1. Non-peer-reviewed corpus pre-covid (2010-2019): (100 in total)

[(0,

'0.043\*"customer" + 0.024\*"address" + 0.022\*"giant" + 0.019\*"trust" + 0.016\*"record" + 0.015\*"trend" + 0.013\*"virus" + 0.012\*"staff" + 0.012\*"release" + 0.008\*"member"'),

(1,

'0.023\*"link" + 0.018\*"video" + 0.016\*"source" + 0.015\*"camera" + 0.014\*"monitoring" + 0.013\*"chance" + 0.013\*"mail" + 0.012\*"hand" + 0.012\*"family" + 0.011\*"thirdparty"'),

(2,

'0.041\*"attack" + 0.039\*"use" + 0.025\*"control" + 0.020\*"device" + 0.015\*"evidence" + 0.014\*"issue" + 0.014\*"management" + 0.014\*"location" + 0.013\*"organization" + 0.012\*"personal\_information"'),

(3,

'0.046\*"business" + 0.032\*"malware" + 0.030\*"work" + 0.027\*"threat" + 0.026\*"industry" + 0.022\*"program" + 0.017\*"encryption" + 0.017\*"vulnerability" + 0.015\*"application" + 0.013\*"technology"'),

(4,

'0.051\*"time" + 0.031\*"system" + 0.030\*"number" + 0.026\*"technology" + 0.025\*"breach" + 0.024\*"email" + 0.021\*"world" + 0.020\*"account" + 0.016\*"secure" + 0.015\*"password"'),

(5,

'0.041\*"online" + 0.039\*"year" + 0.032\*"part" + 0.026\*"percent" + 0.025\*"week" + 0.025\*"report" + 0.022\*"state" + 0.019\*"group" + 0.019\*"day" + 0.018\*"credit"'),

(6,

'0.083\*"company" + 0.039\*"service" + 0.025\*"news" + 0.018\*"home" + 0.018\*"apps" + 0.017\*"phone" + 0.016\*"code" + 0.015\*"spam" + 0.014\*"traffic" + 0.014\*"process"'),

(7,

'0.074\*"software" + 0.065\*"network" + 0.053\*"computer" + 0.025\*"firm" + 0.018\*"hacker" + 0.017\*"tool" + 0.015\*"botnet" + 0.011\*"role" + 0.011\*"ransomware" + 0.011\*"approach"'),

(8,

'0.050\*"today" + 0.041\*"consumer" + 0.027\*"intelligence" + 0.025\*"story" + 0.018\*"director" + 0.017\*"product" + 0.016\*"share" + 0.015\*"security\_firm" + 0.014\*"series" + 0.013\*"tech"'),

(9,

'0.049\*"privacy" + 0.047\*"access" + 0.026\*"government" + 0.021\*"risk" + 0.020\*"help" + 0.017\*"way" + 0.015\*"problem" + 0.014\*"protection" + 0.013\*"survey" + 0.012\*"case"')]

1. Non-peer-reviewed corpus post-covid (2020-2021): (160 in total)

[(0,

'0.050\*"vulnerability" + 0.034\*"flaw" + 0.030\*"series" + 0.022\*"browser" + 0.022\*"application" + 0.020\*"analysis" + 0.020\*"coverage" + 0.016\*"storage" + 0.016\*"patch" + 0.015\*"engineering"'),

(1,

'0.051\*"credit" + 0.035\*"video" + 0.027\*"domain" + 0.026\*"cryptocurrency" + 0.022\*"program" + 0.021\*"researcher" + 0.021\*"conspiracy" + 0.020\*"demand" + 0.019\*"bureau" + 0.019\*"policy"'),

(2,

'0.064\*"day" + 0.030\*"page" + 0.028\*"survey" + 0.025\*"agency" + 0.025\*"investigation" + 0.024\*"article" + 0.022\*"case" + 0.019\*"bitcoin" + 0.018\*"concept" + 0.012\*"month"'),

(3,

'0.042\*"attack" + 0.031\*"traffic" + 0.019\*"list" + 0.017\*"bank" + 0.017\*"place" + 0.015\*"botnet" + 0.012\*"web" + 0.011\*"default" + 0.010\*"internet" + 0.010\*"mortgage"'),

(4,

'0.066\*"firm" + 0.051\*"fraud" + 0.019\*"platform" + 0.015\*"supply" + 0.015\*"part" + 0.015\*"review" + 0.014\*"advantage" + 0.013\*"solution" + 0.011\*"generation" + 0.010\*"morning"'),

(5,

'0.051\*"company" + 0.050\*"software" + 0.041\*"access" + 0.029\*"business" + 0.029\*"service" + 0.028\*"number" + 0.027\*"time" + 0.022\*"network" + 0.022\*"malware" + 0.022\*"group"'),

(6,

'0.055\*"week" + 0.036\*"news" + 0.032\*"computer" + 0.029\*"insurance" + 0.022\*"location" + 0.022\*"hacker" + 0.021\*"home" + 0.018\*"data" + 0.016\*"anyone" + 0.016\*"management"'),

(7,

'0.050\*"help" + 0.038\*"cloud" + 0.036\*"provider" + 0.027\*"vendor" + 0.024\*"file" + 0.023\*"trust" + 0.023\*"network" + 0.023\*"access" + 0.022\*"incident" + 0.022\*"authentication"'),

(8,

'0.052\*"story" + 0.042\*"disclosure" + 0.033\*"name" + 0.033\*"unemployment" + 0.025\*"marketing" + 0.025\*"law\_enforcement" + 0.023\*"country" + 0.021\*"state" + 0.017\*"consumer" + 0.014\*"election"'),

(9,

'0.039\*"breach" + 0.027\*"claim" + 0.023\*"thirdparty" + 0.021\*"money" + 0.017\*"store" + 0.017\*"range" + 0.016\*"cash" + 0.016\*"conference" + 0.013\*"care" + 0.011\*"year"'),

(10,

'0.047\*"online" + 0.043\*"code" + 0.040\*"account" + 0.022\*"source" + 0.020\*"address" + 0.019\*"spam" + 0.019\*"protection" + 0.017\*"password" + 0.015\*"card" + 0.015\*"hack"'),

(11,

'0.051\*"ransomware" + 0.038\*"customer" + 0.030\*"site" + 0.028\*"threat" + 0.023\*"attack" + 0.020\*"someone" + 0.020\*"companys" + 0.017\*"interview" + 0.016\*"today" + 0.015\*"share"'),

(12,

'0.039\*"world" + 0.027\*"tool" + 0.025\*"use" + 0.019\*"encryption" + 0.017\*"government" + 0.015\*"secure" + 0.015\*"experience" + 0.015\*"attention" + 0.015\*"report" + 0.013\*"multifactor"'),

(13,

'0.044\*"way" + 0.027\*"phone" + 0.024\*"technology" + 0.022\*"advertising" + 0.021\*"system" + 0.020\*"state" + 0.018\*"connection" + 0.017\*"tech" + 0.017\*"personal\_information" + 0.015\*"framework"'),

(14,

'0.053\*"part" + 0.043\*"activity" + 0.039\*"organization" + 0.037\*"director" + 0.025\*"cybercrime" + 0.023\*"content" + 0.021\*"support" + 0.016\*"system" + 0.016\*"disinformation" + 0.014\*"board"'),

(15,

'0.080\*"today" + 0.066\*"privacy" + 0.041\*"year" + 0.036\*"identity" + 0.033\*"message" + 0.033\*"response" + 0.025\*"crime" + 0.022\*"use" + 0.020\*"website" + 0.016\*"issue"')]

1. Peer-reviewed corpus keywords (2010-2021) keywords: (110 in total):

[(0,

'0.070\*"health" + 0.062\*"healthcare" + 0.033\*"ransomware" + 0.019\*"fraud" + 0.017\*"consumer" + 0.015\*"email" + 0.012\*"questionnaire" + 0.012\*"consensus" + 0.012\*"device" + 0.011\*"solve"'),

(1,

'0.063\*"policy" + 0.040\*"government" + 0.038\*"cyberspace" + 0.036\*"space" + 0.033\*"governance" + 0.027\*"law" + 0.022\*"economy" + 0.022\*"chapter" + 0.020\*"state" + 0.020\*"insider"'),

(2,

'0.048\*"blockchain" + 0.039\*"protocol" + 0.039\*"scheme" + 0.033\*"encryption" + 0.022\*"insurance" + 0.020\*"memory" + 0.018\*"message" + 0.017\*"certification" + 0.017\*"mission" + 0.016\*"controller"'),

(3,

'0.095\*"system" + 0.043\*"control" + 0.035\*"software" + 0.034\*"process" + 0.032\*"communication" + 0.029\*"design" + 0.024\*"safety" + 0.023\*"infrastructure" + 0.022\*"protection" + 0.021\*"access"'),

(4,

'0.153\*"privacy" + 0.036\*"breach" + 0.032\*"investment" + 0.018\*"surveillance" + 0.016\*"feature\_selection" + 0.014\*"harm" + 0.013\*"disclosure" + 0.013\*"trade" + 0.010\*"firm" + 0.009\*"mass"'),

(5,

'0.027\*"technology" + 0.022\*"development" + 0.020\*"analysis" + 0.020\*"use" + 0.017\*"framework" + 0.017\*"work" + 0.017\*"approach" + 0.013\*"level" + 0.012\*"industry" + 0.012\*"knowledge"'),

(6,

'0.070\*"network" + 0.060\*"detection" + 0.055\*"model" + 0.039\*"attack" + 0.027\*"performance" + 0.025\*"method" + 0.023\*"machine" + 0.021\*"approach" + 0.016\*"traffic" + 0.015\*"intrusion"'),

(7,

'0.052\*"malware" + 0.028\*"cloud" + 0.027\*"service" + 0.018\*"evidence" + 0.017\*"event" + 0.017\*"time" + 0.016\*"group" + 0.014\*"framework" + 0.014\*"series" + 0.014\*"contribution"'),

(8,

'0.104\*"power" + 0.050\*"energy" + 0.032\*"grid" + 0.028\*"distribution" + 0.018\*"storage" + 0.017\*"production" + 0.016\*"manufacturing" + 0.016\*"transmission" + 0.015\*"transformation" + 0.015\*"operation"'),

(9,

'0.074\*"intelligence" + 0.042\*"cybercrime" + 0.035\*"threat" + 0.034\*"home" + 0.025\*"cooperation" + 0.024\*"product" + 0.024\*"market" + 0.015\*"crime" + 0.015\*"crisis" + 0.012\*"mobility"'),

(10,

'0.130\*"risk" + 0.103\*"management" + 0.042\*"business" + 0.041\*"assessment" + 0.027\*"authentication" + 0.025\*"resilience" + 0.019\*"trust" + 0.015\*"compliance" + 0.014\*"organization" + 0.013\*"ecosystem"')]

1. Peer-reviewed corpus pre-covid keywords: (100 in total)

[(0,

'0.039\*"technology" + 0.032\*"management" + 0.022\*"industry" + 0.017\*"level" + 0.016\*"development" + 0.016\*"business" + 0.016\*"infrastructure" + 0.015\*"world" + 0.012\*"space" + 0.011\*"protection"'),

(1,

'0.061\*"privacy" + 0.049\*"policy" + 0.032\*"cyberspace" + 0.017\*"cybercrime" + 0.017\*"governance" + 0.016\*"internet" + 0.014\*"law" + 0.013\*"consumer" + 0.011\*"legislation" + 0.010\*"relationship"'),

(2,

'0.116\*"detection" + 0.046\*"classification" + 0.045\*"malware" + 0.023\*"accuracy" + 0.018\*"feature" + 0.017\*"detect" + 0.014\*"anomaly\_detection" + 0.014\*"module" + 0.013\*"memory" + 0.012\*"computation"'),

(3,

'0.064\*"system" + 0.056\*"control" + 0.038\*"power" + 0.036\*"communication" + 0.026\*"access" + 0.018\*"secure" + 0.018\*"energy" + 0.017\*"architecture" + 0.016\*"operation" + 0.014\*"authentication"'),

(4,

'0.103\*"network" + 0.062\*"traffic" + 0.028\*"technique" + 0.025\*"dataset" + 0.023\*"intrusion\_detection" + 0.023\*"graph" + 0.017\*"ransomware" + 0.017\*"intrusion" + 0.014\*"accuracy" + 0.012\*"password"'),

(5,

'0.033\*"game" + 0.025\*"influence" + 0.016\*"prediction" + 0.015\*"deterrence" + 0.014\*"theory" + 0.014\*"exercise" + 0.013\*"deception" + 0.013\*"uncertainty" + 0.012\*"file" + 0.012\*"aviation"'),

(6,

'0.046\*"blockchain" + 0.028\*"insurance" + 0.022\*"learning" + 0.014\*"workshop" + 0.012\*"cryptocurrency" + 0.011\*"employment" + 0.010\*"shortage" + 0.010\*"literacy" + 0.010\*"index" + 0.009\*"spectrum"'),

(7,

'0.041\*"system" + 0.038\*"analysis" + 0.037\*"approach" + 0.035\*"model" + 0.026\*"framework" + 0.023\*"attack" + 0.023\*"process" + 0.022\*"risk" + 0.021\*"network" + 0.016\*"problem"'),

(8,

'0.031\*"design" + 0.031\*"development" + 0.027\*"knowledge" + 0.023\*"field" + 0.023\*"engineering" + 0.021\*"computer" + 0.020\*"education" + 0.017\*"training" + 0.017\*"project" + 0.017\*"program"'),

(9,

'0.107\*"software" + 0.037\*"vulnerability" + 0.029\*"defense" + 0.016\*"test" + 0.016\*"code" + 0.015\*"damage" + 0.013\*"cyberattack" + 0.011\*"functionality" + 0.011\*"malware" + 0.010\*"execution"')]

1. Peer-reviewed corpus post-covid: (150 in total)

[(0,

'0.079\*"system" + 0.039\*"communication" + 0.026\*"infrastructure" + 0.025\*"access" + 0.023\*"architecture" + 0.021\*"secure" + 0.020\*"platform" + 0.016\*"framework" + 0.016\*"environment" + 0.016\*"protection"'),

(1,

'0.067\*"privacy" + 0.051\*"health" + 0.044\*"healthcare" + 0.016\*"science" + 0.014\*"crisis" + 0.014\*"care" + 0.014\*"innovation" + 0.014\*"cryptography" + 0.013\*"period" + 0.013\*"staff"'),

(2,

'0.035\*"analysis" + 0.025\*"use" + 0.025\*"development" + 0.023\*"technology" + 0.021\*"work" + 0.020\*"framework" + 0.020\*"process" + 0.020\*"software" + 0.019\*"approach" + 0.017\*"design"'),

(3,

'0.034\*"intelligence" + 0.033\*"policy" + 0.029\*"ransomware" + 0.024\*"regulation" + 0.024\*"compliance" + 0.022\*"student" + 0.017\*"assess" + 0.016\*"idea" + 0.015\*"series" + 0.015\*"threat\_intelligence"'),

(4,

'0.058\*"awareness" + 0.054\*"program" + 0.040\*"course" + 0.030\*"web" + 0.020\*"computer\_science" + 0.016\*"exposure" + 0.015\*"online" + 0.014\*"university" + 0.014\*"coverage" + 0.013\*"book"'),

(5,

'0.043\*"event" + 0.035\*"breach" + 0.032\*"code" + 0.029\*"vulnerability" + 0.017\*"motivation" + 0.016\*"discovery" + 0.016\*"prototype" + 0.014\*"list" + 0.014\*"quantum" + 0.013\*"trade"'),

(6,

'0.085\*"system" + 0.067\*"control" + 0.066\*"power" + 0.049\*"energy" + 0.033\*"operation" + 0.029\*"resilience" + 0.029\*"distribution" + 0.019\*"reliability" + 0.014\*"generation" + 0.013\*"load"'),

(7,

'0.026\*"encryption" + 0.023\*"behaviour" + 0.020\*"key" + 0.019\*"home" + 0.015\*"wide\_range" + 0.015\*"video" + 0.015\*"port" + 0.014\*"questionnaire" + 0.013\*"big\_data" + 0.012\*"acceptance"'),

(8,

'0.097\*"malware" + 0.027\*"image" + 0.022\*"combination" + 0.020\*"company" + 0.020\*"language" + 0.019\*"certification" + 0.019\*"solution" + 0.017\*"team" + 0.016\*"visualization" + 0.016\*"smart\_city"'),

(9,

'0.073\*"risk" + 0.059\*"management" + 0.036\*"safety" + 0.031\*"business" + 0.027\*"solution" + 0.017\*"value" + 0.017\*"organization" + 0.016\*"adoption" + 0.016\*"technology" + 0.015\*"influence"'),

(10,

'0.063\*"network" + 0.063\*"model" + 0.052\*"detection" + 0.042\*"attack" + 0.032\*"performance" + 0.031\*"method" + 0.025\*"approach" + 0.021\*"problem" + 0.019\*"traffic" + 0.018\*"threat"'),

(11,

'0.047\*"loss" + 0.035\*"law" + 0.032\*"class" + 0.029\*"collaboration" + 0.023\*"insurance" + 0.022\*"cooperation" + 0.021\*"speed" + 0.018\*"reduction" + 0.017\*"social\_media" + 0.017\*"banking"'),

(12,

'0.027\*"investment" + 0.023\*"wireless" + 0.019\*"channel" + 0.019\*"disinformation" + 0.014\*"transaction" + 0.014\*"phenomenon" + 0.012\*"aviation" + 0.011\*"author" + 0.010\*"radiation" + 0.010\*"aircraft"'),

(13,

'0.087\*"game" + 0.020\*"feedback" + 0.018\*"detector" + 0.015\*"log" + 0.013\*"honeypot" + 0.012\*"attacker" + 0.012\*"gamification" + 0.011\*"stress" + 0.009\*"variance" + 0.008\*"zeroday"'),

(14,

'0.036\*"construction" + 0.036\*"optimization" + 0.029\*"point" + 0.021\*"text" + 0.020\*"assurance" + 0.020\*"employee" + 0.019\*"chapter" + 0.018\*"water" + 0.017\*"baseline" + 0.017\*"insider"')]