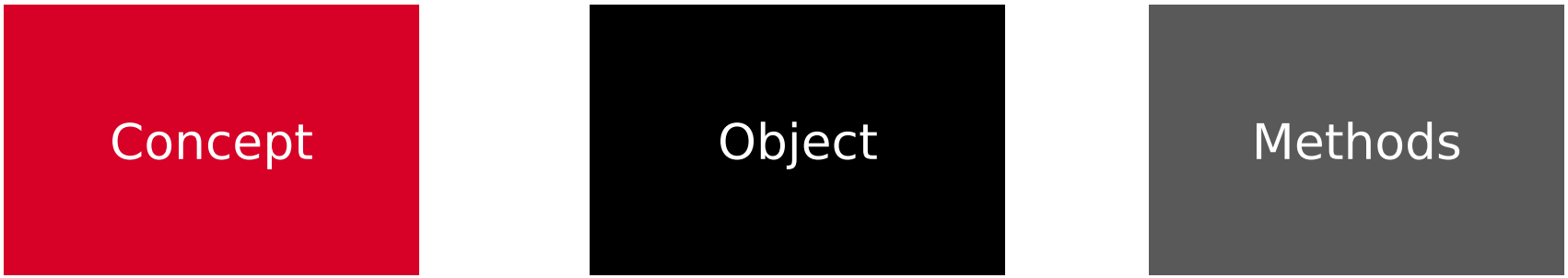
A. Calibration exercise

Previously…
How to Measure Anything…





Previously…
What is ‘measurement’?



Previously…
What is ‘measurement’?

Concept

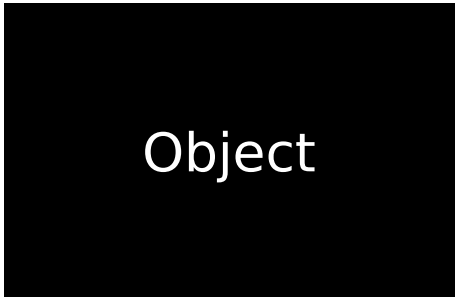
Measurement: A quantitatively-expressed reduction of uncertainty based on one or more observations

Previously…
Bayesian Measurement

Concept

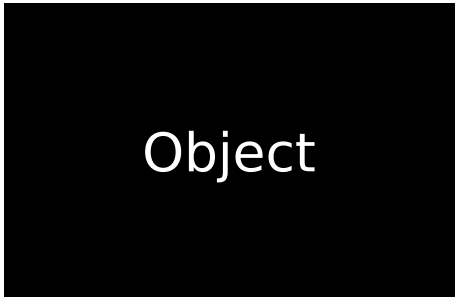
Probability: the state of uncertainty of an observer (a.k.a. ‘degree of belief’)

Previously…
The Object of Measurement



1. If it matters at all, it is detectable/observable
2. If it is detectable, it can be detected as an amount (or range of possible amounts)
3. If it can be detected as a range of possible amounts, it can be measured

Previously…
Risk



- **Risk:** A state of uncertainty where some of the possibilities involve a loss, catastrophe or other undesirable outcome
- **Measurement of Risk:** A set of possibilities, each with quantified possibilities and quantified losses

Previously…
Statistics

Methods

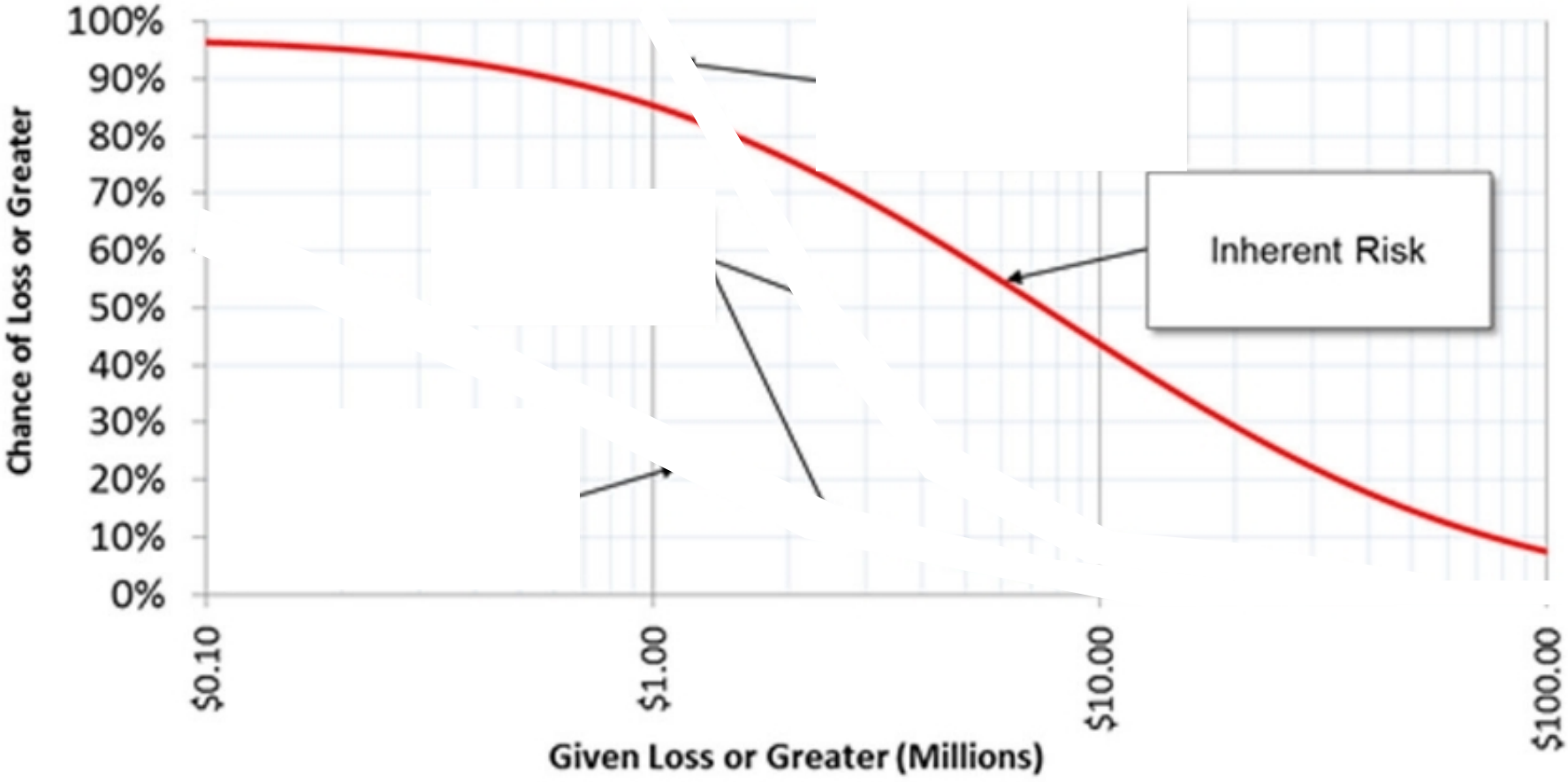
- ‘Cybersecurity is **not** some exceptional area **outside the domain of statistics** but rather **exactly the kind of problem** statistics was made for.’

Previously…
One-for-one Substitution

Instead of:	We substitute:
• Rating **likelihood** on a scale of 1 to 5	• Estimating the **probability of the event happening** in a given period of time
• Rating **impact** on a scale of 1 to 5	• Estimating a 90% confidence interval for a **monetized loss**
• Plotting likelihood and impact scores on a **risk matrix**	• Using the quantitative likelihood and impact to generate a **loss exceedance curve**
• Further dividing the risk matrix into **risk categories** and guessing what you should do	• Comparing the loss exceedance curve to a **risk tolerance curve** and prioritising actions based on return on mitigation

Previously...
Monte Carlo Simulations

Previously...
Loss Exceedance Curve



Previously…
Real Example: *CSBS 2020*

Questions from the Q&A

1. ‘How can you quantify losses in non-monetary terms?’

Decomposition

Event	Event Probability (per Year)	Impact (90% Confidence Interval)		Random Result (zero when the event did not occur)
		Lower Bound	Upper Bound	
AA	.1	$50,000	$500,000	0
AB	.05	$100,000	$10,000,000	$8,456,193
AC	.01	$200,000	$25,000,000	0
AD	.03	$100,000	$15,000,000	0
AE	.05	$250,000	$30,000,000	0
AF	.1	$200,000	$2,000,000	0
AG	.07	$1,000,000	$10,000,000	$2,110,284
AH	.02	$100,000	$15,000,000	0
⇩	⇩	⇩	⇩	⇩
ZM	.05	$250,000	$30,000,000	0
ZN	.01	$1,500,000	$40,000,000	0
			Total:	$23,345,193

Decomposition

Event	Probability of event occurring in a year	Type of Event if One Occurs			Confidentiality and Integrity 90% Confidence Interval ($000)		Availability 90% Confidence Intervals			
							Duration of Outage (hours)		Cost per Hour ($000)	
		Only Conflnt	Only Availability	Both Types	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound
AA	.1	.2	.7	.1	$50	$50	.25	4	$2	$10
AB	.05	.3	.5	.2	$100	$10,000	.25	8	$1	$10
AC	.01	.1	.8	.1	$200	$25,000	.25	12	$40	$200
AD	.03	0	0	1.0	$100	$15,000	.25	2	$2	$10
AE	.05	0	.6	.4	$250	$30,000	1	24	$5	$50

Hubbard & Seiersen (2016). *How to Measure Anything in Cybersecurity Risk*.

Decomposition 101

- **Clear**
  - Does everybody know what you mean? Do *you* know what you mean?
- **Measurable**
  - What do you see when you see more of it?
- **Useful**
  - What would you do differently if you knew this?

Example: ‘Threat Actor Skill Level’

- **Clear**
  - Can you define what you mean by ‘skill level’? Is this really an unambiguous unit of measure or even a clearly defined discrete state?
- **Measurable**
  - How would you even detect this? What basis do you have to say that skill levels of some threats are higher or lower than others?
- **Useful**
  - Even if you had unambiguous definitions for this, and even if you could observe it in some way, how would the information have bearing on some action of your firm?

Avoid Over-decomposition

- Imagine someone standing in front of you **holding a crate**, about 0.5 m x 0.25 m x 0.25 m.
- They ask you to provide a **90% CI** on the weight of the crate, solely by looking at it.
- They’re not jacked, so it probably doesn’t weigh **over 150 kg**
- You give a 90% CI that it weighs **between 10–150 kg**
- This range is large, but can you try to narrow it by estimating the number of items in the crate and the weight per item?
- Would your estimate **be better** as a result?
- **It would probably be worse.**

Uninformative Decompositions

- ‘What you have done is decomposed the problem into multiple **purely speculative estimates** that you then use to try and do some math.’

- ‘Decompositions should be **less abstract** to the expert than the aggregated amount. If you find yourself decomposing a dollar impact into factors like threat skill level then you should have **less uncertainty** about the new factors than you did about the original, direct estimate of monetary loss.’

Rules of Decomposition

1. Decompositions should leverage what you are **better at estimating** or data you can obtain

2. Try to **check your decompositions** against a directly estimated range with a simulation.
   - You might decide to toss the decomposition if it produces results you think are absurd, or you might decide your original range is the one that needs updating.

For Want of a Nail...

For want of a nail the shoe was lost.
For want of a shoe the horse was lost.
For want of a horse the rider was lost.
For want of a rider the message was lost.
For want of a message the battle was lost.
For want of a battle the kingdom was lost.
And all for the want of a horseshoe nail.

...as Applied to Military Risk

- Top-level/existential risk: **loss of kingdom**
  - Decompose into **risk of battle loss**
    - Decompose into **risk of message loss** (i.e., comms disruption)
      - Decompose into **risk of messenger loss** (i.e., asset loss)
        - Decompose into **risk of rider loss** (i.e., chance of injury or fatality)
          - etc.?

Estimating Risk of ‘Rider Loss’

Remember

‘You almost always have more data than you think.’

JSP 375, Vol. 1, Ch. 16

Minor injury accident or incident	Any injury, accident or incident that results in: 1) Up to seven days lost time and is not reportable under RIDDOR or causes minor damage. 2) Mild heat illness (heat illness not requiring admission to hospital).

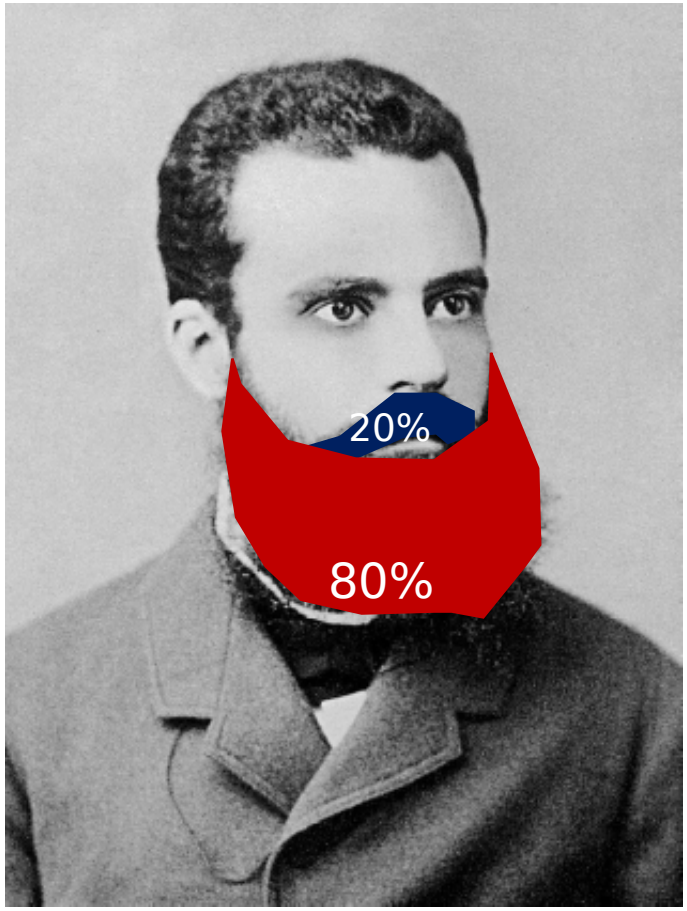
JSP 375, Vol. 1, Ch. 16

Specified Injuries	Any injury, accident or incident that results in: 1) A fracture, other than to fingers, thumbs and toes. 2) Amputation of an arm, hand, finger, thumb, leg, foot or toe. 3) Permanent loss of sight or reduction of sight. 4) Crush injuries leading to internal organ damage. 5) Serious burns (covering more than 10% of the body, or damaging the eyes, respiratory system or other vital organs). 6) Scalpings (separation of the skin from the head) which requires hospital treatment. 7) Unconsciousness caused by head injury or asphyxia. 8) Severe heat illness (heat illness requiring admission to intensive care). 9) Any other injury which leads to hypothermia, heat-induced illness or requires resuscitation or admittance to hospital for more than 24 hours.

Determining Likelihoods

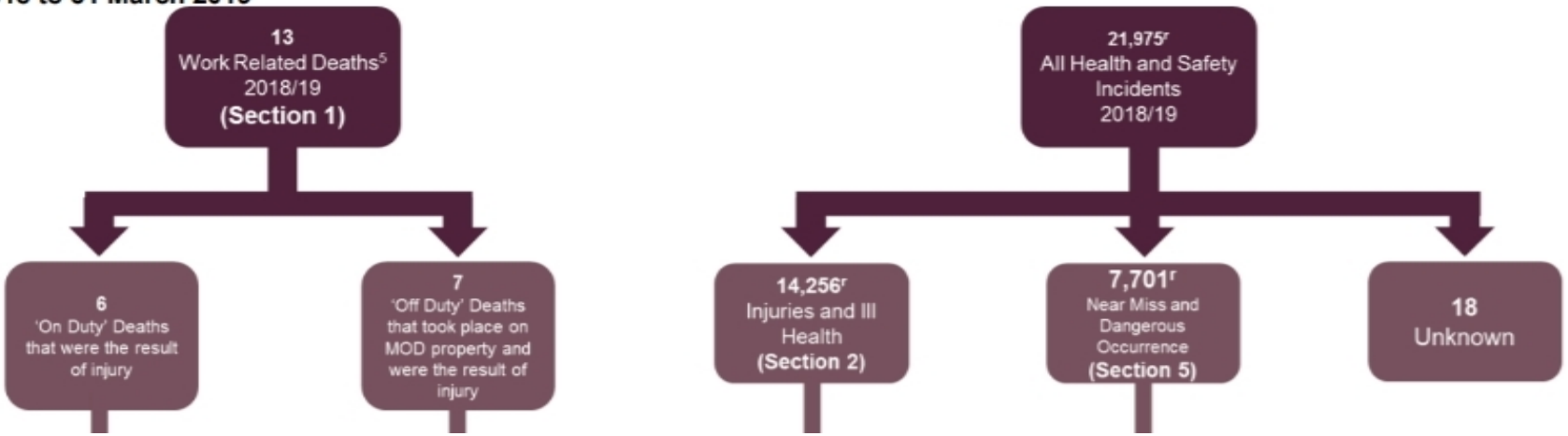
- Minor injuries, incidents or diseases
  - 80%
-
- Major injuries, incidents or diseases
  - 10%
-
- Specified injuries
  - 9.5%
-
- Death
  - 0.5%

20%



Or…

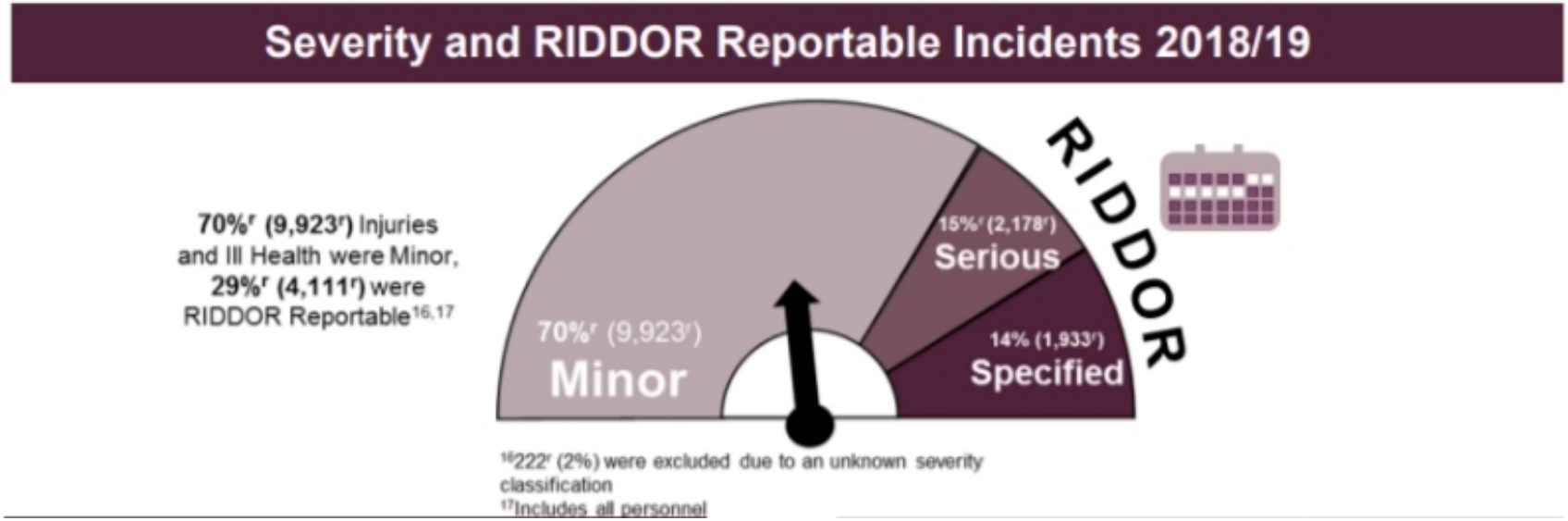
Figure 1: Health and Safety Incidents, Numbers
1 April 2018 to 31 March 2019



MoD, *MOD Health and Safety Statistics: Annual Summary & Trends Over Time 2014/15–2018/19* (2019, rev. 2020)

Or…

- Deaths = 13 deaths out of 14,256 incidents, or 0.1%



MoD, *MOD Health and Safety Statistics: Annual Summary & Trends Over Time 2014/15–2018/19* (2019, rev. 2020)

Comparing

- Minor injuries, incidents or diseases
    - 80% **(actually 70%)**
- Major injuries, incidents or diseases
    - 10% **(actually 15%)**
- Specified injuries
    - 9.5% **(actually 14%)**
- Death
    - 0.5% **(actually 0.1%)**

Now What?

- Now that you have a measurable value (№ of injuries/deaths of various severities), **can you turn it into a monetary amount**?

- Probably!
  - e.g., see Knieser & Viscusi (2019) (“The Value of a Statistical Life”), which calculated the US Government’s valuing of an individual human life as being equal to **$10m**
  - Quality-adjusted life years (QALYs)
  - Work hours lost

Shop Around

Current Mishap Definitions and Reporting Criteria

Mishap Class	Total Property Damage	Fatality/Injury
A	$2,500,000 or more and/or aircraft destroyed	Fatality or permanent total disability
B	$600,000 or more but less than $2,500,000	Permanent partial disability or three or more persons hospitalized as inpatients
C	$60,000 or more but less than $600,000	Nonfatal injury resulting in loss of time from work beyond day/shift when injury occurred
D	$25,000 or more but less than $60,000	Recordable injury or illness* not otherwise classified as a Class A, B, or C

DoD, “Current Mishap Definitions and Reporting Criteria” (https://navalsafetycenter.navy.mil/Resources/Current-Mishap-Definitions/, accessed 2021-11-15)

Decomposing Your Simulation

Event	Event Prob.	Prob. of Minor Injuries	Minor Injuries Bounds	Prob. of Major Injuries	Major Injuries Bounds	Prob. of Deaths	Deaths Bounds
• AA	• 0.3	• 0.2	• 10–20	• 0.2	• 2–3	• 0.05	• 5–10
• AB	• 0.2	• 0.2	• 44–200	• 0.4	• 20–30	• 0.02	• 100–400
• AC	• 0.5	• 0.6	• 1–3	• 0.5	• 1–2	• 0.04	• 1–40

Stop Right There

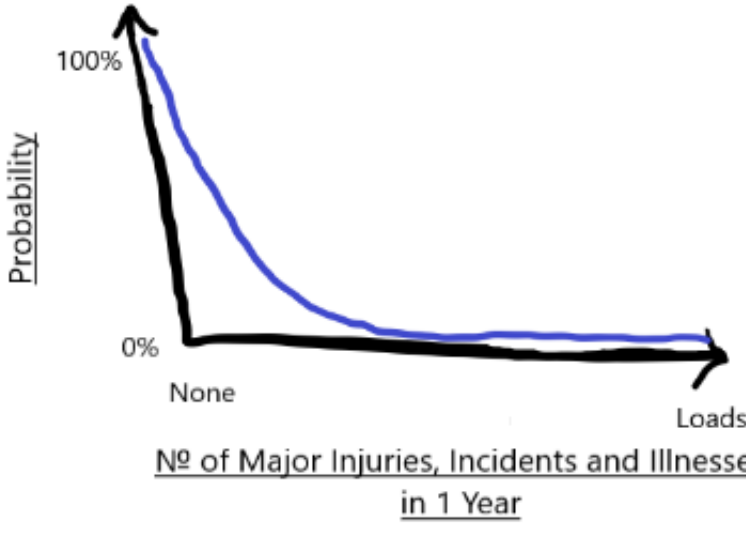
- Depending on your goals, **do you actually need a monetary amount**?

- Is this extra level of abstraction actually providing you with more clarity? **Reducing your uncertainty** further?

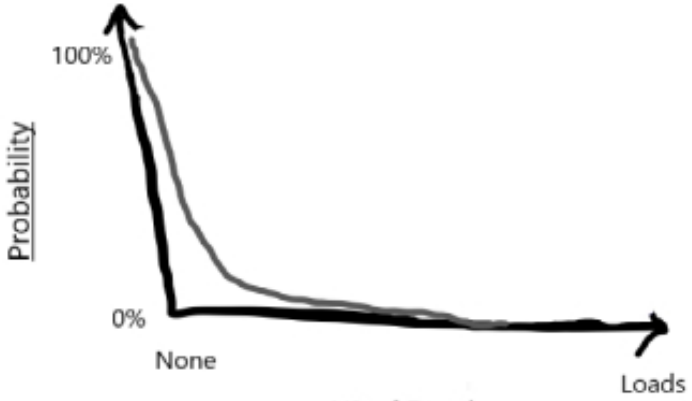
- **A number is a number**, whether it has a pound sign before it or not

Generated LECs

№ of Minor Injuries, Incidents and Illnesses in 1 Year

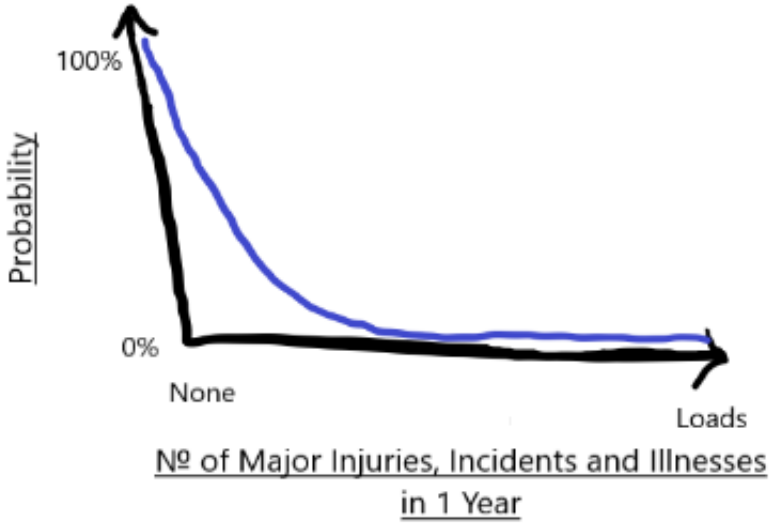


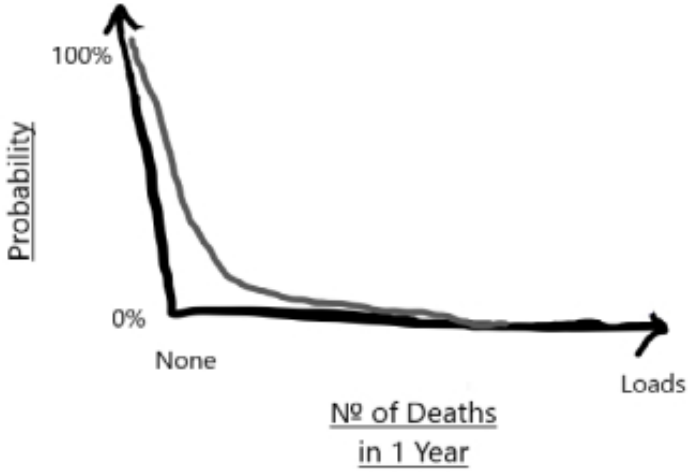
№ of Major Injuries, Incidents and Illnesses in 1 Year



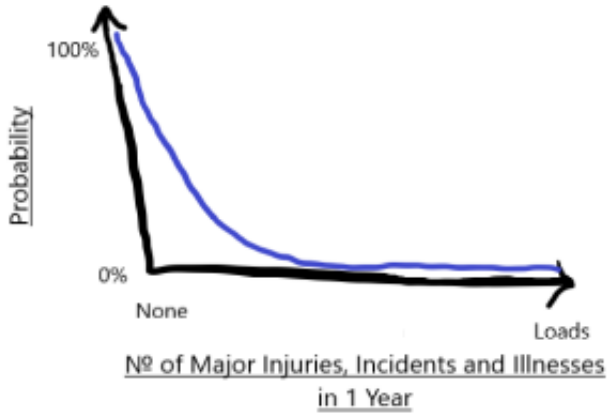
№ of Deaths in 1 Year

Moving Up in the World

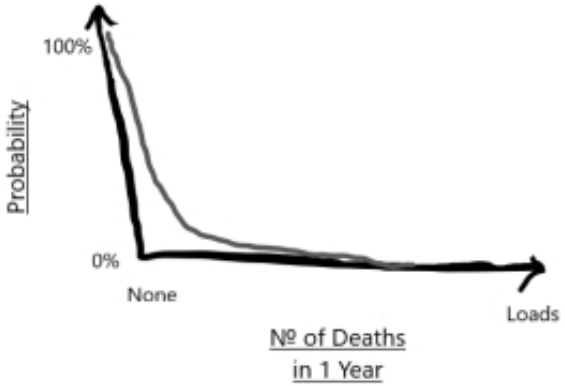


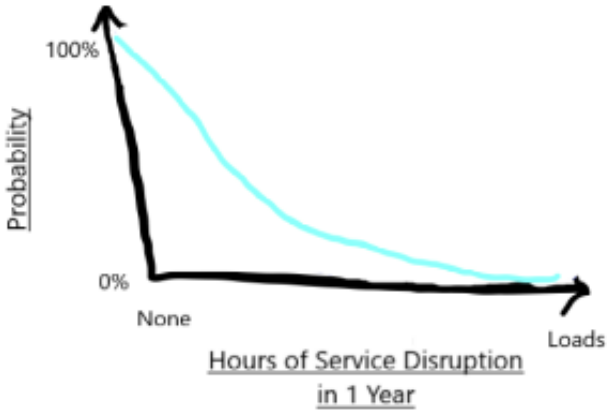


Pairings


100%
Probability
0%
None
Loads
№ of Major Injuries, Incidents and Illnesses
in 1 Year


100%
Probability
0%
None
Loads
£s of Property Damage
in 1 Year


100%
Probability
0%
None
Loads
№ of Deaths
in 1 Year


100%
Probability
0%
None
Loads
Hours of Service Disruption
in 1 Year

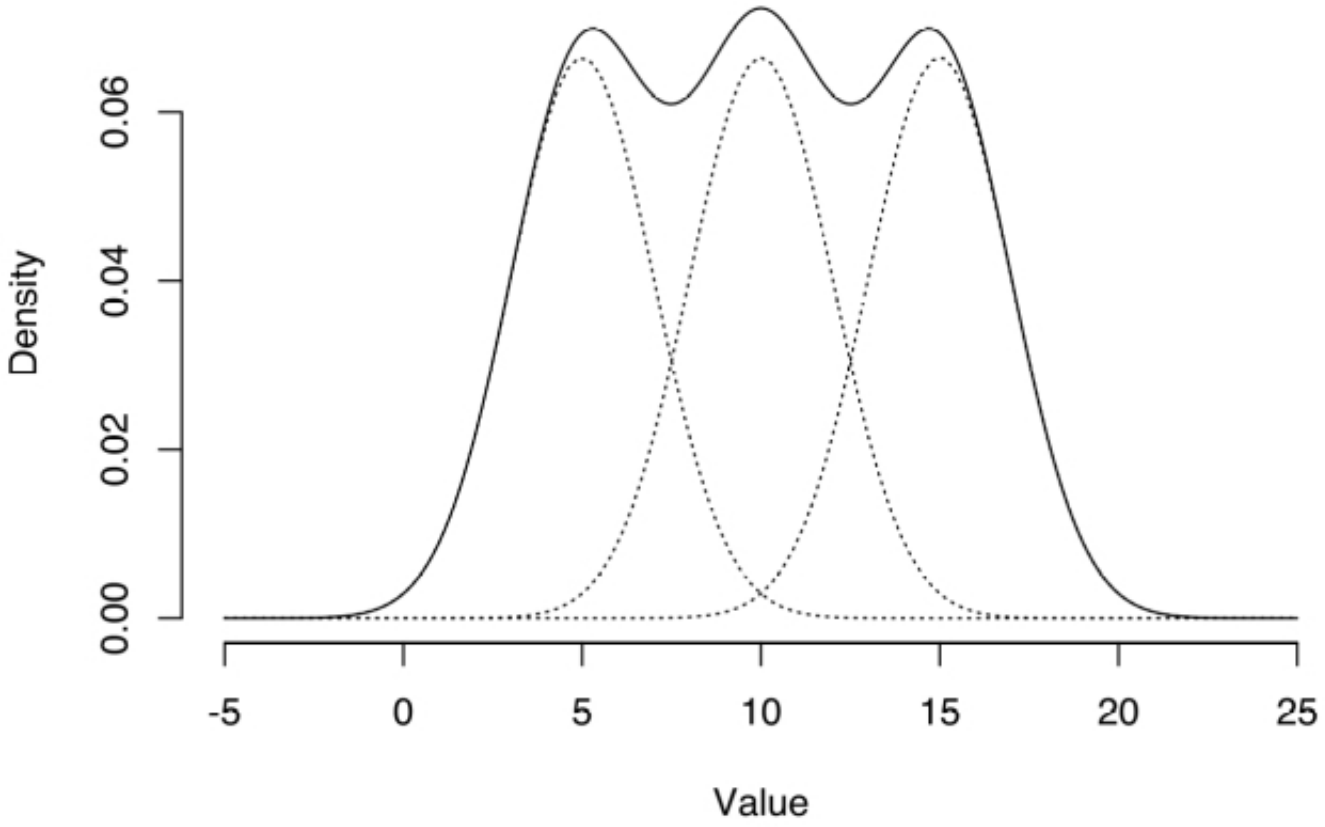
Combining Random Variables

We can combine means directly, but we can't do this with standard deviations. We can combine variances as long as it's reasonable to assume that the variables are independent.

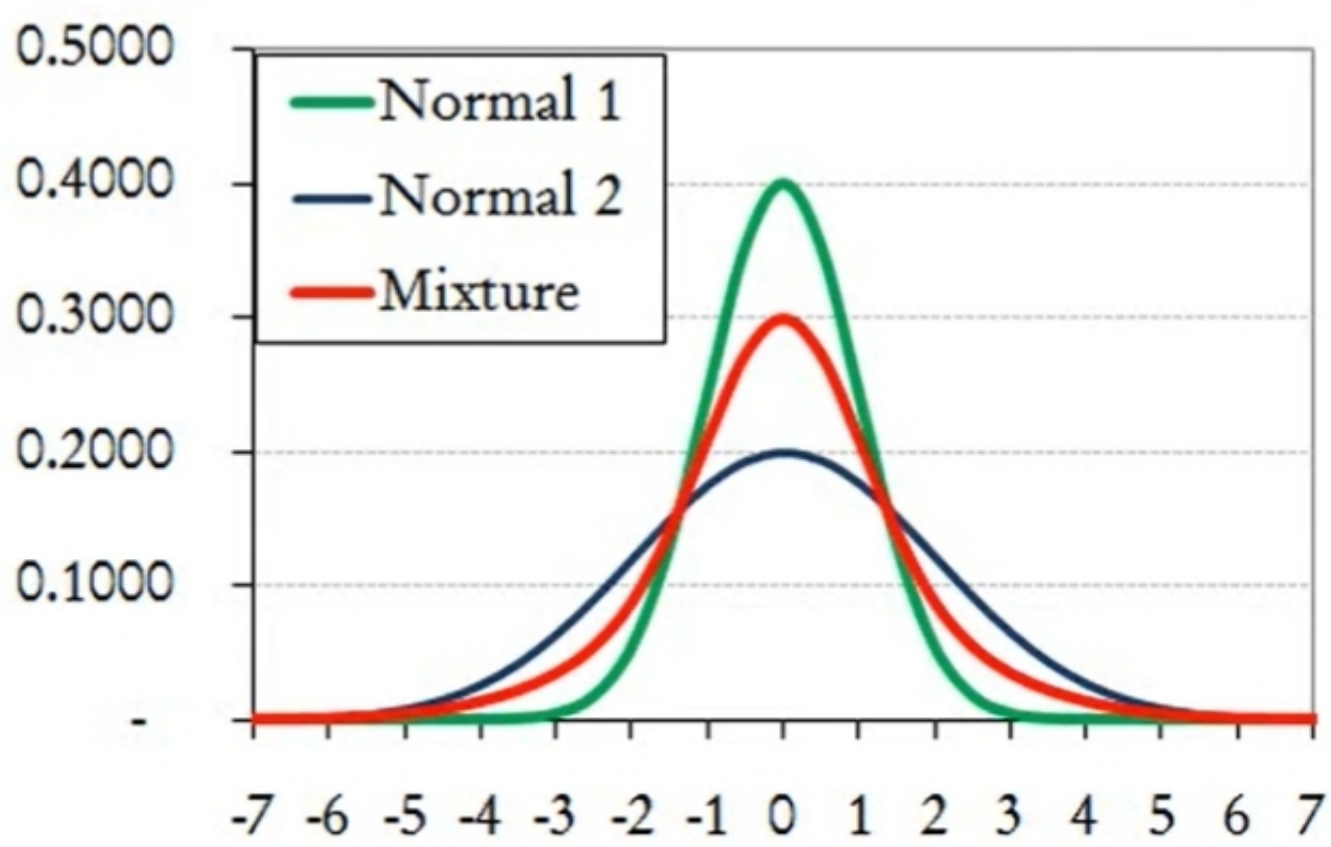
	Mean	Variance
Adding: $T = X + Y$	$\mu_T = \mu_X + \mu_Y$	$\sigma_T^2 = \sigma_X^2 + \sigma_Y^2$
Subtracting: $D = X - Y$	$\mu_D = \mu_X - \mu_Y$	$\sigma_D^2 = \sigma_X^2 + \sigma_Y^2$

Khan Academy, “Combining Random Variables” (https://www.khanacademy.org/math/ap-statistics/random-variables-ap/combining-random-variables/a/combining-random-variables-article, accessed 2021-11-15)

Mixture Distributions



Mixture Distributions



Bionic Turtle, “FRM: Normal mixture distribution” (https://www.youtube.com/watch?v=IkXxHwQZU2g, accessed 2021-11-15)

Combining LECs

Table 4. Separately ranked losses (X)

Hurricane			Earthquake		
Rank (Exceedance Probability)	Simulation Year	Loss (USD millions)	Rank (Exceedance Probability)	Simulation Year	Loss (USD millions)
1 (10%)	3	1,200	1 (10%)	9	750
2 (20%)	10	888	2 (20%)	5	215
3 (30%)	5	544	3 (30%)	3	0

AIR, “Modeling Fundamentals: Combining Loss Metrics” (https://www.air-worldwide.com/publications/air-currents/2012/Modeling-Fundamentals--Combining-Loss-Metrics/, accessed 2021-11-15)

Combining LECs

Table 5. Losses are combined first, *then* ranked (✓)

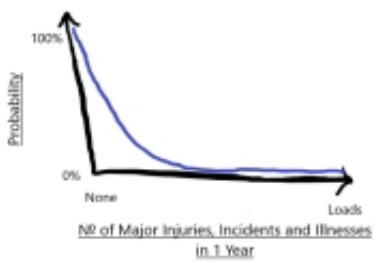
Simulation Year	Combined Loss (USD millions)	Rank (Exceedance Probability)	Simulation Year	Ranked Combined Loss (USD millions)
1	45	1 (10%)	3	1,200
2	9	2 (20%)	10	888
3	1,200	3 (30%)	9	764

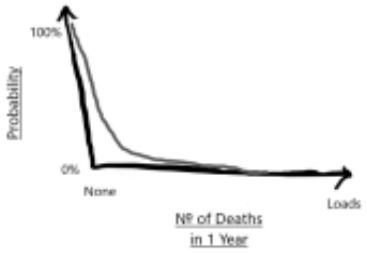
AIR, “Modeling Fundamentals: Combining Loss Metrics” (https://www.air-worldwide.com/publications/air-currents/2012/Modeling-Fundamentals--Combining-Loss-Metrics/, accessed 2021-11-15)

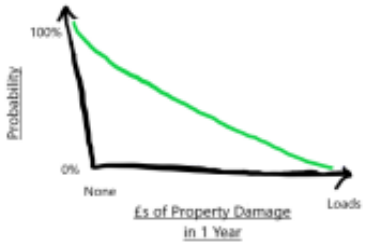
Defining Your Models

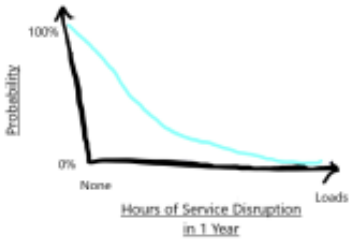
- Figure out you how you will compare/weight different scales
  - For simplicity, let’s just say we normalise each of our LEC-axes to a scale of $[0,1]$

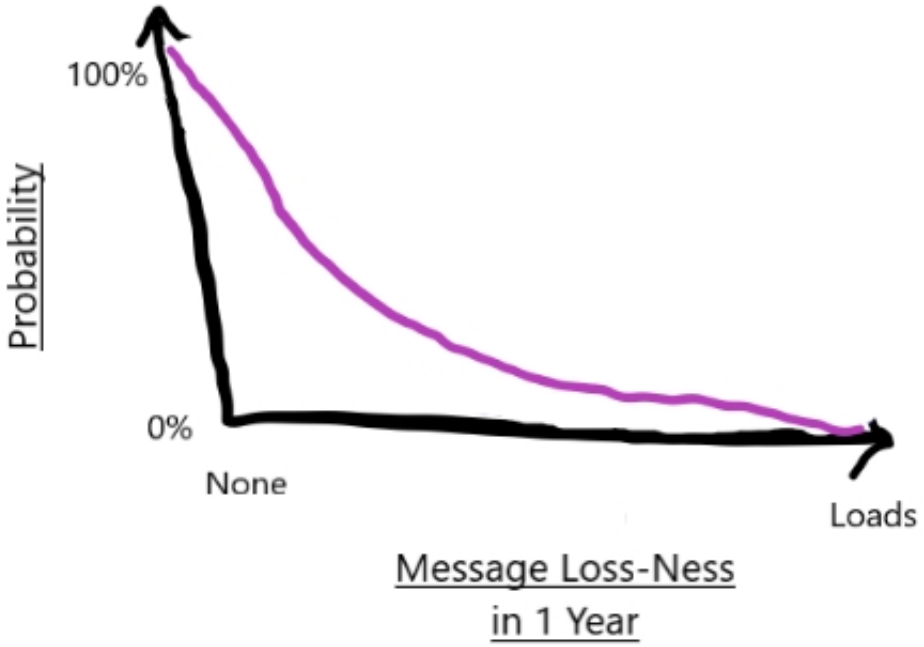
Putting It All Together











Questions from the Q&A

1. ‘How can you quantify losses in non-monetary terms?’

2. ‘What published research is there on the efficacy of implementing specific controls?’

Axon *et al.* (2021)

- Axon *et al.* “Practitioners’ Views on Cybersecurity Control Adoption and Effectiveness” (2021). https://dl.acm.org/doi/fullHtml/10.1145/3465481.3470038, accessed 2021-11-15.
  - ‘Cybersecurity practitioners possess a wealth of field knowledge in this area, yet there has been little academic work collecting and synthesising their views.’
  - ‘RQ1: How *effective* do security practitioners perceive different cybersecurity controls to be in addressing organisational cyber-risk?’
  - ‘RQ2: How are different cybersecurity controls *deployed* in practical environments?’

Such *et al.* (2016)

- Such *et al*. “Information Assurance Techniques: Perceived Cost Effectiveness” (2016). https://eprints.lancs.ac.uk/id/eprint/78969/1/information_assurance_techniques.pdf, accessed 2021-11-15.
  - ‘Despite this importance, the characteristics of these assurance techniques have not been comprehensively explored within academic research from the perspective of industry stakeholders.’

Table 5: Effectiveness of Assurance Techniques — P: Poor; F: Fair; G: Good; VG: Very Good; E: Excellent

Assurance Technique	Effectiveness					Total Resp.
	P	F	G	VG	E	
Review of [...] PPP	2%	27%	**46%**	19%	5%	93
[...] Self Assessment Form	31%	**35%**	29%	4%	2%	84
Architectural Review	1%	6%	**46%**	39%	8%	83
Configuration Review	-	21%	**41%**	35%	3%	80
Source Code Review	7%	9%	**49%**	28%	7%	69
Observation	2%	**39%**	38%	18%	4%	56
Interview	7%	22%	**36%**	32%	3%	72
Red Team Exercises	5%	5%	27%	**38%**	26%	66
Penetration Tests	-	5%	34%	**47%**	15%	88
Vulnerability Scan	6%	28%	**37%**	24%	5%	86
Social Engineering	5%	22%	**45%**	15%	14%	65
Threat Assessment	-	19%	**47%**	29%	5%	73
Static Analysis	2%	**41%**	39%	18%	-	44
Dynamic Analysis	2%	**40%**	36%	19%	2%	42
Fuzzing	5%	**39%**	32%	24%	-	41
Formal Verification	3%	28%	**40%**	30%	-	40
Cryptographic Validation	-	20%	**49%**	24%	7%	41
Emanation Security Analysis	6%	**40%**	34%	20%	-	35
Witnessed Test	10%	28%	**45%**	15%	3%	40
Public Review	16%	**37%**	26%	13%	8%	38

Such, Vidler, Seabrook & Rashid (2015)

- Such, Vidler, Seabrook & Rashid. “Cyber Security Controls Effectiveness: A Qualitative Assessment of Cyber Essentials” (2015). https://eprints.lancs.ac.uk/id/eprint/74598/4/SCC_2015_02_CS_Controls_Effectiveness.pdf, accessed 2021-11-15.
  - ‘The purpose of this report is to investigate the effectiveness of the Cyber Essentials controls in mitigating ‘commodity-level’ attacks attempting to exploit vulnerabilities in Small and Medium Enterprise (SME) networks.’

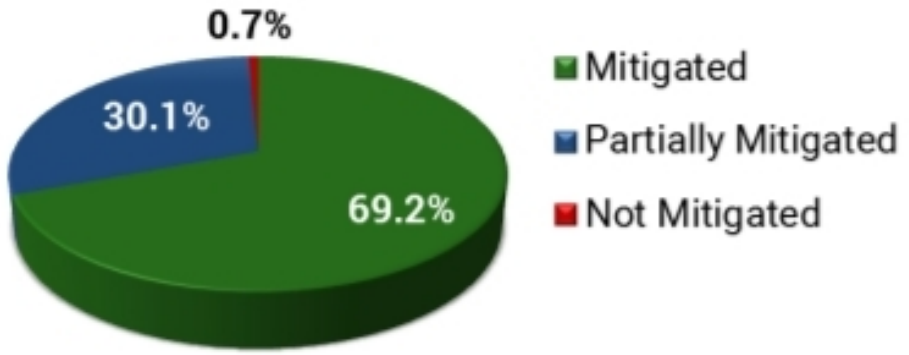


Figure 1: *Cyber Essentials Aggregated Vulnerability Mitigation Results*

Questions from the Q&A

1. ‘How can you quantify losses in non-monetary terms?’

2. ‘What published research is there on the efficacy of implementing specific controls?’

3. ‘How easy is it to augment the data generating the LEC with firewall output, etc.? **Are you aware of any tooling?**’

Here’s One I Made Earlier

Threat Intelligence Service

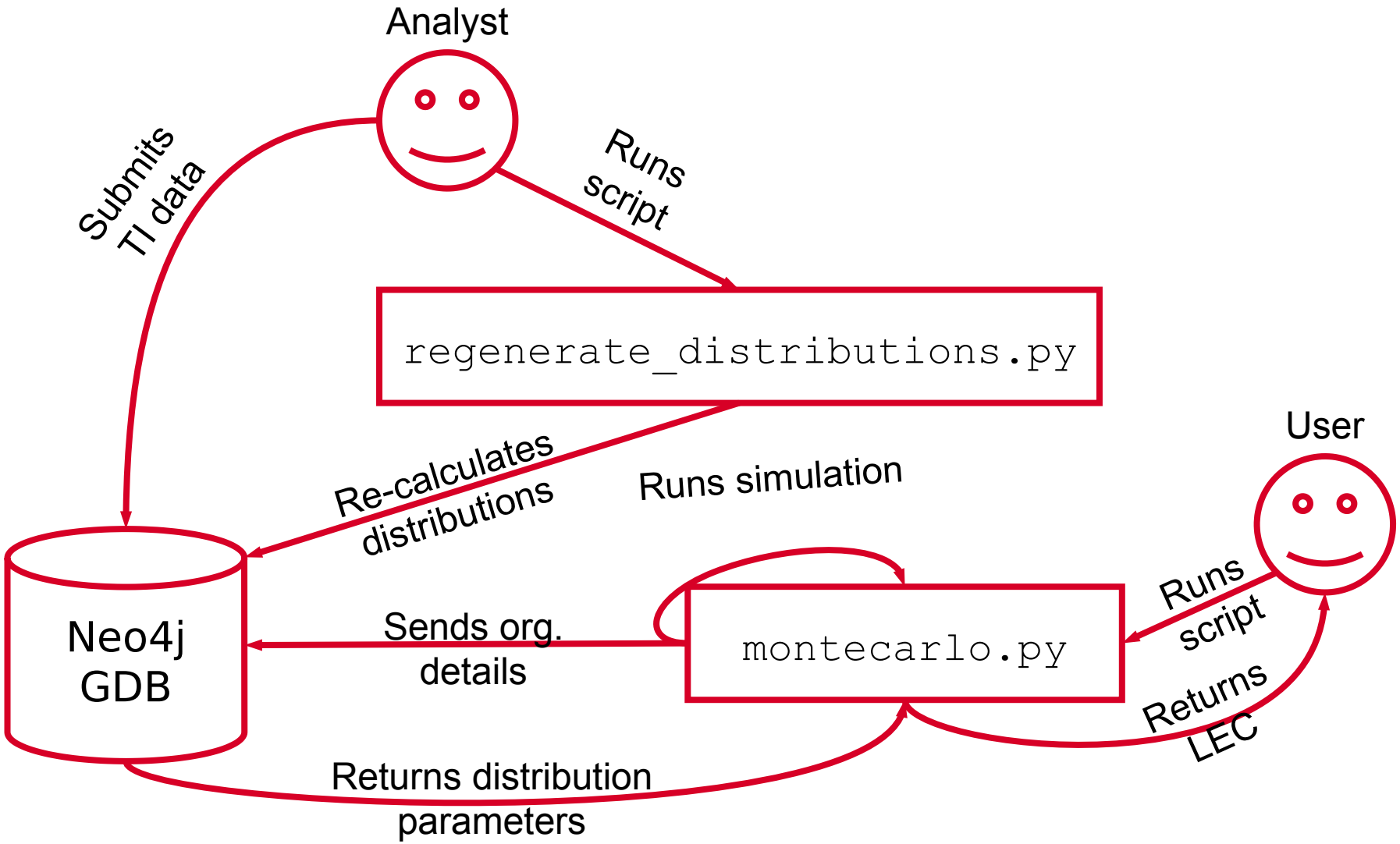
- https://code.bengoldsworthy.net/Rumperuu/Threat-Intelligence-Service

- Comprises a handful of scripts
  - Python (and a prototype in R)

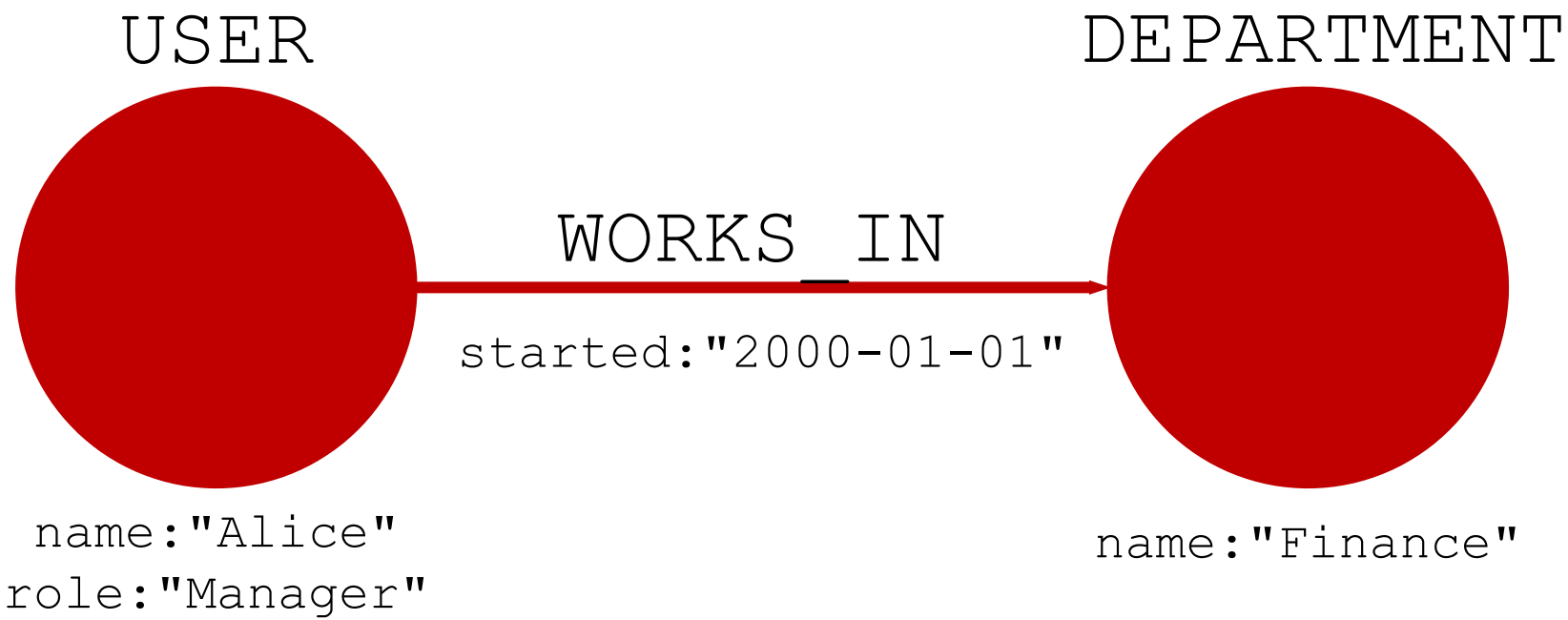
- Underpinned by Neo4j Graph Database
  - Cypher Query Language

- (Currently) licensed under the CRAPL, ‘an academic-strength open source license’
  - https://matt.might.net/articles/crapl/

Architecture

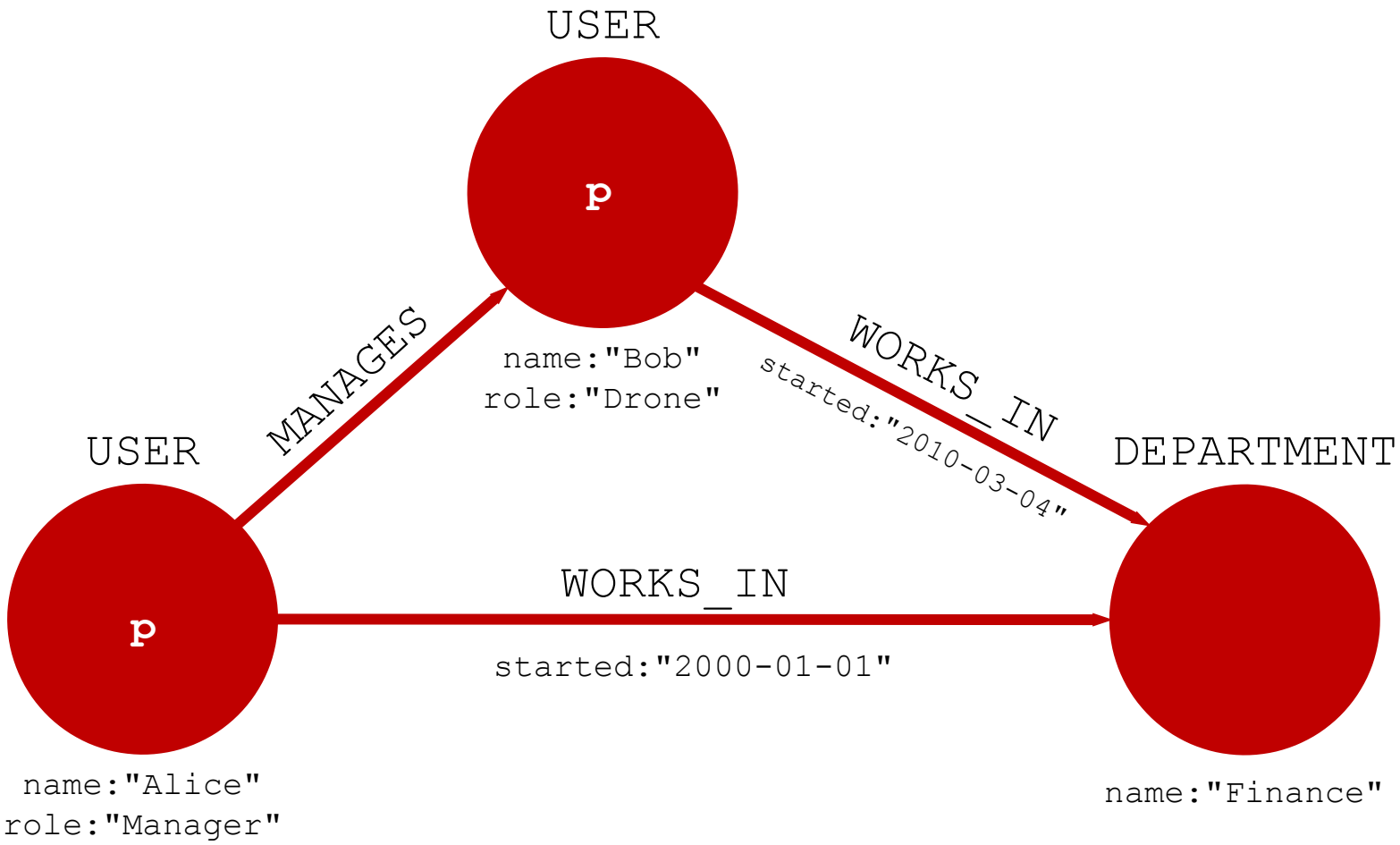


Intro. to Graph Databases



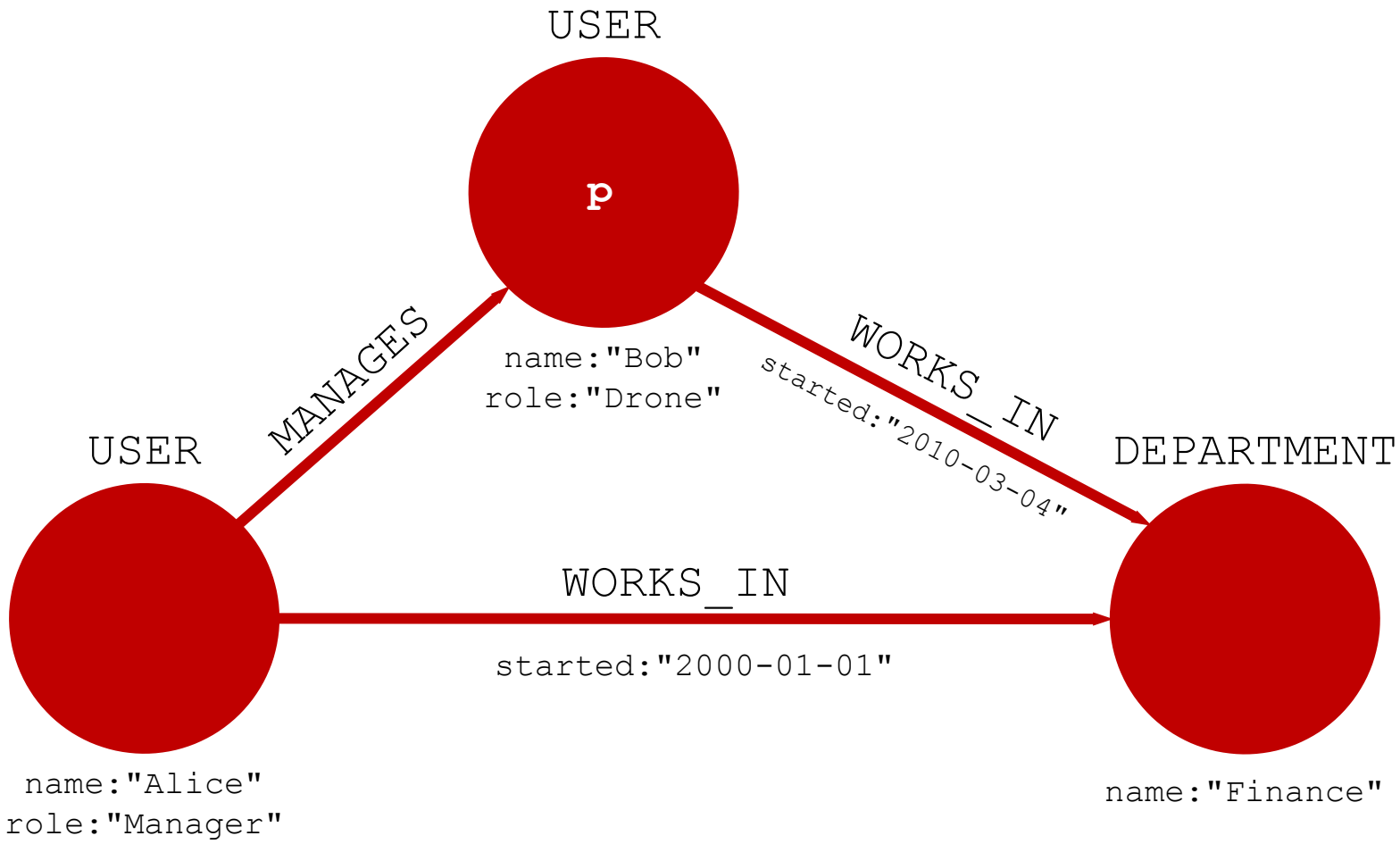
Intro. to Graph Databases

`MATCH (p:USER) RETURN p`



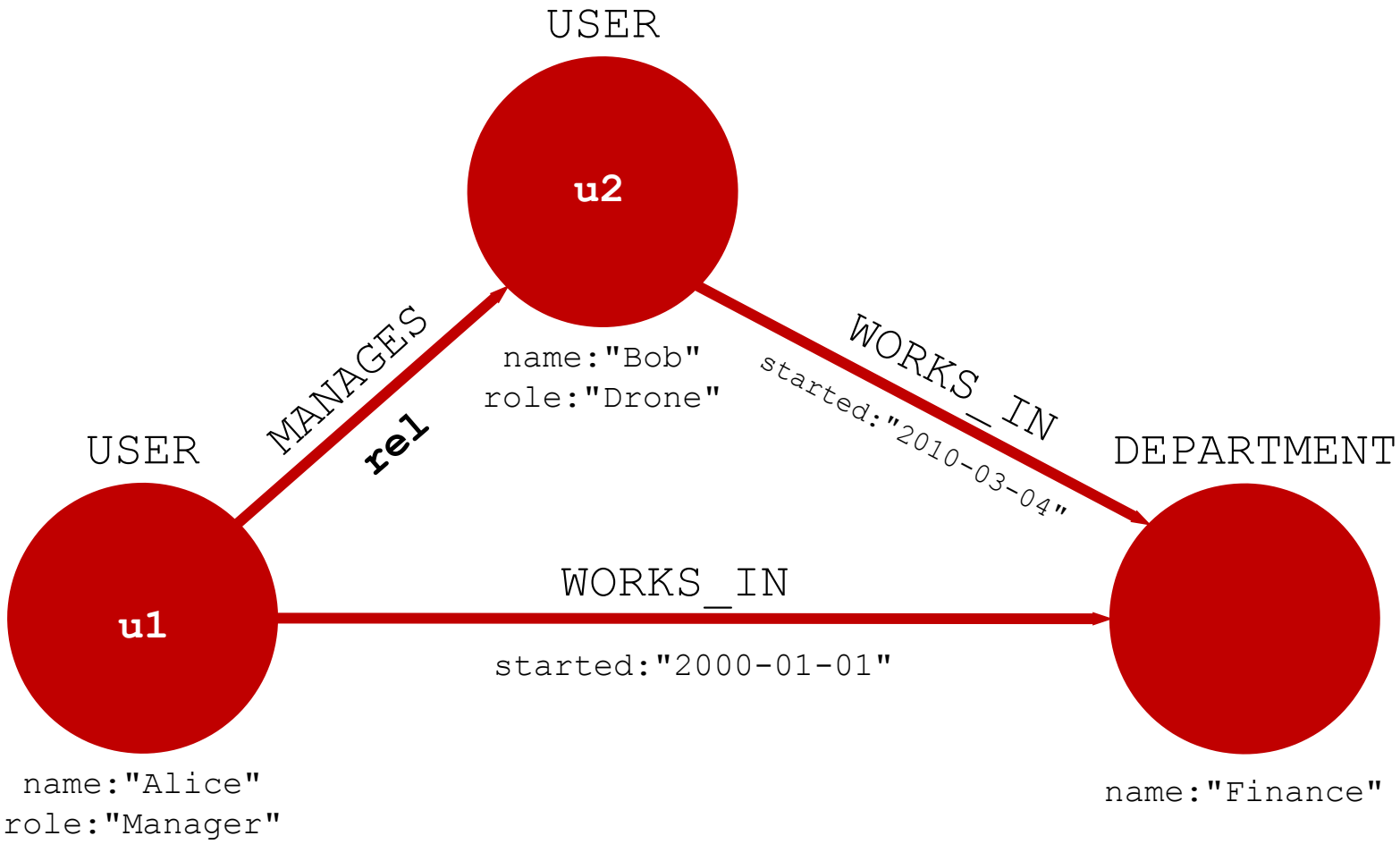
Intro. to Graph Databases

MATCH (p:USER {name:"Bob"}) RETURN p

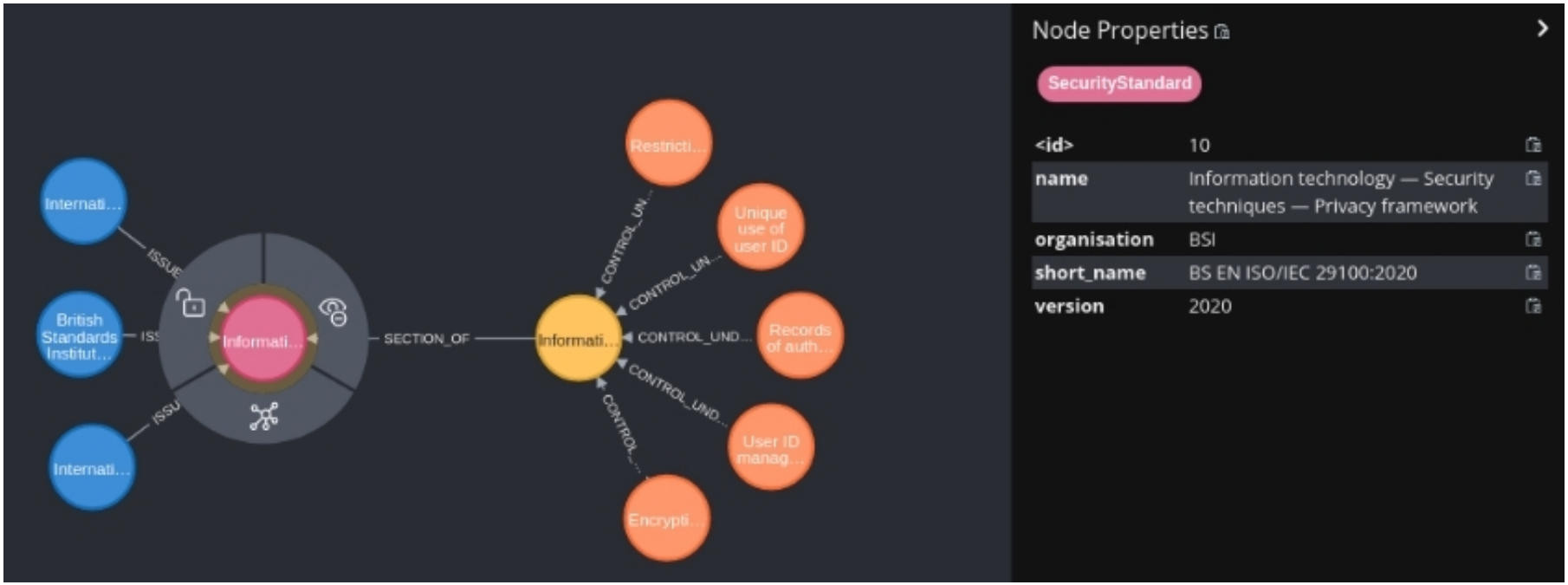


Intro. to Graph Databases

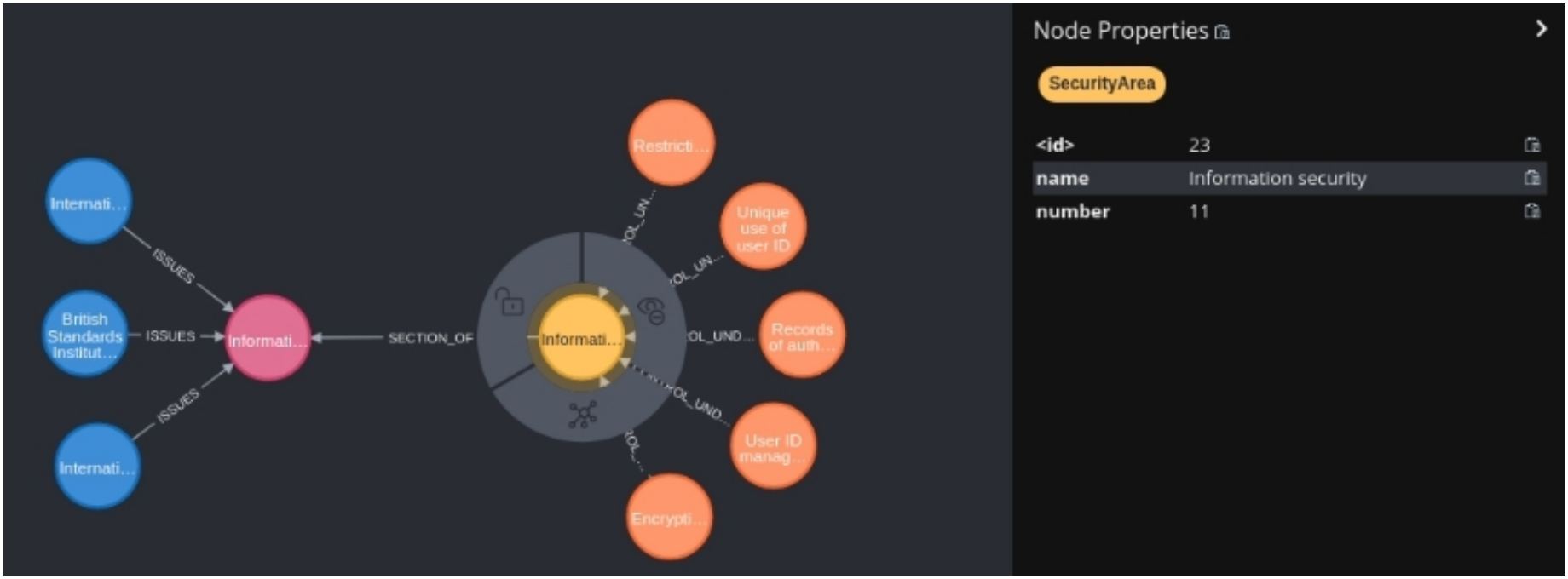
```
MATCH (u1:User)-[rel]->(u2:User)
RETURN u1, rel, u2
```



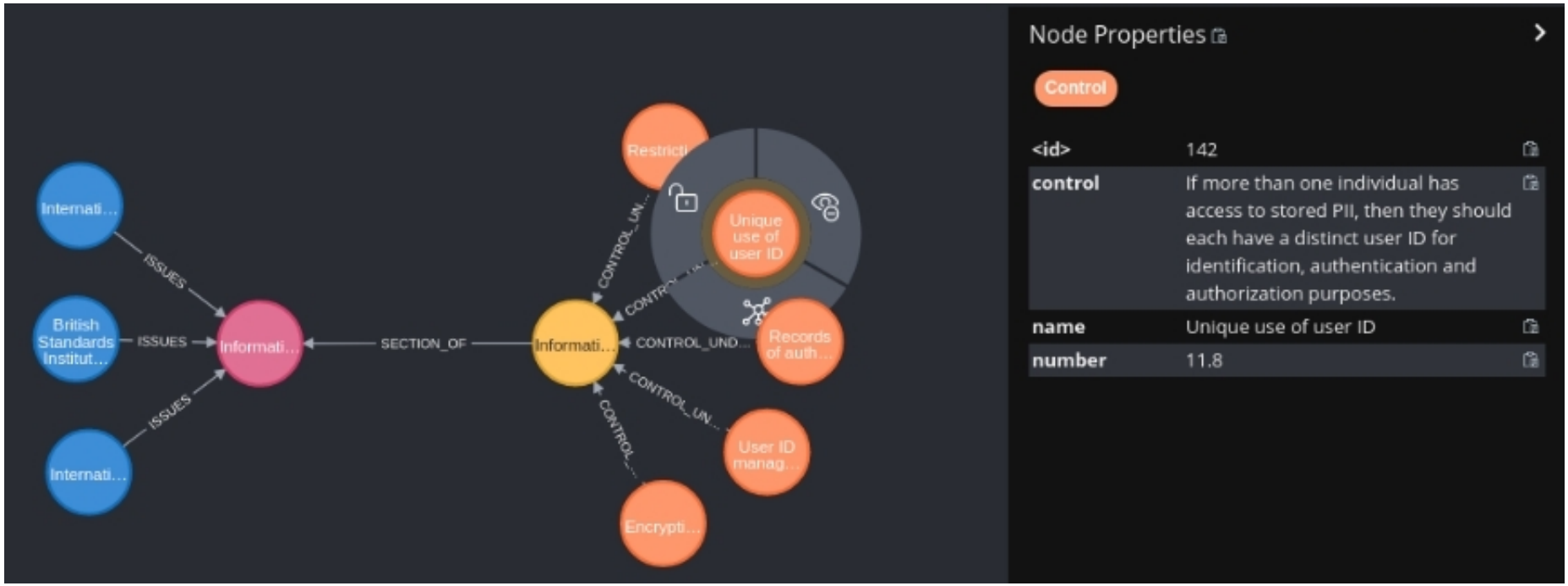
TI Data Structure
Standards and Controls



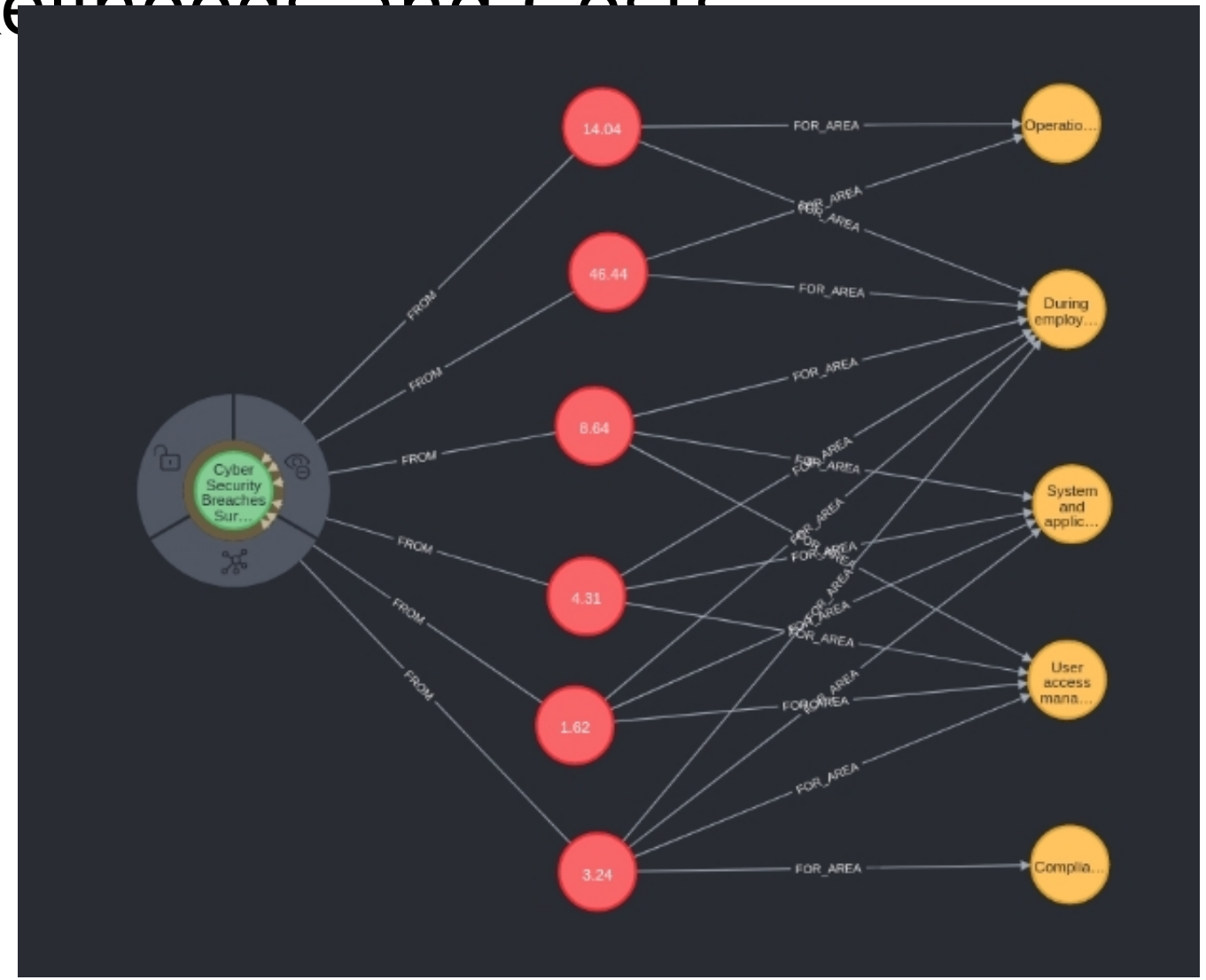
TI Data Structure
Standards and Controls



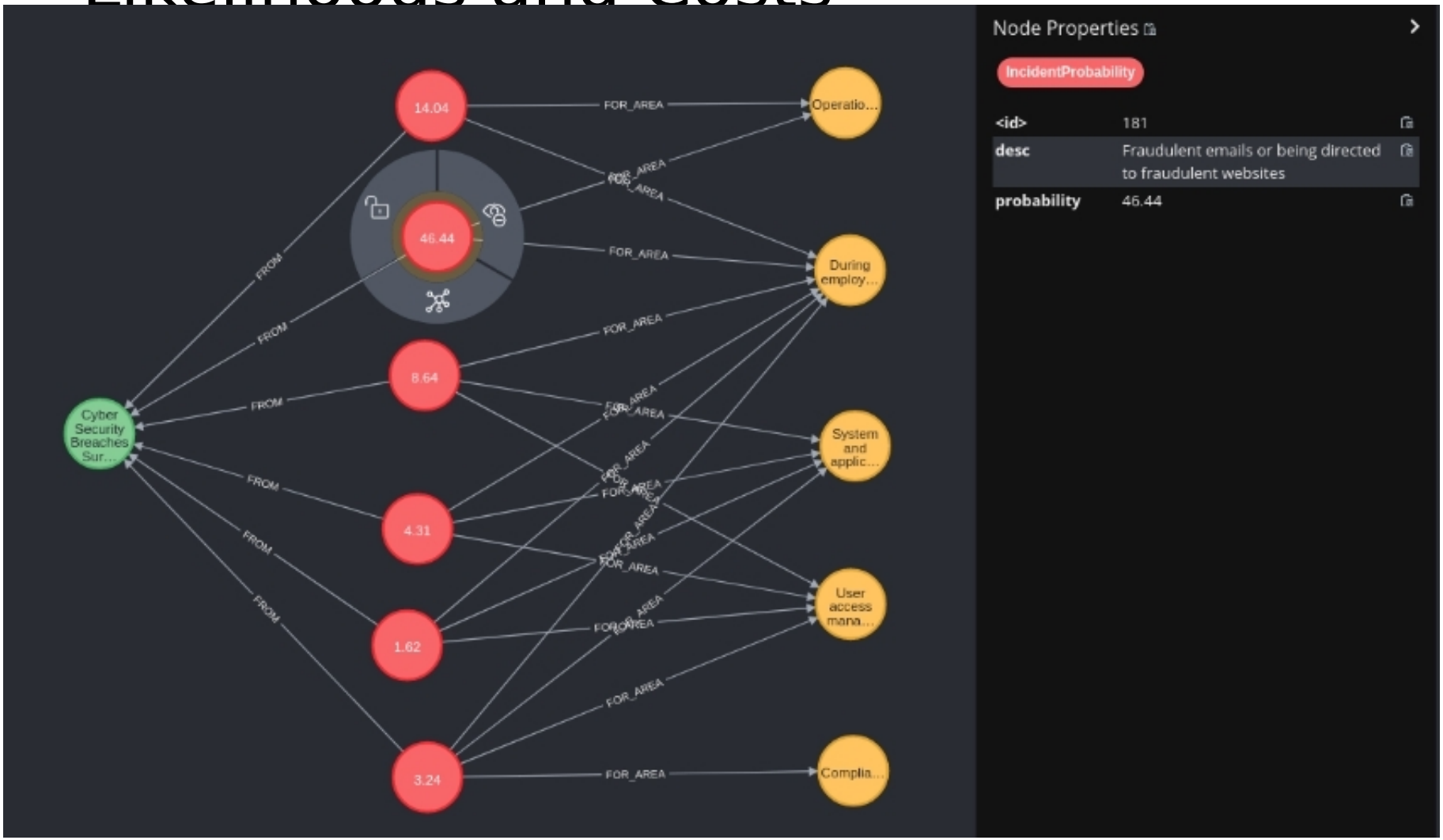
TI Data Structure
Standards and Controls



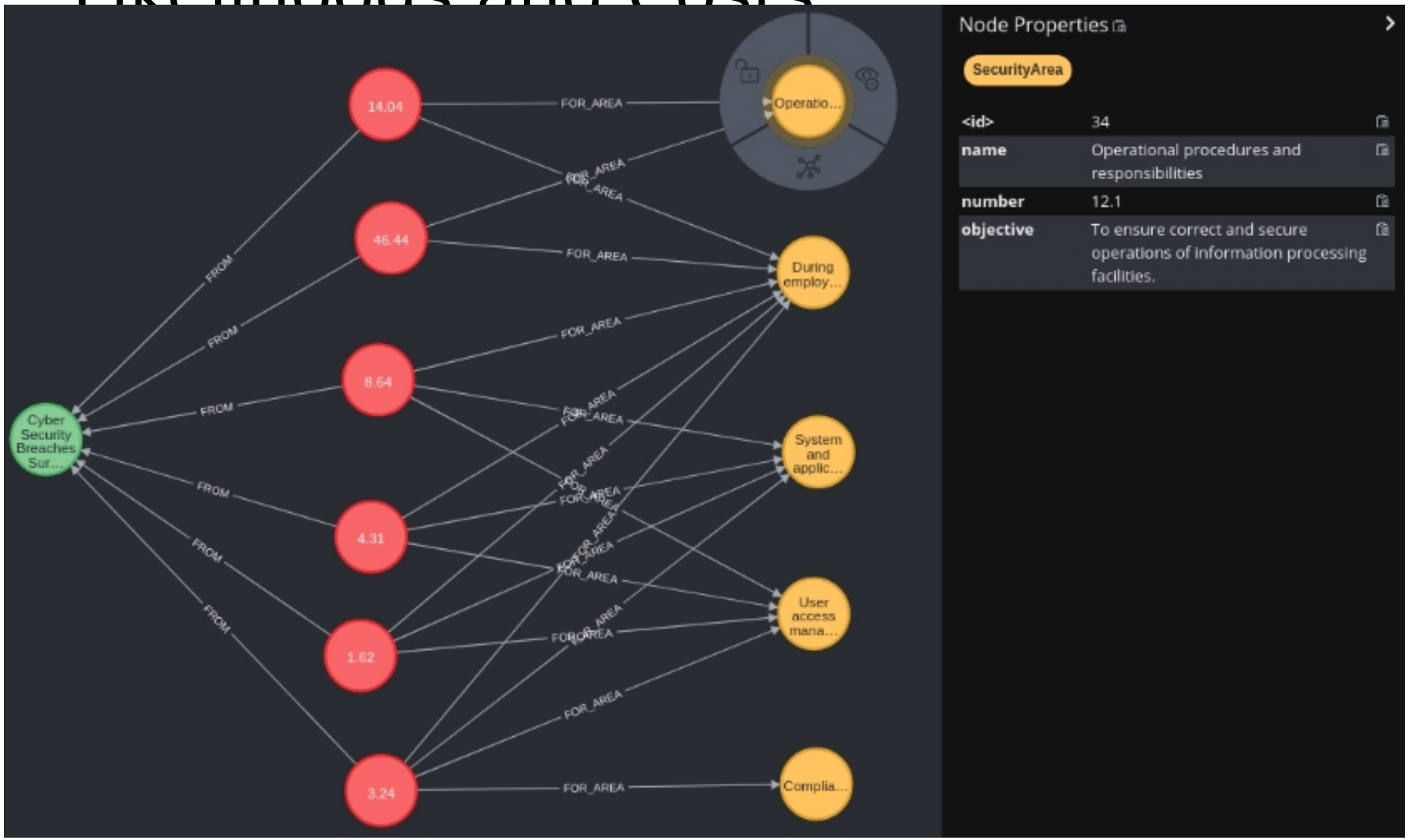
TI Data Structure
Likelihoods and Costs



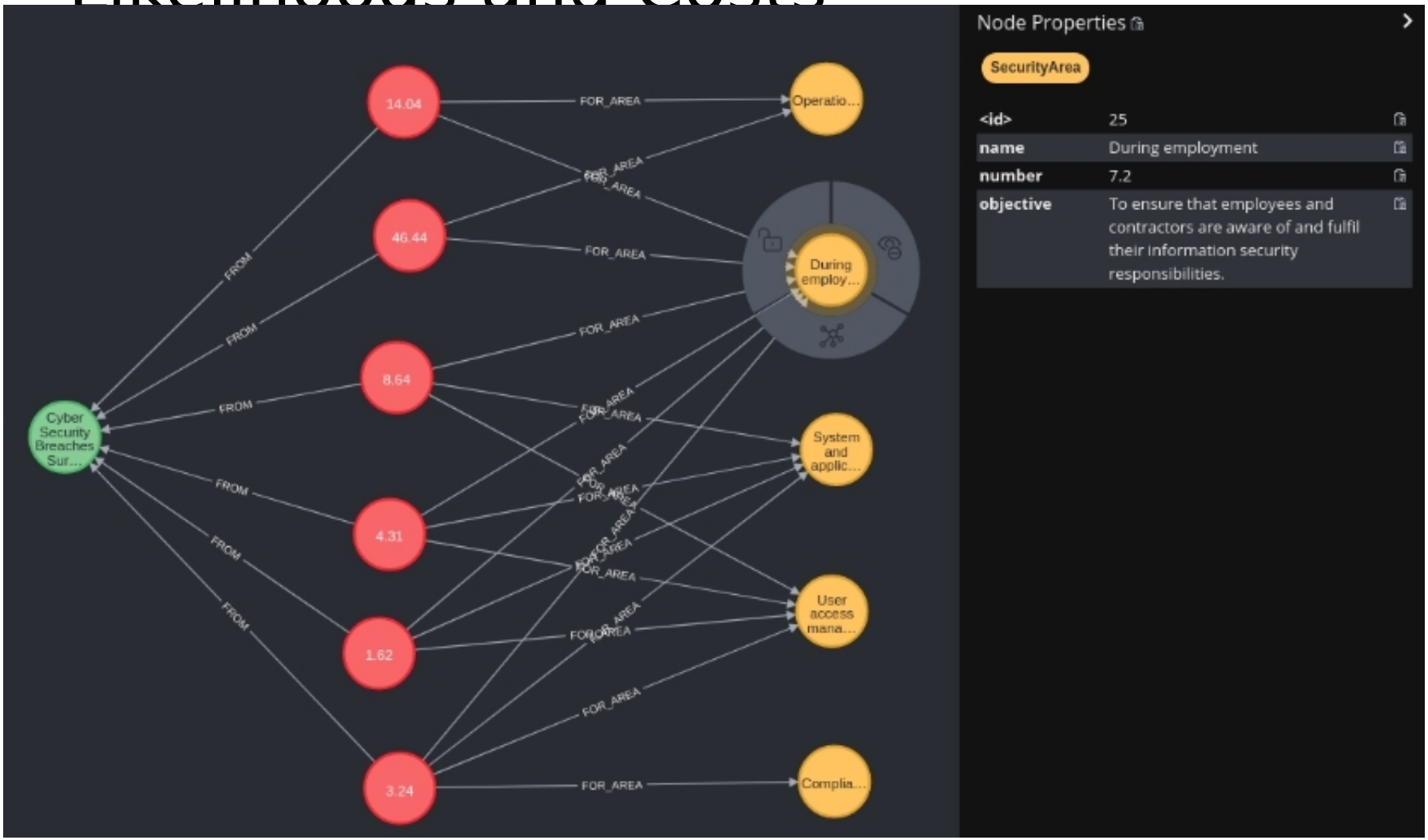
TI Data Structure Likelihoods and Costs



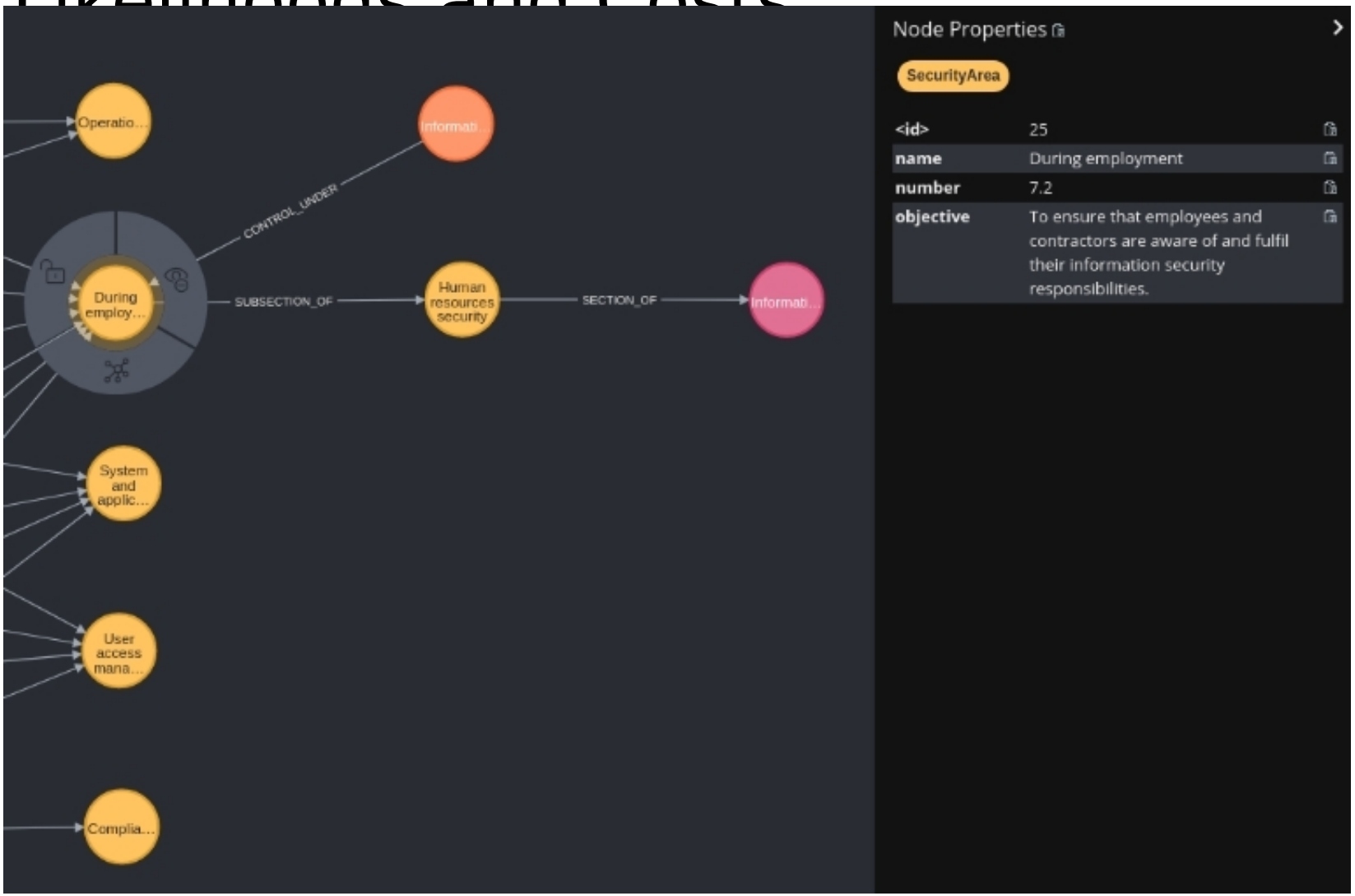
TI Data Structure Likelihoods and Costs



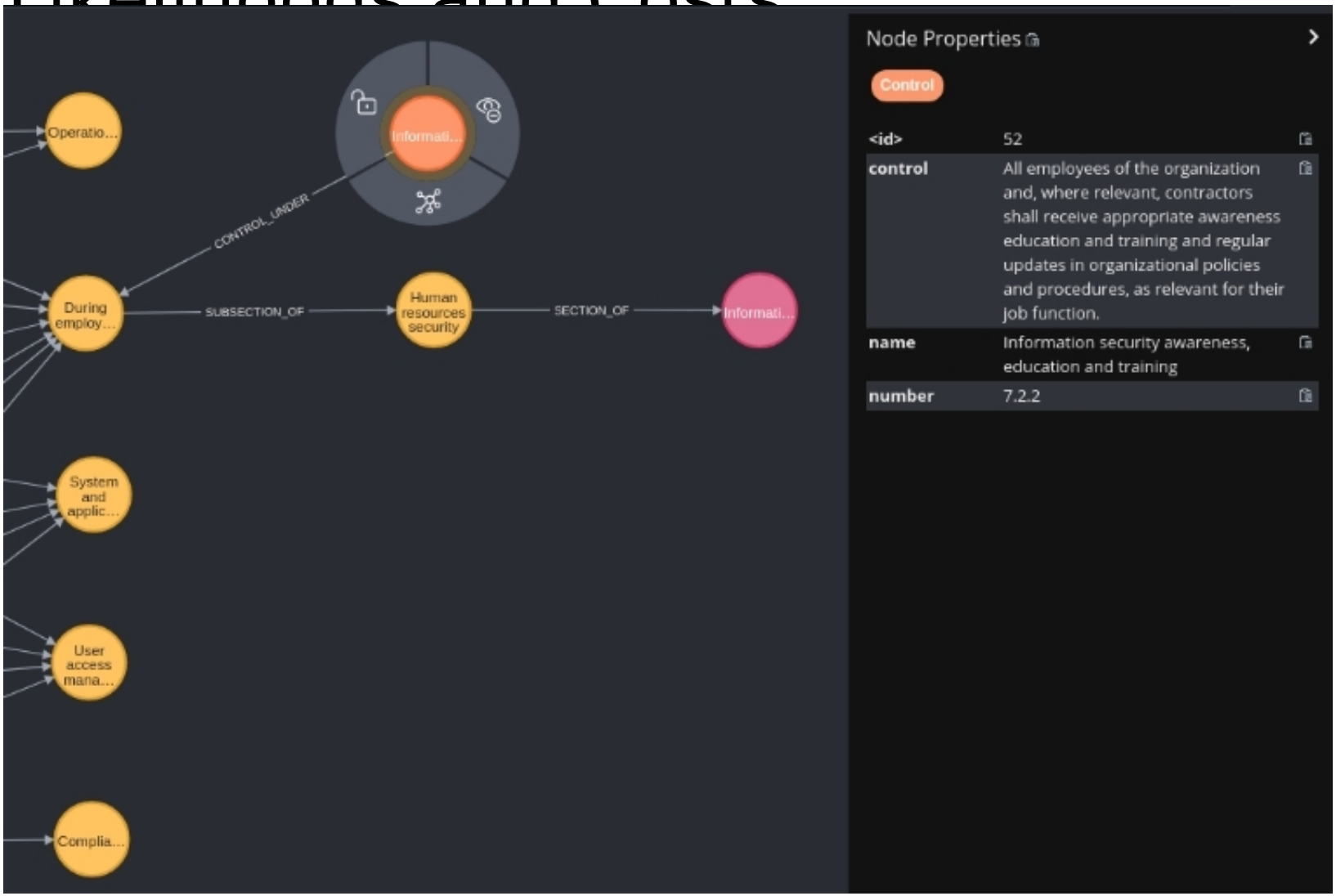
TI Data Structure
Likelihoods and Costs



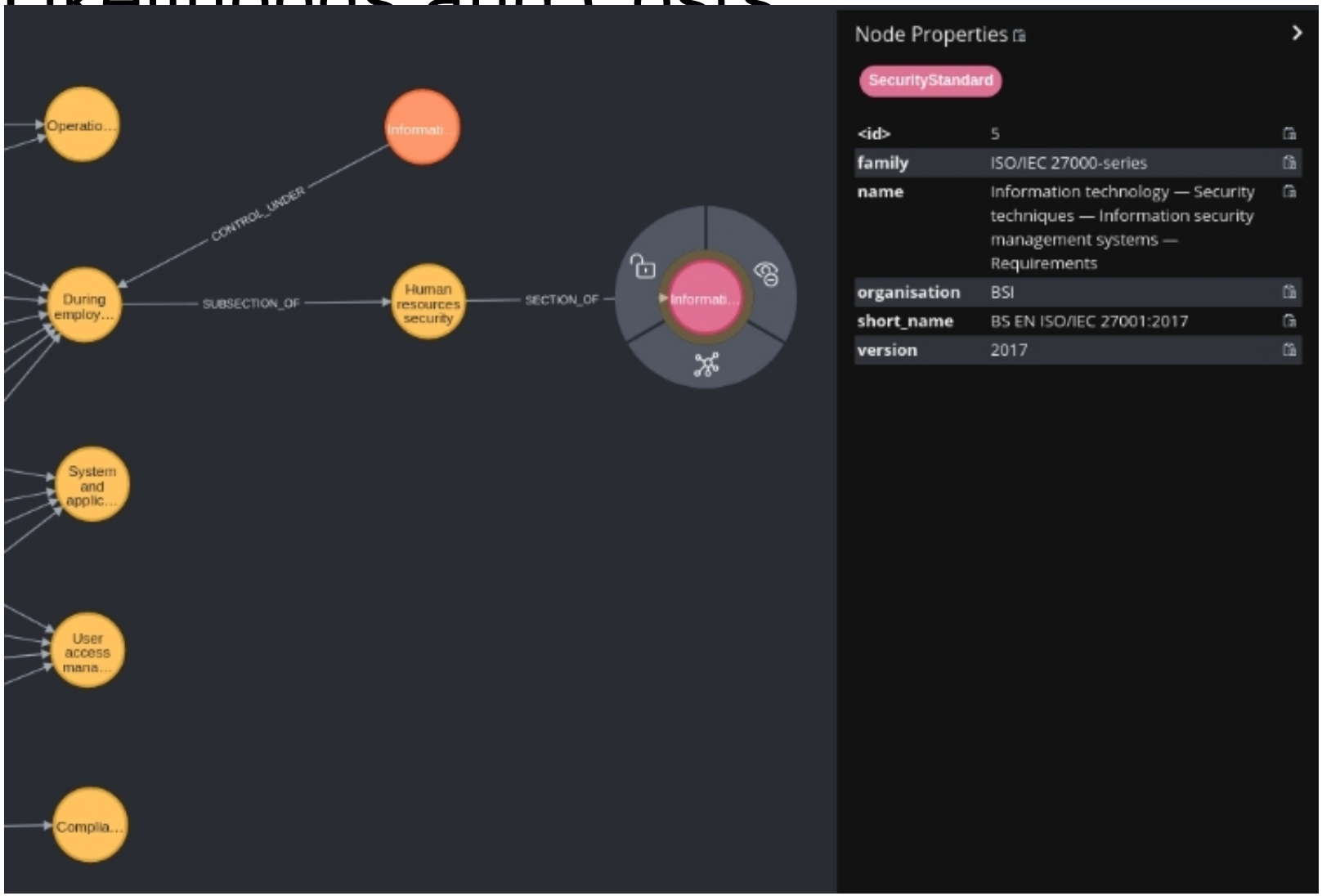
TI Data Structure Likelihoods and Costs



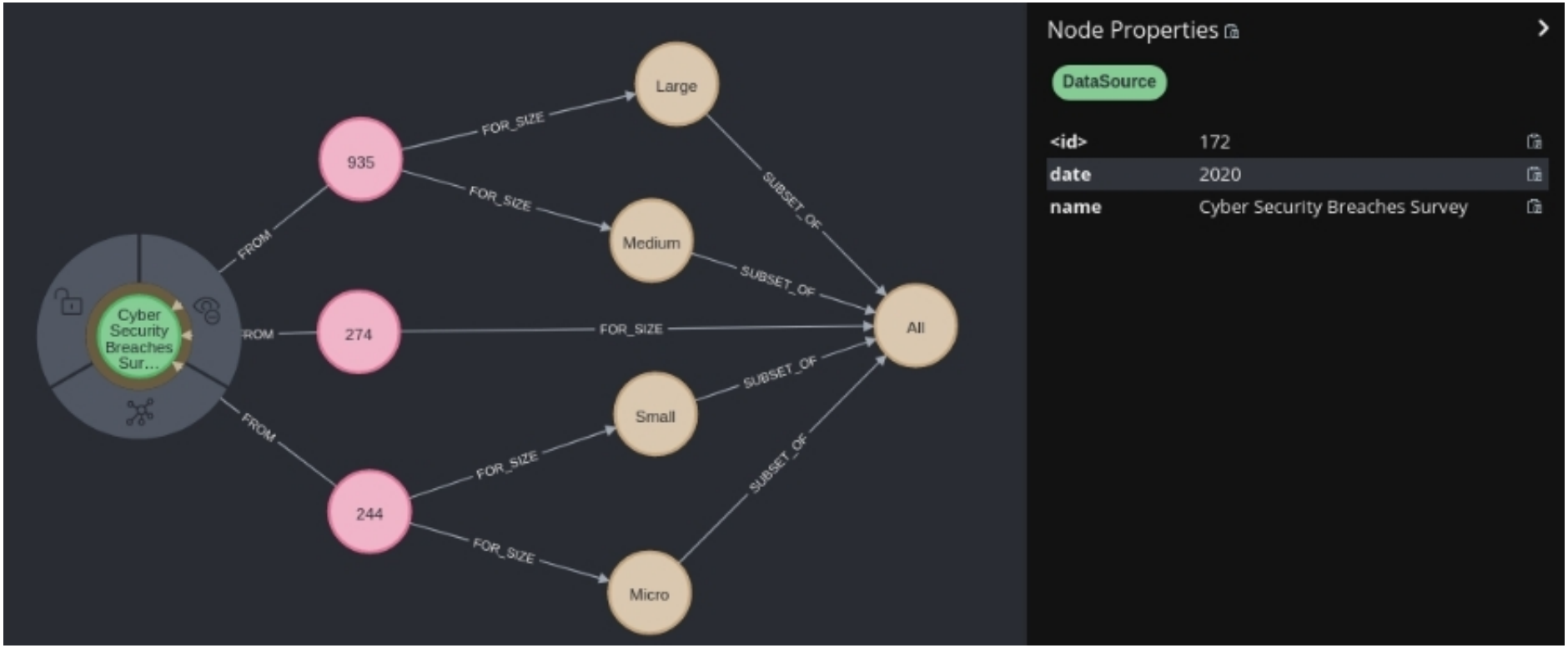
TI Data Structure Likelihoods and Costs



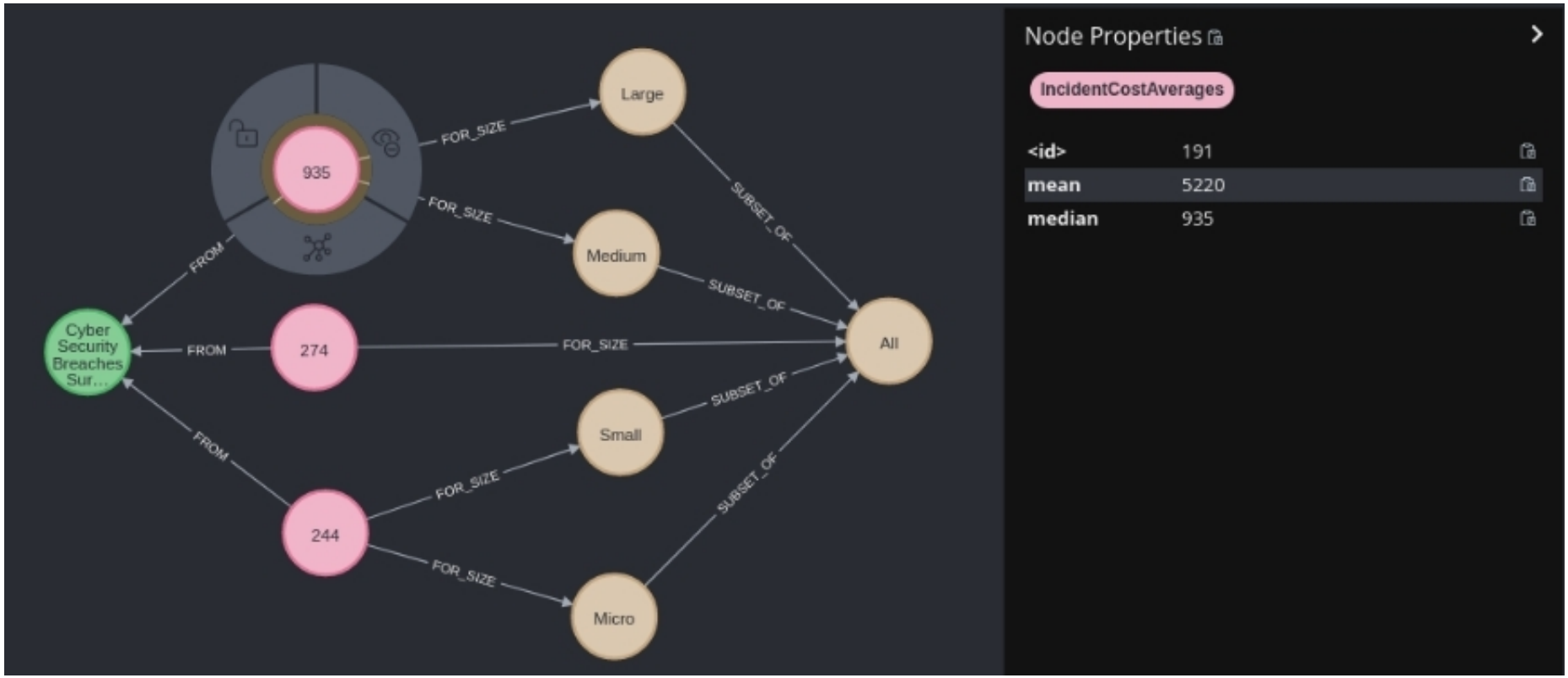
TI Data Structure Likelihoods and Costs



TI Data Structure Likelihoods and Costs



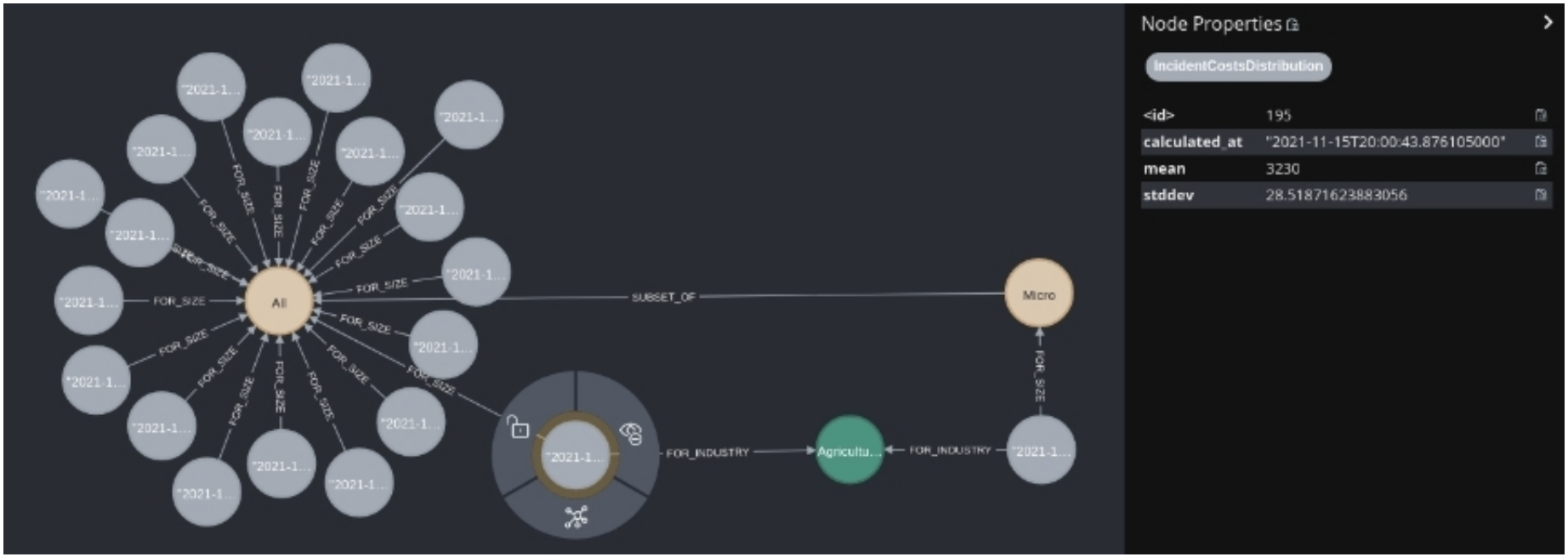
TI Data Structure Likelihoods and Costs



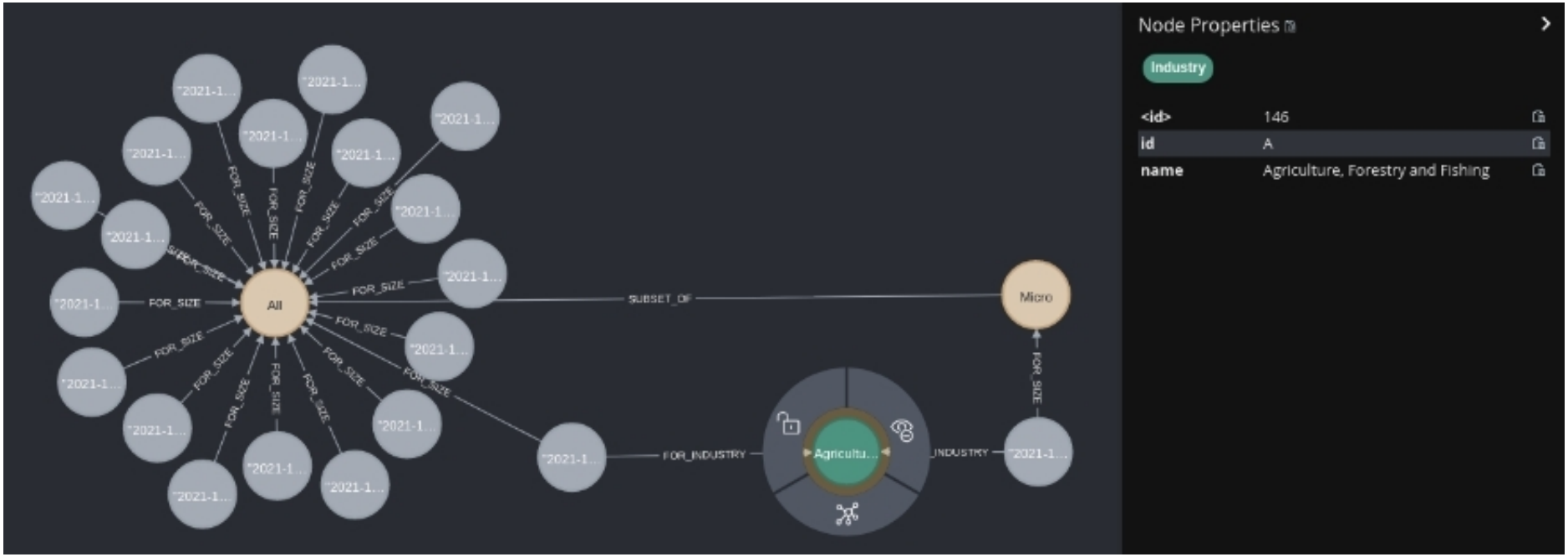
(Re)generating Distributions

```
Level 31: New incident frequency distribution successfully generated for '('All', 'Activities of extraterritorial organisations and bodies')'.
Level 31: New incident costs distribution successfully generated for '('All', 'Activities of extraterritorial organisations and bodies')'.
Level 31: New incident frequency distribution successfully generated for '('Micro', 'All')'.
Level 31: New incident costs distribution successfully generated for '('Micro', 'All')'.
Level 31: New incident frequency distribution successfully generated for '('Micro', 'Agriculture, Forestry and Fishing')'.
Level 31: New incident costs distribution successfully generated for '('Micro', 'Agriculture, Forestry and Fishing')'.
Level 31: New incident frequency distribution successfully generated for '('Micro', 'Mining and quarrying')'.
Level 31: New incident costs distribution successfully generated for '('Micro', 'Mining and quarrying')'.
Level 31: New incident frequency distribution successfully generated for '('Micro', 'Manufacturing')'.
Level 31: New incident costs distribution successfully generated for '('Micro', 'Manufacturing')'.
Level 31: New incident frequency distribution successfully generated for '('Micro', 'Electricity, Gas, Steam and air conditioning')'.
Level 31: New incident costs distribution successfully generated for '('Micro', 'Electricity, Gas, Steam and air conditioning')'.
```

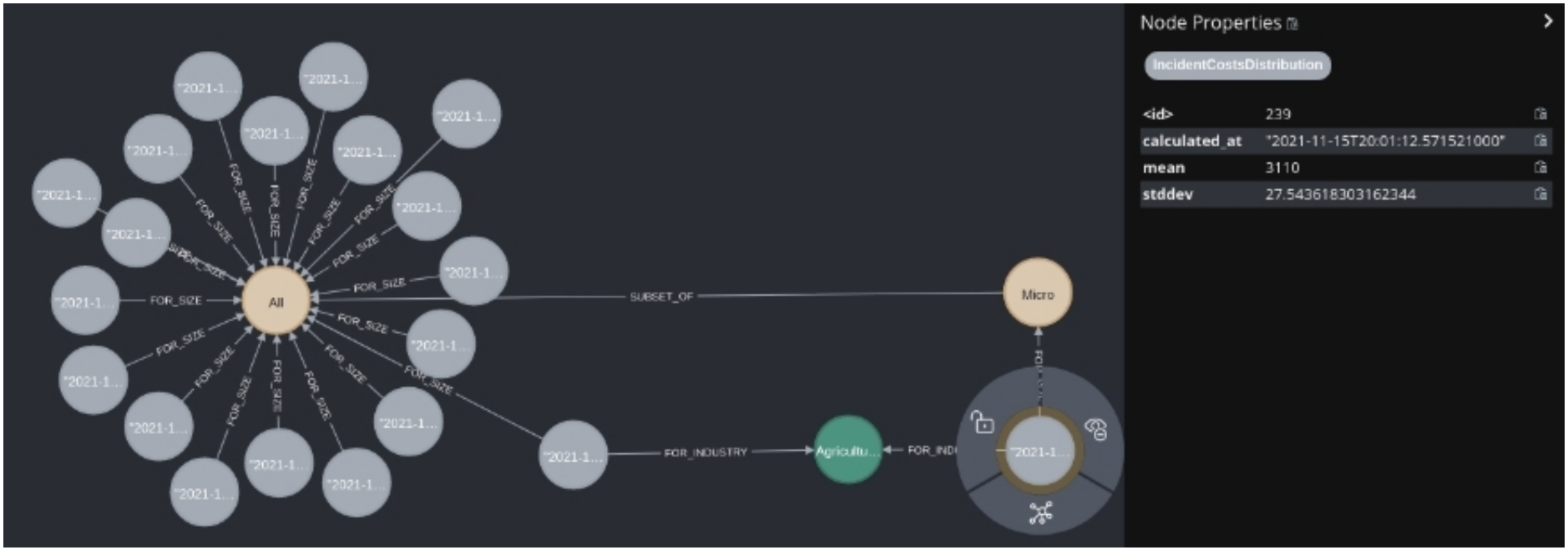
TI Data Structure
Calculated Distribution Parameters



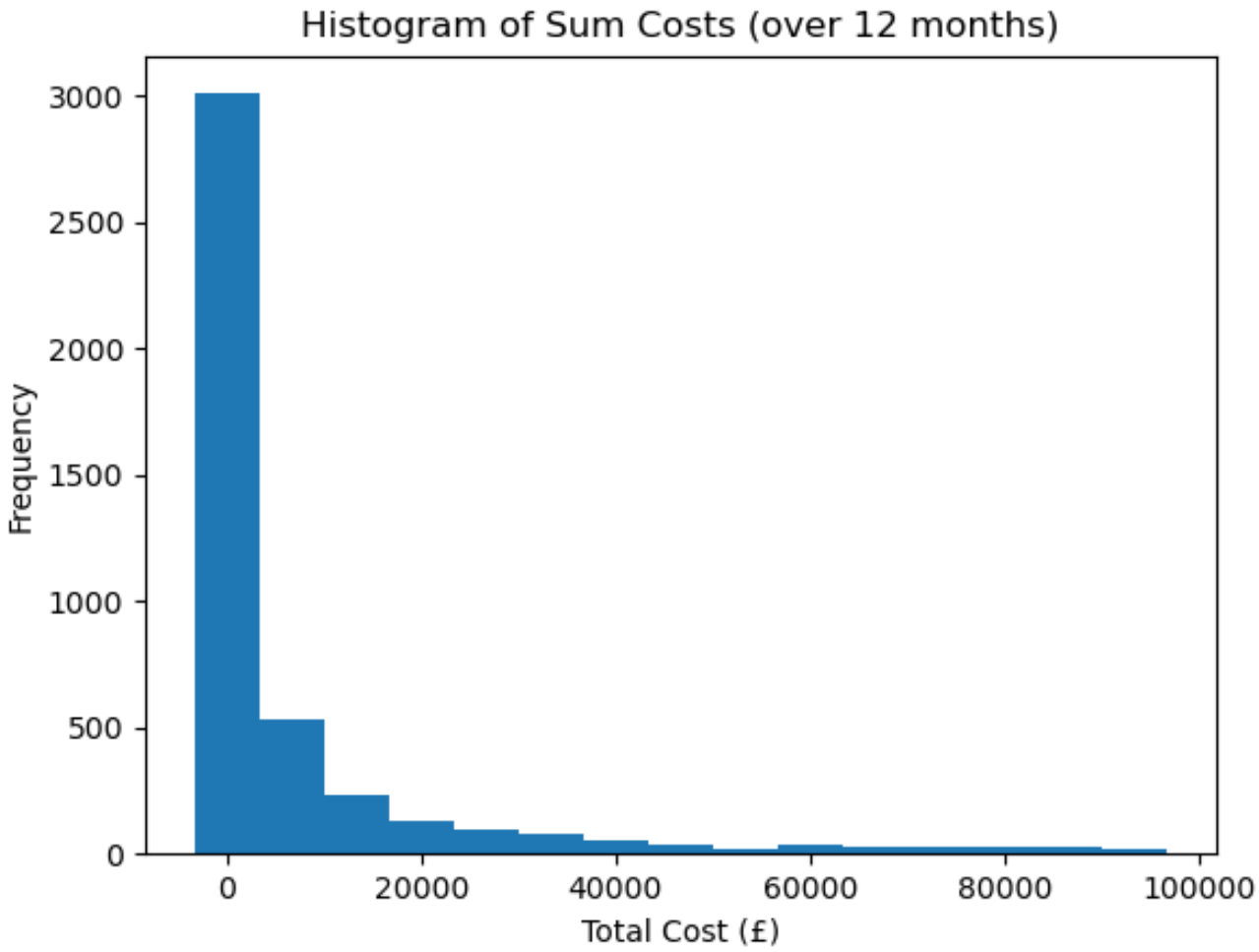
TI Data Structure
Calculated Distribution Parameters



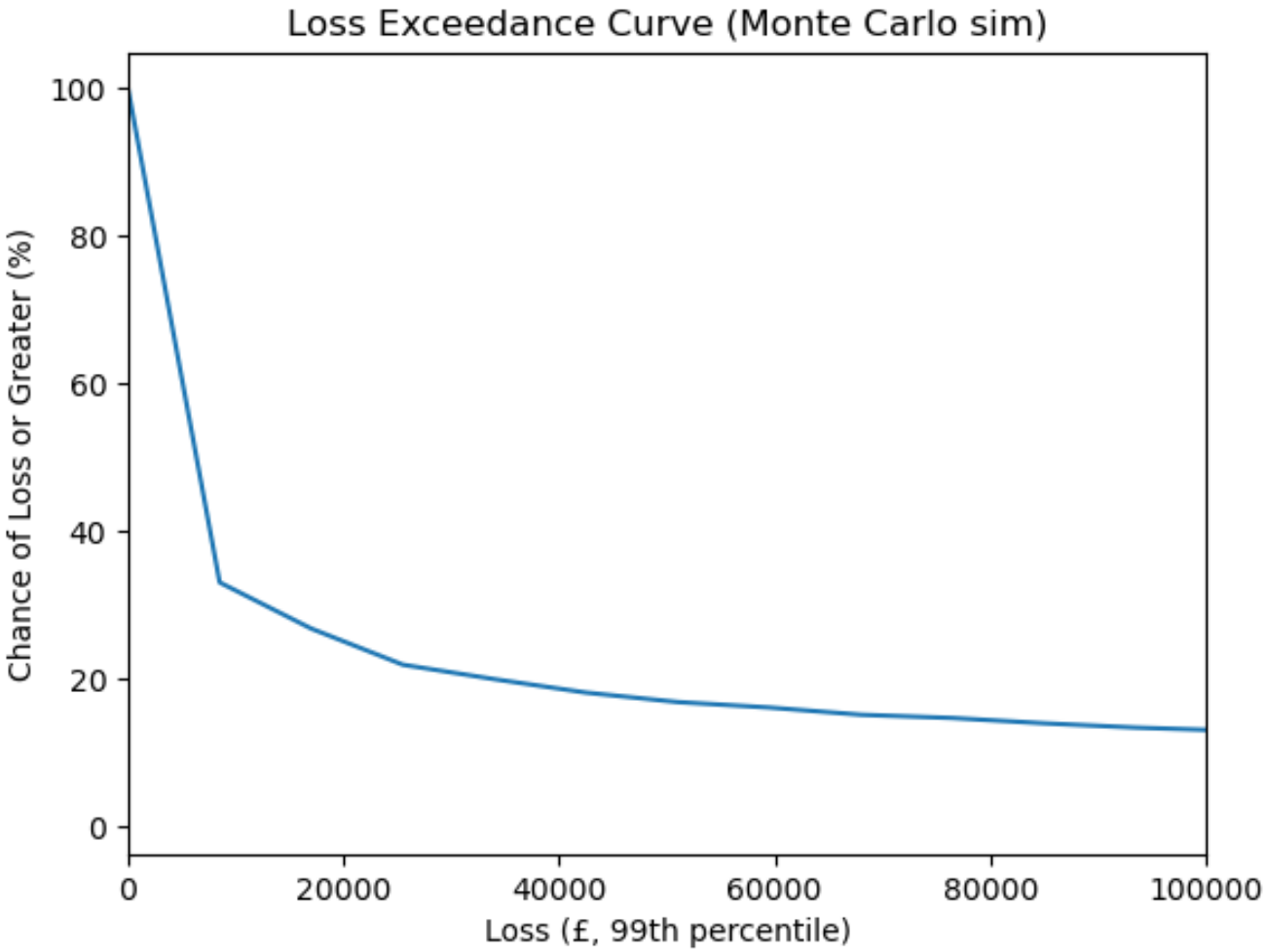
TI Data Structure
Calculated Distribution Parameters



Simulation Results



Simulation Results



Future Development Plans

- Watch this space...

Questions, Quibbles, Quomments?

Appendix
Calibration Test

- ‘Much research has already been done on this point and two findings are clear:
    1. Most people are bad at assigning probabilities; but
    2. Most people can also be trained to be very good at it.’

- This Appendix contains a short test you can use to assess your own, uncalibrated estimation abilities.

Appendix
Calibration Test

1. For the first set of questions, answer with a lower bound and an upper bound that represent your 90% confidence interval
   - That is, you should feel 90% confident that the answer to the question will be within the range you have given
   - E.g., if the question asks when a battle in WWII took place, you know your 90% CI *must* be between 1939–1945, but if you're a military buff you might be able to narrow that range down to a single year

2. For the second set of questions, answer true or false and then your level of confidence that you are correct (out of 50%, 60%, 70%, 80%, 90% or 100% confidence)

Appendix
Calibration Test

#	Question	Lower Bound	Upper Bound
1	In 1938, a British steam locomotive set a new speed record by going how fast (mph)?		
2	In what year did Sir Isaac Newton publish the *Universal Laws of Gravitation*?		
3	How many mm long is a typical business card?		
4	The Internet (then called ‘ARPANET’) was established as a military communications system in what year?		
5	In what year was William Shakespeare born?		
6	What is the air distance between New York and Los Angeles (km)?		
7	What percentage of a square could be covered by a circle of the same width?		
8	How old was Charlie Chaplain when he died?		
9	What is the weight, in lbs, of the first edition of *How to Measure Anything*?		
10	The TV show *Big Brother* first aired on what date?		

Appendix
Calibration Test

#	Question	True/ False	Confidence That You Are Correct (circle one)
1	The ancient Romans were conquered by the ancient Greeks		50% 60% 70% 80% 90% 100%
2	There is no species of three-humped camels		50% 60% 70% 80% 90% 100%
3	A gallon of oil weighs less than a gallon of water		50% 60% 70% 80% 90% 100%
4	Mars is always further away from Earth than Venus		50% 60% 70% 80% 90% 100%
5	Germany won the first World Cup		50% 60% 70% 80% 90% 100%
6	Napoleon was born on the island of Corsica		50% 60% 70% 80% 90% 100%
7	‘M’ is one of the three most commonly-used letters in English		50% 60% 70% 80% 90% 100%
8	In 2002, the price of the average new desktop computer purchased in the US was under $1,500		50% 60% 70% 80% 90% 100%
9	Lyndon B. Johnson was a governor before becoming vice president		50% 60% 70% 80% 90% 100%
10	A kilogram is more than a pound		50% 60% 70% 80% 90% 100%

Appendix
Calibration Test

- You can now test whether the ranges you gave truly reflect your 90% CI:
    1. Take one of your answer ranges
    2. Imagine you are offered the chance to win £1,000 one of two ways:
        A. You win £1,000 if the true answer is within your range, and nothing if it is not; or
        B. You spin a dial divided into two unequal slices, one of which comprises 90% (or, for the second set of questions, the CI value you circle) of it, and if the dial lands on that slice you win £1,000
    3. Which way would you prefer?

- If you have truly given your 90% CI, you will have no preference.
- If you prefer option A, your initial estimate was probably overconfident
- If you prefer option B, your initial estimate was probably underconfident
- Review your ranges if they do not truly reflect your 90% CI

Appendix
Calibration Test

#	Question	Answer
1	In 1938, a British steam locomotive set a new speed record by going how fast (mph)?	126
2	In what year did Sir Isaac Newton publish the *Universal Laws of Gravitation*?	1685
3	How many mm long is a typical business card?	85
4	The Internet (then called 'ARPANET') was established as a military communications system in what year?	1969
5	In what year was William Shakespeare born?	1564
6	What is the air distance between New York and Los Angeles (km)?	3,944
7	What percentage of a square could be covered by a circle of the same width?	78.5%
8	How old was Charlie Chaplain when he died?	88
9	What is the weight, in lbs, of the first edition of *How to Measure Anything*?	0.56
10	The TV show *Big Brother* first aired on what date?	2000

Appendix
Calibration Test

#	Question	True/ False
1	The ancient Romans were conquered by the ancient Greeks	True
2	There is no species of three-humped camels	True
3	A gallon of oil weighs less than a gallon of water	True
4	Mars is always further away from Earth than Venus	False
5	Germany won the first World Cup	False
6	Napoleon was born on the island of Corsica	True
7	‘M’ is one of the three most commonly-used letters in English	False
8	In 2002, the price of the average new desktop computer purchased was under $1,500	True
9	Lyndon B. Johnson was a governor before becoming vice president	False
10	A kilogram is more than a pound	True

Appendix
Calibration Test

- How did you do?

- As you were giving 90% CIs for the first set of questions, you would expect to find 9 out of 10 answers falling within your ranges (if you were properly calibrated)

- For the second set of questions, convert each of your circled confidence percentages into a decimal (e.g., 60% = 0.6) and add them up. The result is how many answers you would expect to have gotten correct (again, if you were properly calibrated)

- You can find a variety of practical techniques for improving your calibration in the *How to Measure Anything…* books