Intertopic Distance Map (via multidimensional scaling)

PC2

PC1

4

3

2

1

Marginal topic distribution

2%

5%

10%
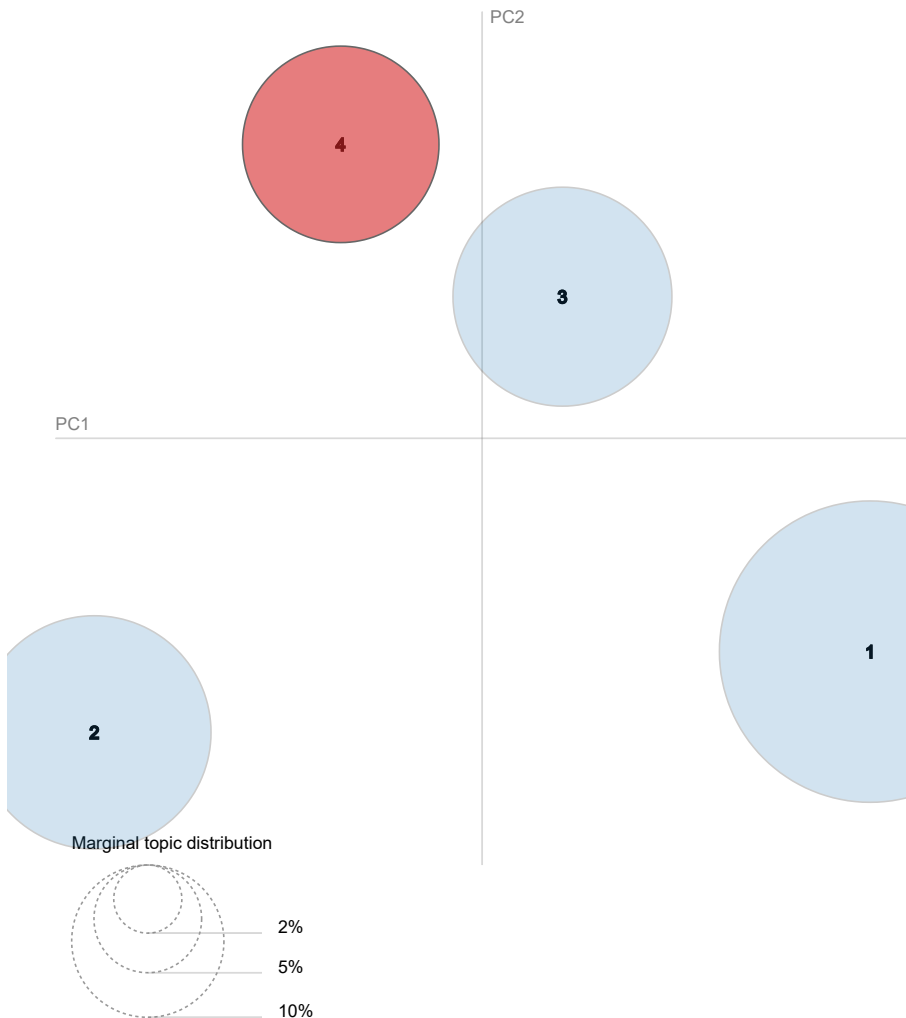
Top-30 Most Relevant Terms for Topic 4 (16.7% of tokens)

	0	20	40	60	80	100

datum
rule
threat
financial
intelligence
iodef
host
way
security
information
cti
business
traffic
opportunity
context
vault
otx
alien
clutter
prioritisation
use
level
customer
discovery
feature
range
device
source
client
analytic

Overall term frequency

Estimated term frequency within the selected topic
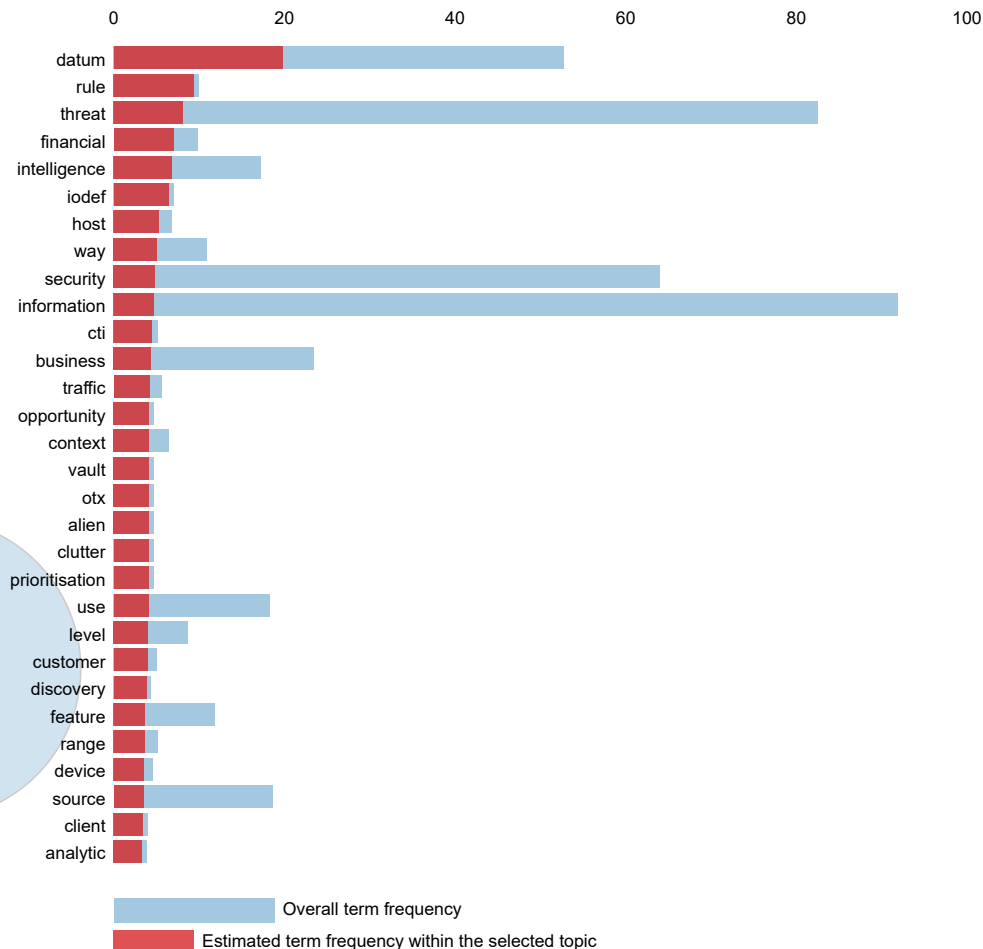
1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
2. relevance(term w | topic t) = λ * p(w | t) + (1 - λ) * p(w | t)/p(w); see Sievert & Shirley (2014)