# Data is Everywhere



Jay Jacobs jay@cyentia.com

# INSTITUTE



# Whatcha Been Doing?

- (Mostly) Full-time with Cyentia Institute
- Conducting sponsored research
- Building Cyentia Library

CYENTAA INSTITUTE	About 🖛 Library 🕶 Podcast Blog Contact 🛅 🎔 🔍
Cybersecurity Resear At Cyentia Institute, we understand that reliable research is essential for	a stronger, more informed professional
community. We leverage a unique process to select and curate high qua	ality research for community use.
Q Search the Library	
All Years Y All Tags Y <u>Clear all</u>	

### **CYENT**A

### **VOICE OF THE** ANALYST STUDY

An Inside Perspective on Security Operations



Commissioned by respond



THE 2017

REPORT

Sponsored by POINT

NSTITUTE

### STRIKING SECURITY GOLD

As the premier security conference in the world, RSA Cov

nrough which to study the topics and trends within our incusity. The Conference's slogan of Where the World Talks Security' shows that's not just an accident; it's the goal.

this report, which has its roots in a smilar question asked by an eight-year-old day an app "What's the RSA Conference about Dadd/?" That in or of the 25th anniversary of RSA Confer

s Impan, information, Call for Paper (CPP) de godmine of details and insight about the sessions just waiting to m re was kind enough to supply the one for our digital luable nuggets of insight abv

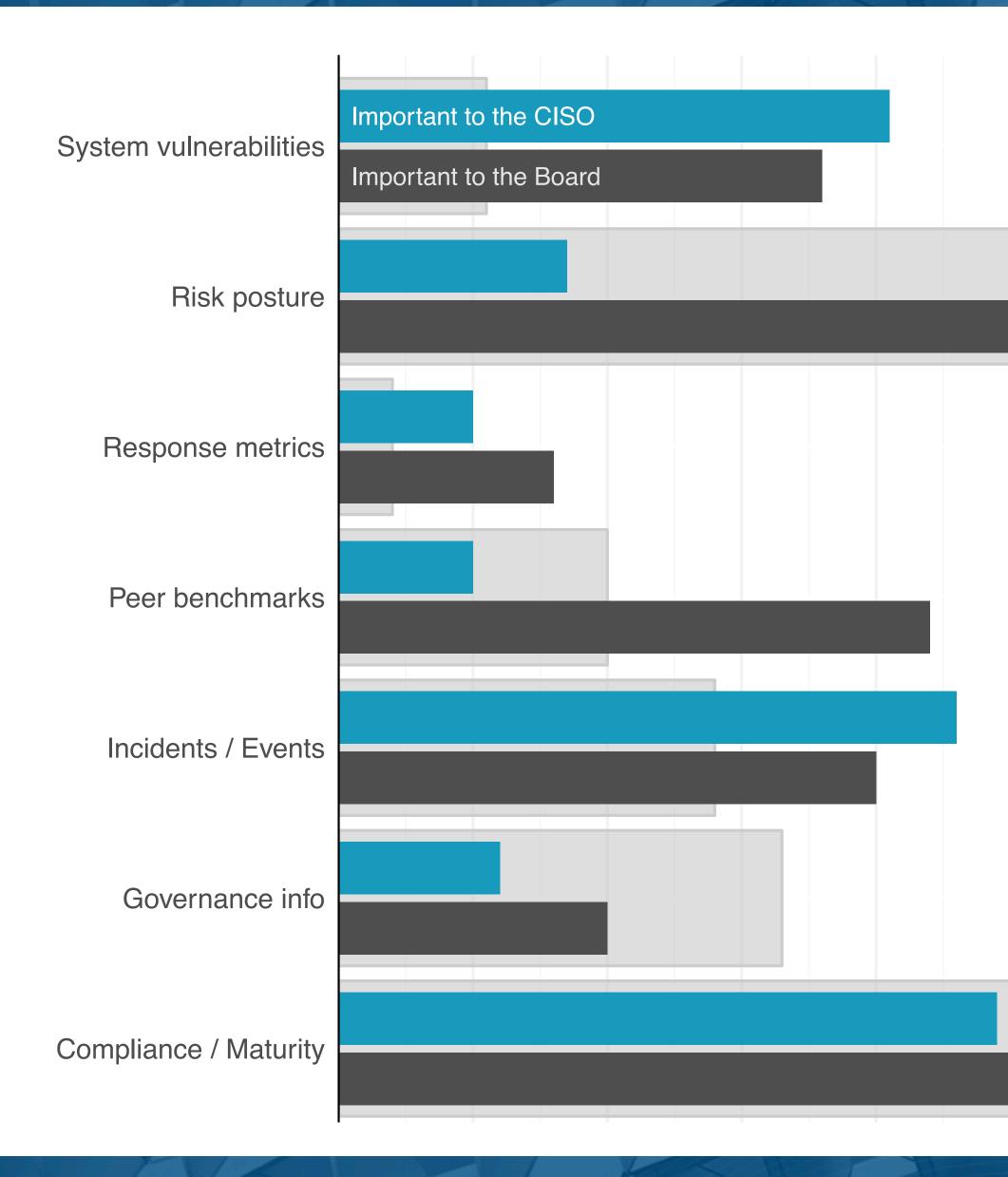
RSA Conference Where the world







# Cyber Balance Sheet 2017



INSTITUTE

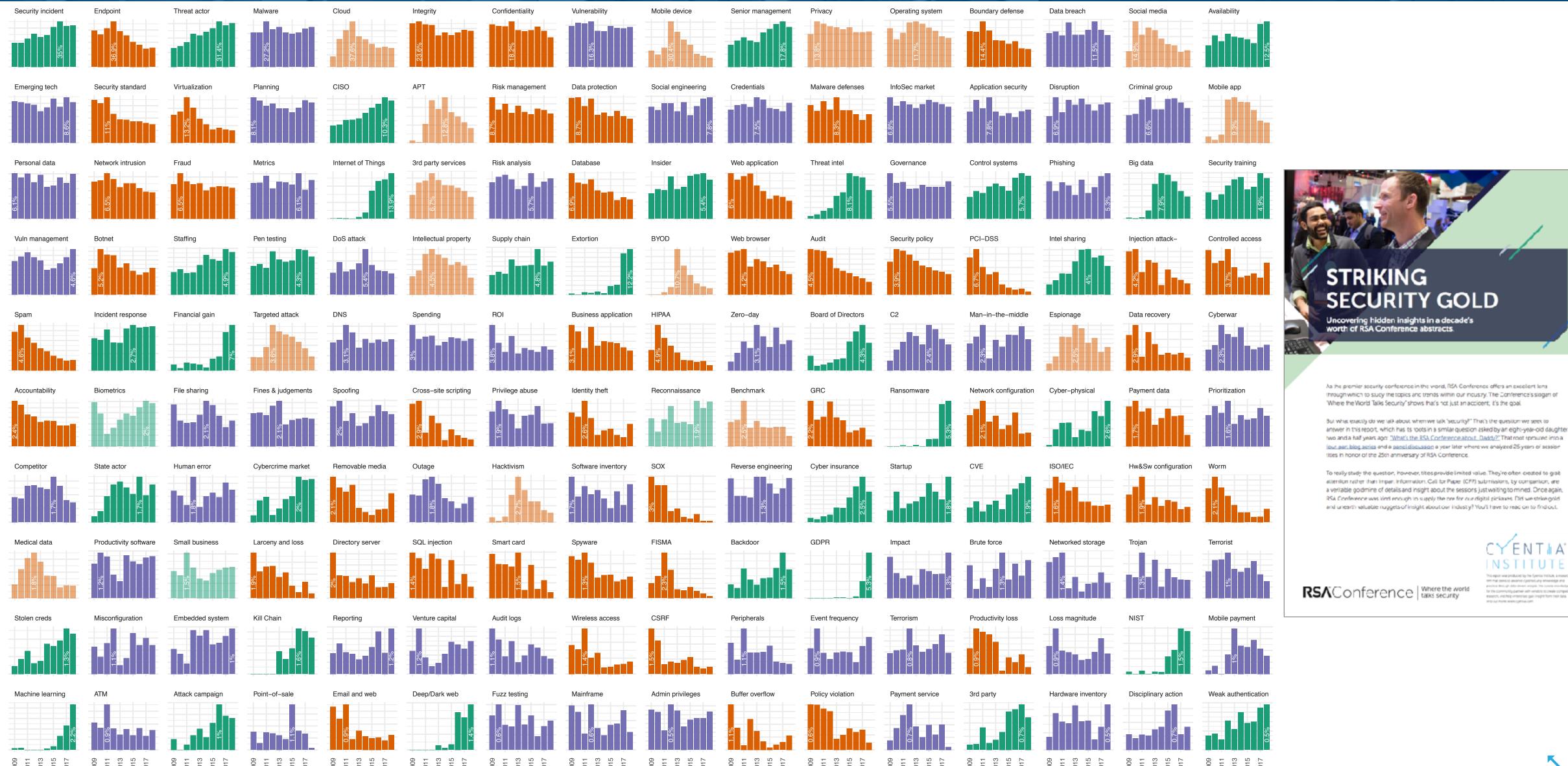
What the Board gets







# **RSAC: Topics and Trends**



### FN NSTIT

Source: Cventia Institute with data from RSA Conference







# **RSAC: Topics and Trends**



### ΕN INSTITUTE

Source: Cventia Institute with data from RSA Conference







# RSAC: Landscape of Topics

								Forced bro
							Cross-site scripting	
						Input handling	injection njection attack–	
				Application see	çurity	Buffer overflow		ClickJacking
		NISD Metrics Account		k configuration Vuln managemen	Web applicatio Fuzz testing CVE Utuability Pen testing Pass-the-ha	erability		
	COBIT Governant ISO/IEC	Hardwo NIST CIS FISMA Reporting Spenotics Risk management Impact ce Board of Directors nior management Former employee Ver	Benchmark Control	trength strength cryptanalys htrolled access Virtualization Internet of Things St Startup Supply chain Dis dependent of Wireless	Embedded system Sis Database Set sniffer DHCP Control systems Control systems Enviced Peripherals Cyberwar December 2000 Peripherals Cyberwar December 2000 Bacor	t Threat capability Threat capability Threat capability Web browser Weak arthenticatio		
HITRUST	Security standard	Accountability	ard party Disciplinary action Olicy violation Class Consumer tech D	sified data ata recovery big data Event frequency File sharing Small bi sine Competitor	Terrorist Sm Die app Network control M Hacktivism Errolla	Threat actor Software inventory Corportunistic attack Misconfiguration	Pharming sruption Vents attack vatening hole	Web defaceme
	GLB HIPAA		A judgements A	Endpoint Biometrics Natural hazard levice Payment service Mobile pa	ial media APT ACTORS t ATM Deep/Dark web	State actor Malward Argeted attack Spoo State Espionage Malward Malwa	e defenses Botnet <sup>ofing</sup> Rootkit Backdoor Spam Stolen creds Adware	[
ERPA			Dataloss amount Data breach			Skimmers scraper Antiforensics Fraud	Attack campaign Social e Spyware Phishing	ngineering
			Data Diedel		Bank Identity theft	data	Fina Password dumper	ncial gain Estortion
			Copyrighted data					Rans



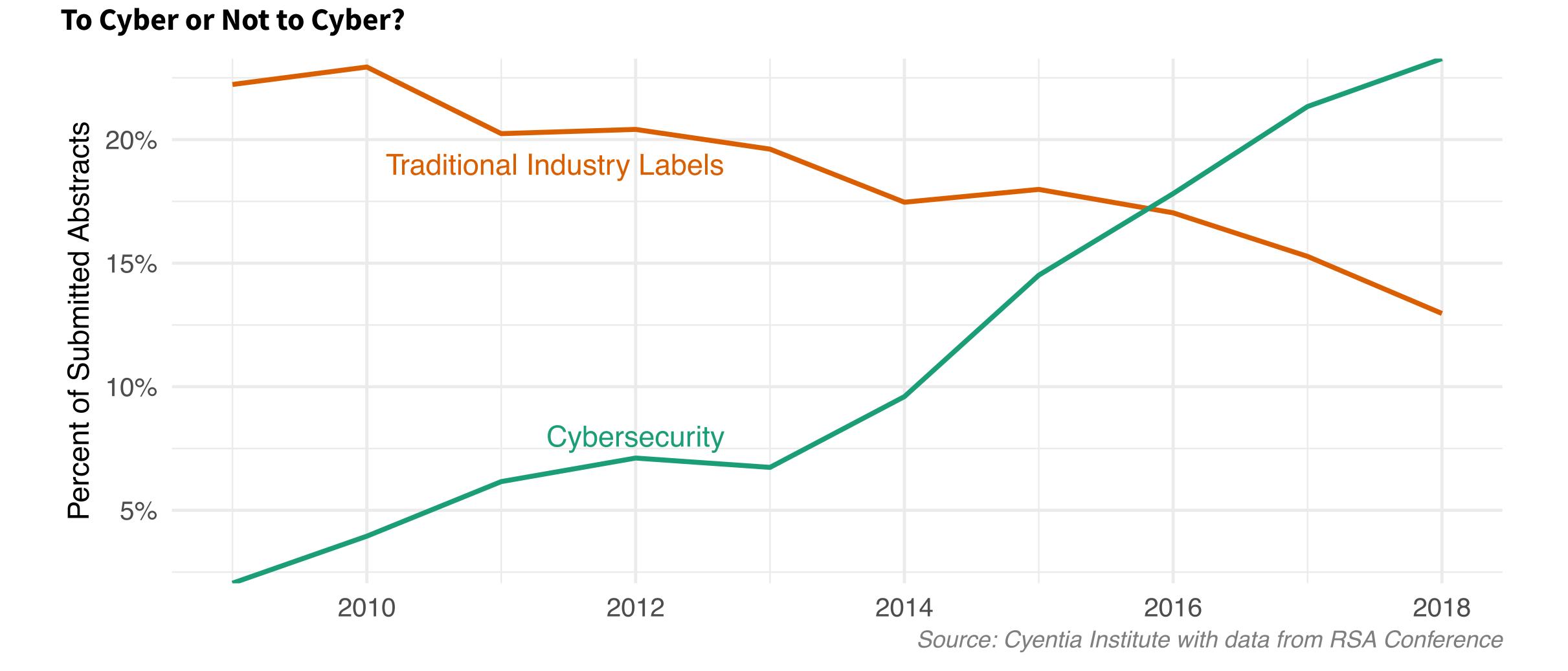


nt

)ownloader

omware

## RSA Conference





 $\checkmark$ 

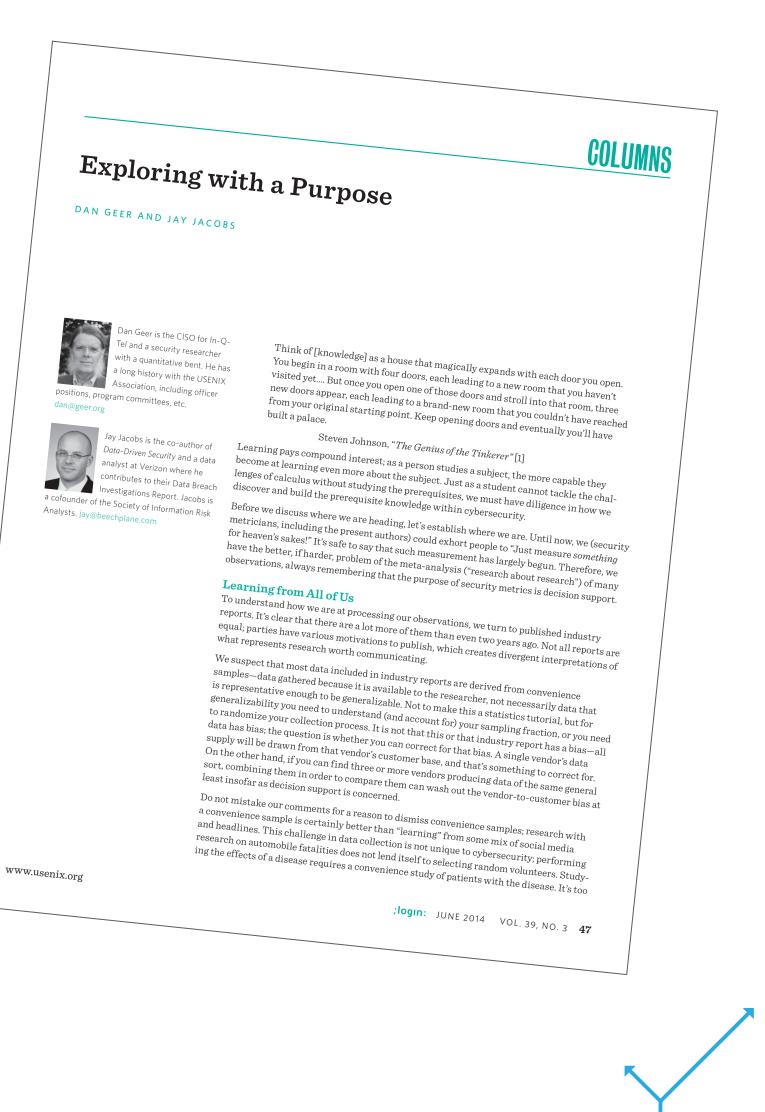
# Where we are headed

"What we (the security metrics people) must now do is learn how to do meta-analysis in our domain..."

- 1. Meta-Analysis and standing on the shoulders of giants: Cochrane Library
- 2. Case study: Ransomware
- 3. The Cyentia Library: present and future

NSTI

- Geer, Jacobs, 2014





# Cochrane Library

Cochrane Library

### What is a systematic review?

A systematic review attempts to identify, appraise and synthesize all the empirical evidence that meets pre-specified eligibility criteria to answer a given research question. Researchers conducting systematic reviews use explicit methods aimed at minimizing bias, in order to produce more reliable findings that can be used to inform decision making. (See Section 1.2 in the Cochrane Handbook for Systematic Reviews of Interventions.)



### Trusted evidence. Informed decisions. Better health.

http://www.cochranelibrary.com/

# Cochrane Library

Cochrane Library

### What is a systematic review?

### A systematic review attempts to identify, appraise and synthesize all the empirical evidence that meets pre-specified eligibility criteria to answer a given research question. Use more reliable

findings that can be used to inform decision making. (See Section 1.2 in the Cochrane Handbook for Systematic Reviews of Interventions.)



### Trusted evidence. Informed decisions. Better health.

http://www.cochranelibrary.com/

# Systemic Reviews

A systematic review attempts to identify, appraise and synthesize all the empirical evidence that meets pre-specified eligibility criteria to answer a given research question.

Given a Research Question:

- Identify sources of evidence and information
- Appraise the quality of the evidence
- Synthesize and aggregate the evidence together (meta-analysis)

➡ Identify Sources

### Research Question



Appraise Quality

Synthesize Evidence

# **Developing Research Questions**

A great research question:

- ... is interesting
- ... frames the object of measurement

**Poor Research Questions** 

"How Secure is this web app?"

"What risks do we face?

# • ...can be supported by observation/evidence

**Better Research Questions** 

"What is the probability this web app will have a vulnerability exploited in the next 12 months?"

"What is the probability of these events occurring this year?"

Breakdown broad topics into a series of research questions

Research Question — Identify Sources — Appraise Quality — Synthesize Evidence



NSTIT

# Identify sources

### https://www.cyentia.com/library/

# CYENT A

### Cybersecurity Research Library

At Cyentia Institute, we understand that reliable research is essential for a stronger, more informed professional community. We leverage a unique process to select and curate high quality research for community use.

VIEW LIBRARY CLASSIFICATION SYSTEM



Research Question \_\_\_\_\_

Identify Sources

### INSTITUTE

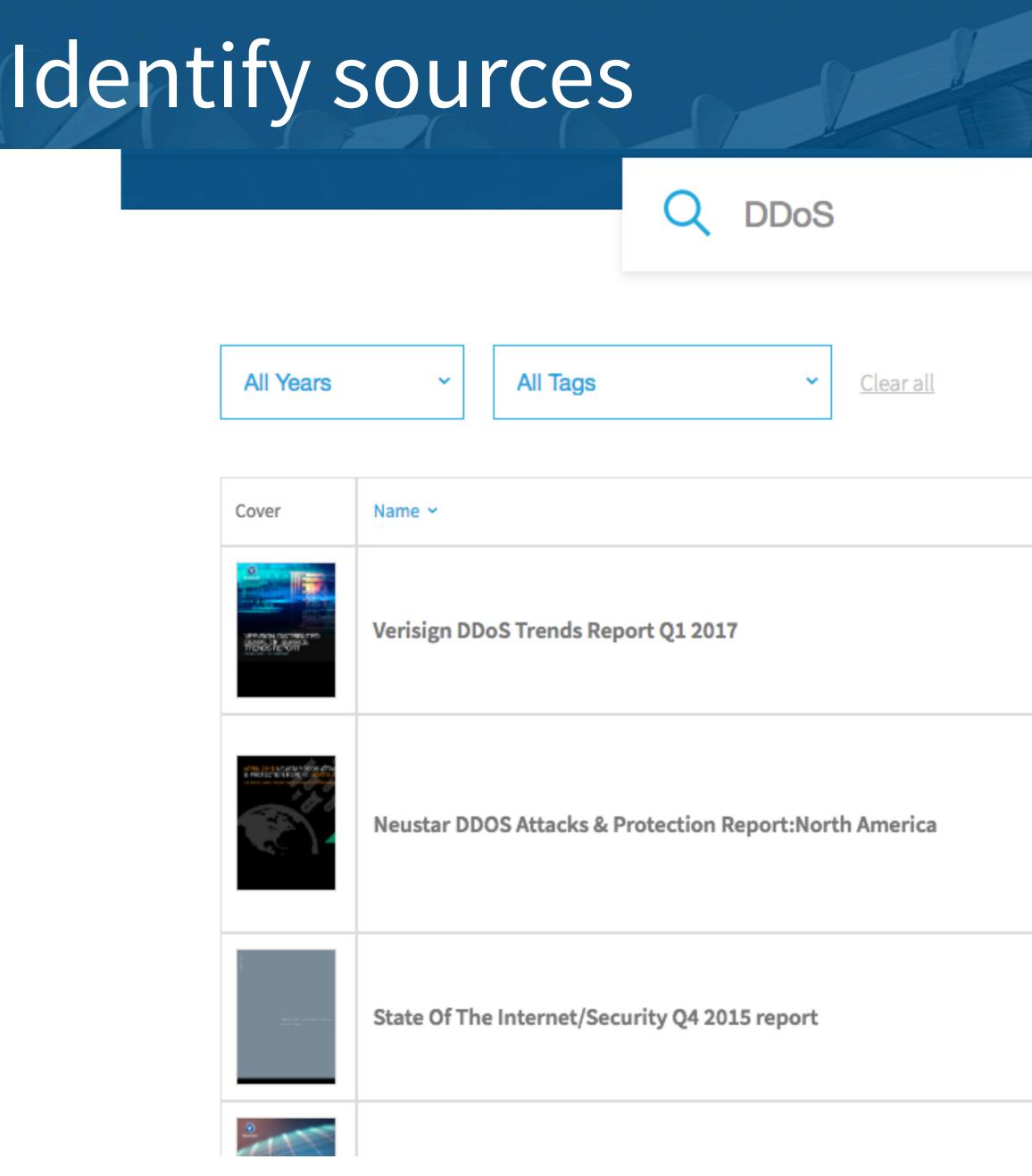
Q

About 🔻 Podcast

Appraise Quality –

Synthesize Evidence





**Research Question** 

Identify Sources

### CYENT I A INSTITUTE



Year ~	Туре ~	Торіс	Subtopic
2017	Industry report	Information Assets, Security attributes, Threats	Actors and motives, Availability, Events and TTPs
2015	Industry report	Controls, GRC Management, Impact and Loss, Market trends, Security attributes, Threats	Actors and motives, Availability, CIS "Top20" Controls
2015	Industry report	Controls, Information Assets, Security attributes, Threats	Actors and motives, Availability, Data
		Controls, Information	Actors and motives,



Synthesize Evidence



# Appraise the Quality

- "Quality" is study-specific (survey vs collected data), but always contains:
  - 1. Source of data, collection process (selection bias)
  - 2. Sample size, sub-sample slices (sampling error)
  - 3. Data Interpretation (e.g. statistics)
    - Appraising quality is subtle, complex and often subjective

Research Question — Identify Sources



Appraise Quality -

→ Synthesize Evidence



# Synthesize Evidence

of the size of the effect and/or to **resolve uncertainty** when reports disagree.

• Offset convenience samples

• Research in security is relatively simple: counts, proportions, means, etc.

Research Question ————— Identify Sources ————



# A meta-analysis uses a statistical approach to combine the results from multiple studies in an effort to increase power (over individual studies), improve estimates

https://en.wikipedia.org/wiki/Meta-analysis

Appraise Quality \_\_\_\_\_

Synthesize Evidence

# Meta-Analysis: Combining Proportions

Think about picking marbles from an urn:

- First person picked 19 out of 50 red
- Second person picked 32 out of 75 red
- total: 51 out of 125 were red

the same urn

### ...Assuming the studies are drawing from the same "urn" or are representative of

### Can visualize and talk about confidence in proportions with the **beta distribution**



# **Beta Distribution**

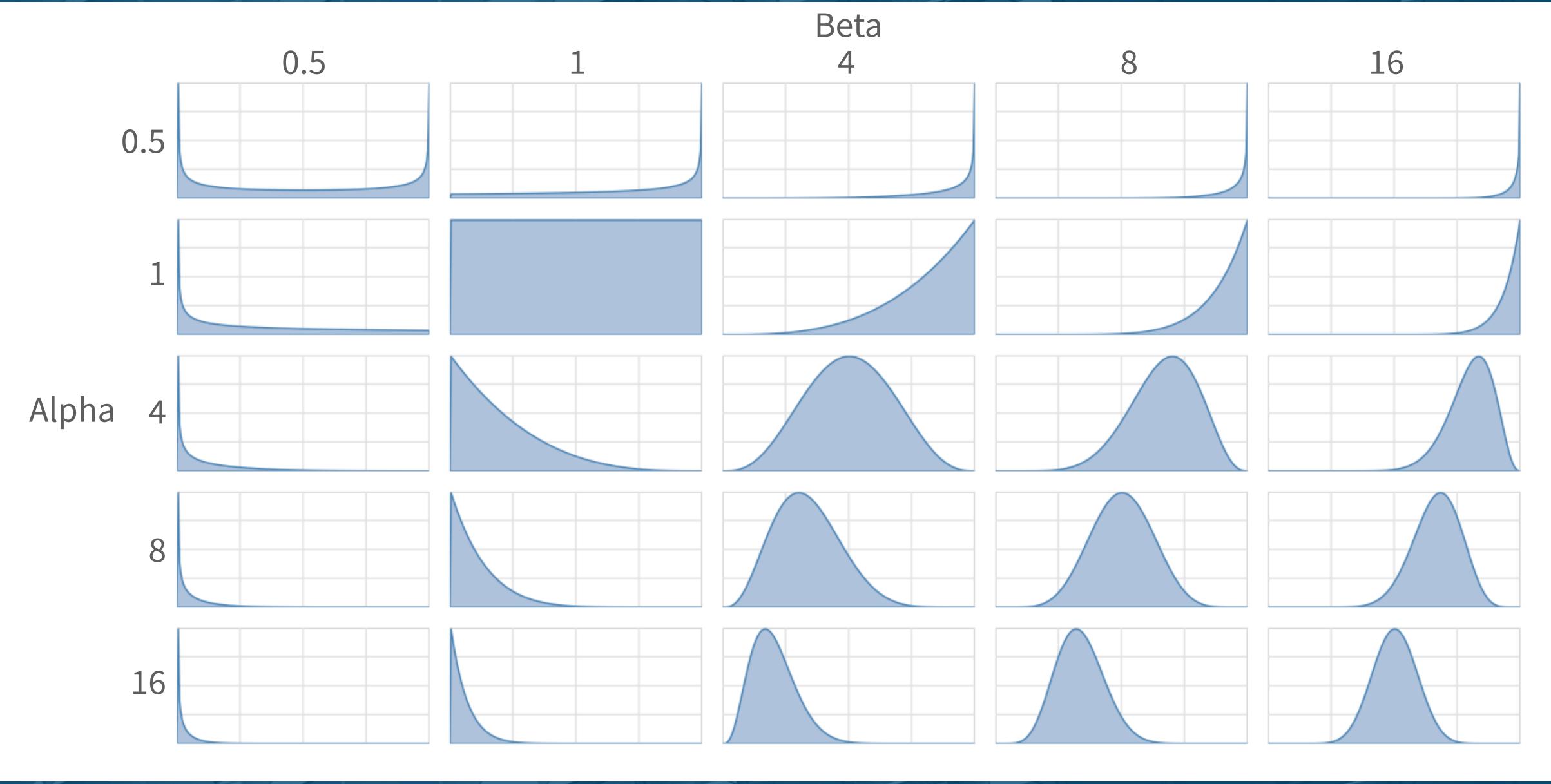
- "[The beta distribution] represents all the possible values of a probability when we don't know what that probability is." - David Robinson, stats.stackexchange.com
- Basis for betaPERT, conjugate prior for bayesian inference
- Has two parameters: alpha ( $\alpha$ ) and beta ( $\beta$ )
  - $\alpha$  are counts of class 1 (success/heads/red/breached/infected)
  - β are counts of class 2
- 50 out of 250 machines infected with malware:

beta( $\alpha$ =50,  $\beta$ =200)





# Visualizing the Beta

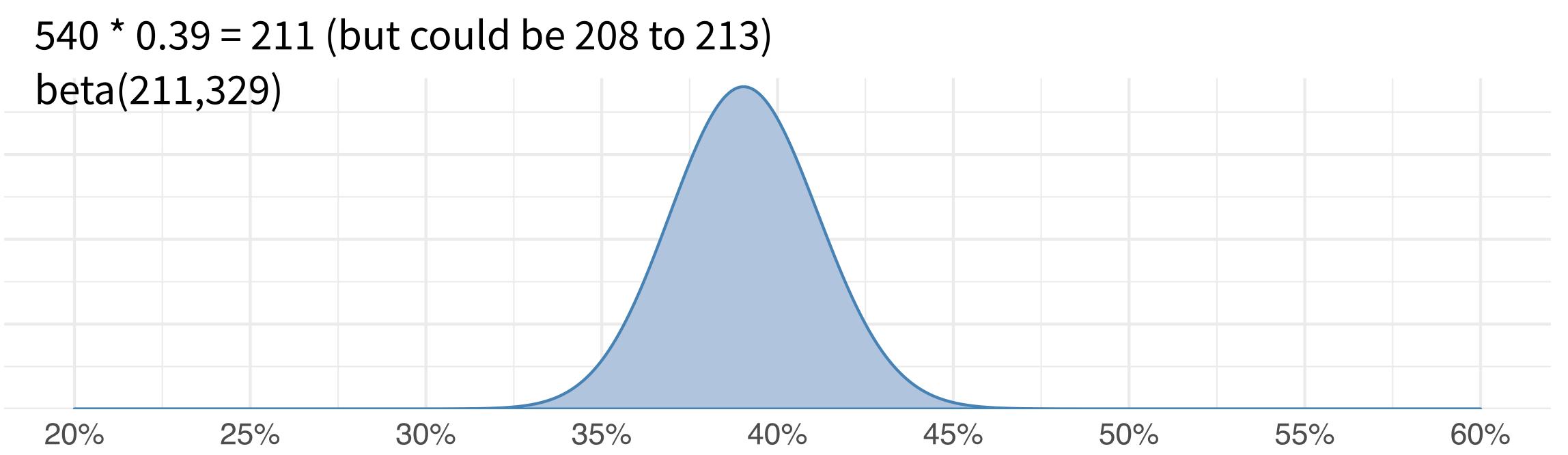


# INSTITUTE



# Applying the beta

- Osterman does a ransomware study and surveys 540 people
- Claims the "average ransomware penetration rate" is 39 percent
- How confident should we be about that 39%?











# CYENT<sub>I</sub>A® INSTITUTE





www.kaspersky.com

## Measuring Ransomware: The Setup

Three broad research questions

- How many orgs are affected by ransomware (prevalence)?
- How many orgs are paying the ransom amount (payment rate)?
- How much does ransomware cost (ransom amount)?

BSI, Ergebnisse der Umfrage zur Betroffenheit durch Ransomware (2016)

Fortinet, Q4 2016 Threat Landscape Report (2017)

IBM, Ransomware: How Consumers and Businesses Value Their Data (2016)

Kaspersky, Cost of Cryptomalware : SMBs at the Gunpoint (2016)

Osterman Research / Malwarebytes, Understanding the Depth of the Global Ransomware Problem (2016)

Ponemon Institute / Carbonite, The Rise of Ransomware (2017)

Symantec report (2012)

Dell Secureworks blog post (2013)

University of Kent study (2015)

BitDefender report (2016)

Datto report (2016)

Kaspersky - Consumer Security Risks (2016)

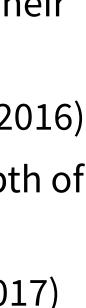
TrustLook blog post (2017)

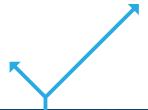
Cisco Annual Security Report (2016)

Cyber Extortion Risk Report, NYA International (2015)

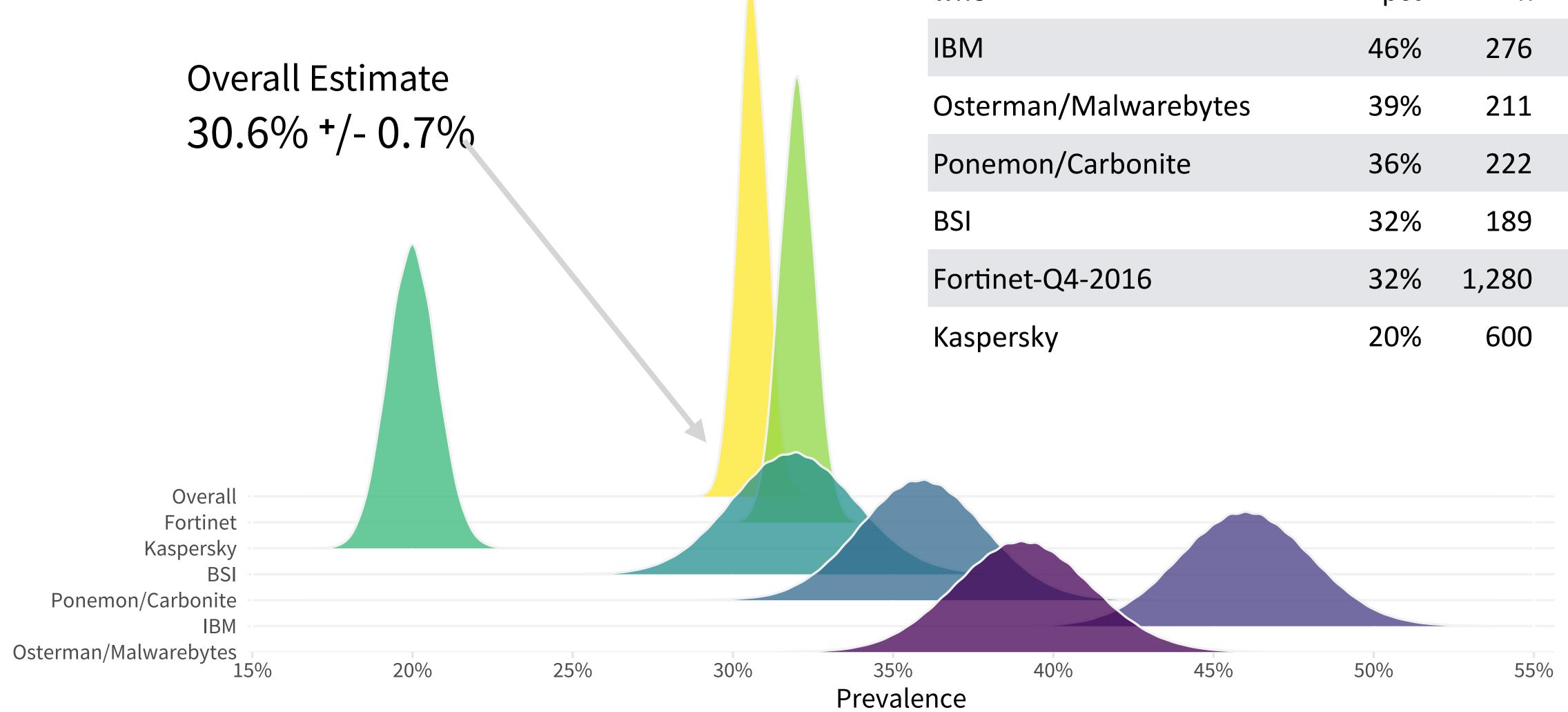


NSTIT





## Ransomware Prevalence



INSTITUTE

who	pct	X	
IBM	46%	276	
Osterman/Malwarebytes	39%	211	
Ponemon/Carbonite	36%	222	
BSI	32%	189	
Fortinet-Q4-2016	32%	1,280	4,
Kaspersky	20%	600	3,

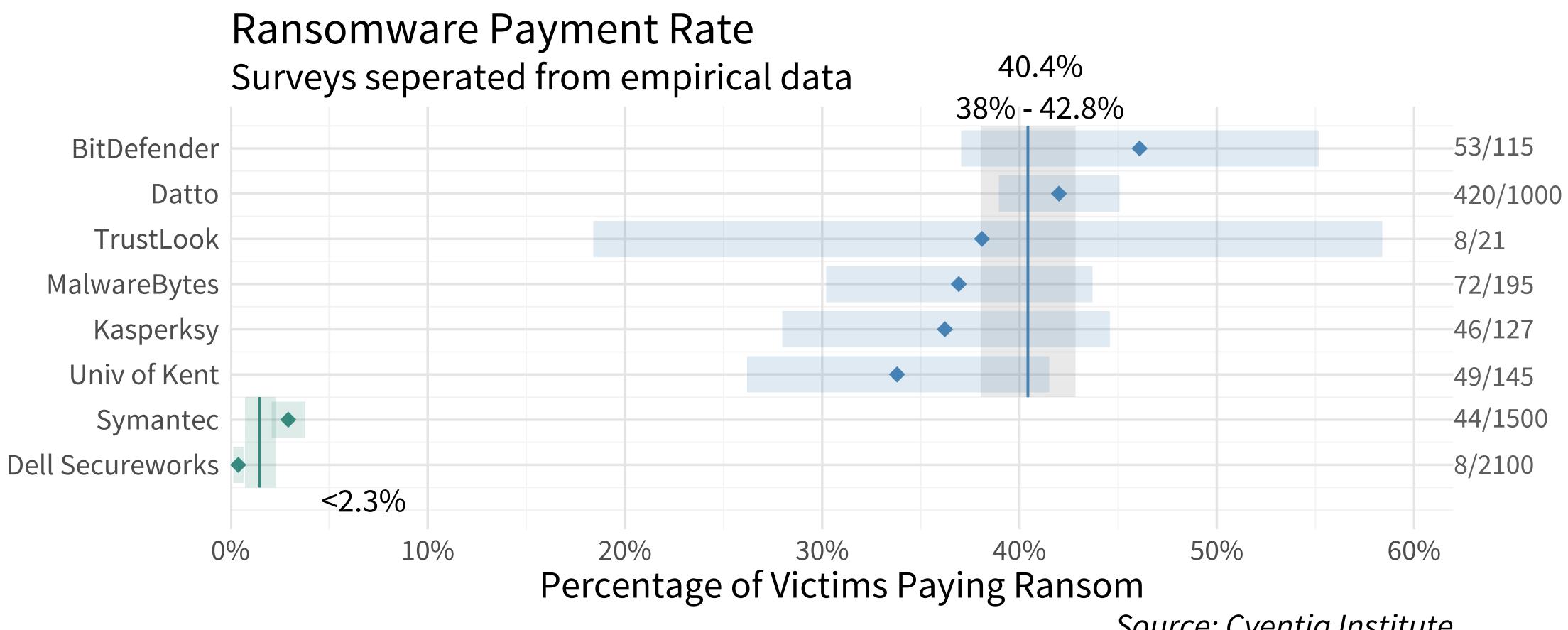
Source:CyentiaInstitute







# How many orgs are paying?



INSTITUTE

Source: Cyentia Institute



# CIGI Study



### Methodology

- This survey was conducted by Ipsos on behalf of the Centre for International Governance Innovation ("CIGI") between December 23, 2016, and March 21, 2017.
- The survey was conducted in 24 economies—Australia, Brazil, Canada, China, Egypt, France, Germany, Great Britain, Hong Kong (China), India, Indonesia, Italy, Japan, Kenya, Mexico, Nigeria, Pakistan, Poland, South Africa, South Korea, Sweden, Tunisia, Turkey and the United States-and involved 24,225 Internet users.
- Twenty of the countries utilized the Ipsos Internet panel system while Tunisia was conducted via CATI, and Kenya, Nigeria and Pakistan utilized face-to-face interviewing, given online constraints in these countries and the length
- In the US and Canada respondents were aged 18-64, and 16-64 in all other countries.
- Approximately 1000+ individuals were surveyed in each country and are weighted to match the population in each country surveyed. The precision of Ipsos online polls is calculated using a credibility interval. In this case, a poll of 1,000 is accurate to +/- 3.5 percentage points. For those surveys conducted by CATI and face-to-face, the margin of error is +/-3.1, 19 times out of 20.

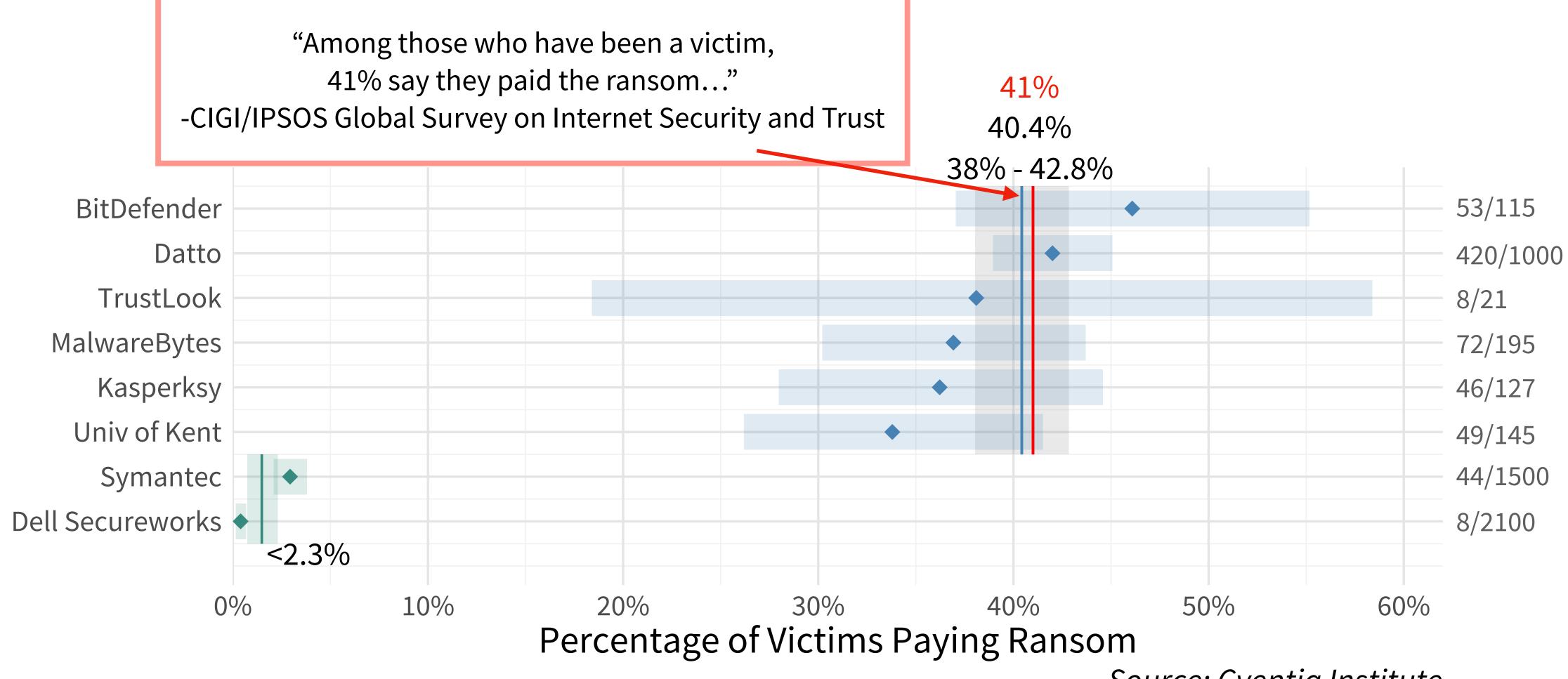


### • Early 2017 study

- 24,225 Internet users
- Across 24 countries (individual surveys) conducted)
- Weighted to match populated of country



# How many orgs are paying?

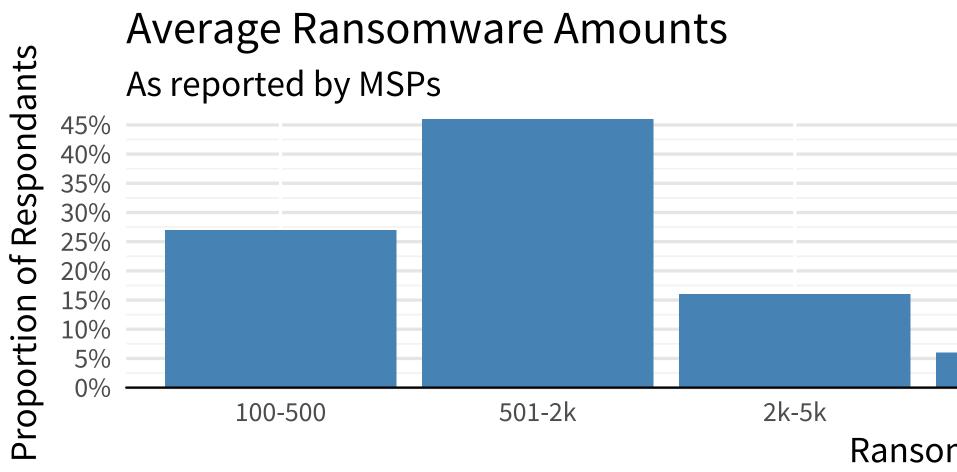


INSTITUTE

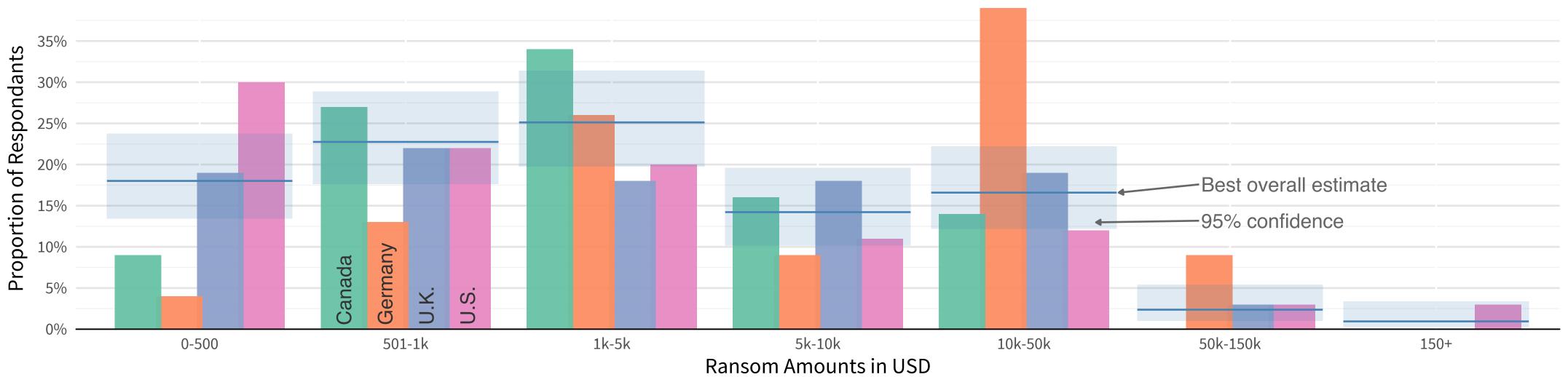
Source: Cyentia Institute



## Ransom Amounts



### Ransomware Amounts by Country





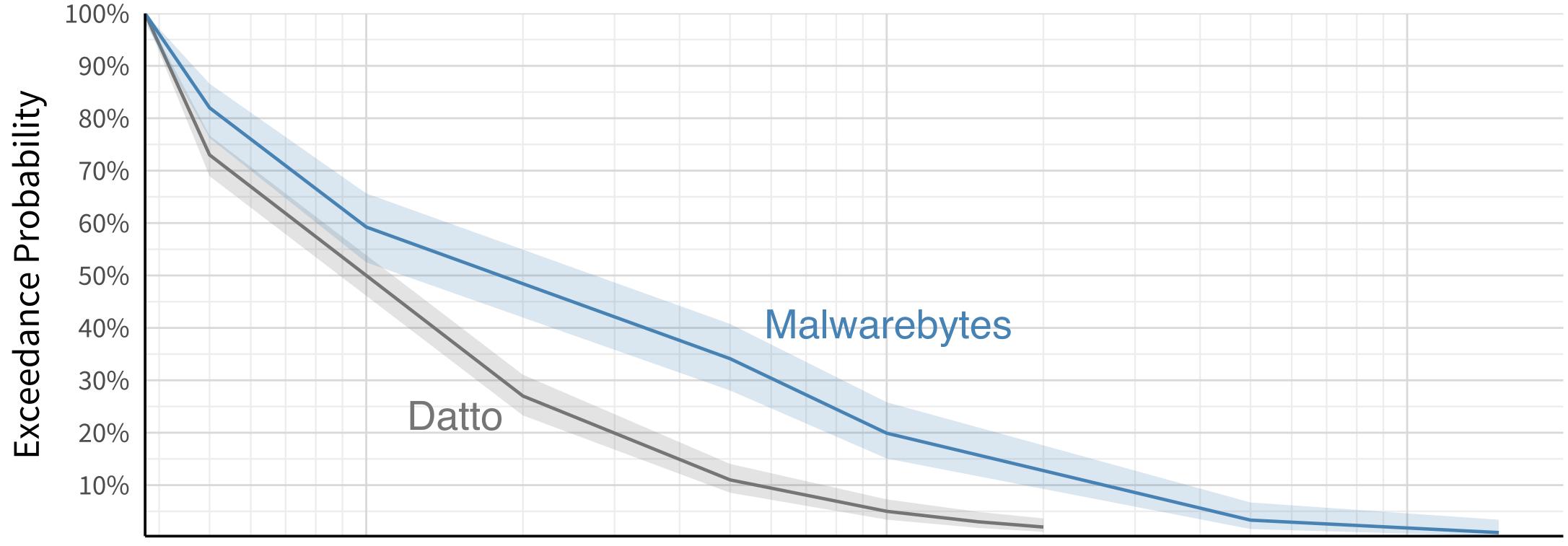
5k-	-10k 10	k-15k 15	k-20k 2	.0k+
Ransom Amo	ounts (in USD)			

Source: Cyentia Instutue, data from: Datto's State of the Channel Ransomware Report, 2016

Source: Cyentia Institute, data from "MalwareBytes/Osterman Research, "Understanding the Depth of the Global Ransomware Proble

# Exceeding Ransom Amount





1,000

### 10,000 Ransom Amounts (in USD)

Source: Cyentia Instutue, data from: MalwareBytes/Osterman Research, "Understanding the Depth of the Global Ransomware Problem", Datto's "State of the Channel Ransomware Report 2016"





100,000

# Challenges: Lessons Learned

- Experiment successful!
- Quality of vendor reports was terrible, rejected 2 out 3 on average "Not all reports are equal; parties have various motivations to publish, which creates divergent interpretations of what represents research worth communicating." - Geer, Jacobs 2014
  - Very poor, circular or missing citations
  - Terminology is loose and/or confusing
  - Object of measurement and framing is muddled or misaligned
  - ... is Ponemon: 51% (perception), 36% (included), 1.2% (excluded on wording)
  - Getting a simple sample size shouldn't be this hard
- Synthesizing the evidence was relatively straight-forward.

### While Library helped, identifying and narrowing down sources was a challenge \*\*



NSTIT

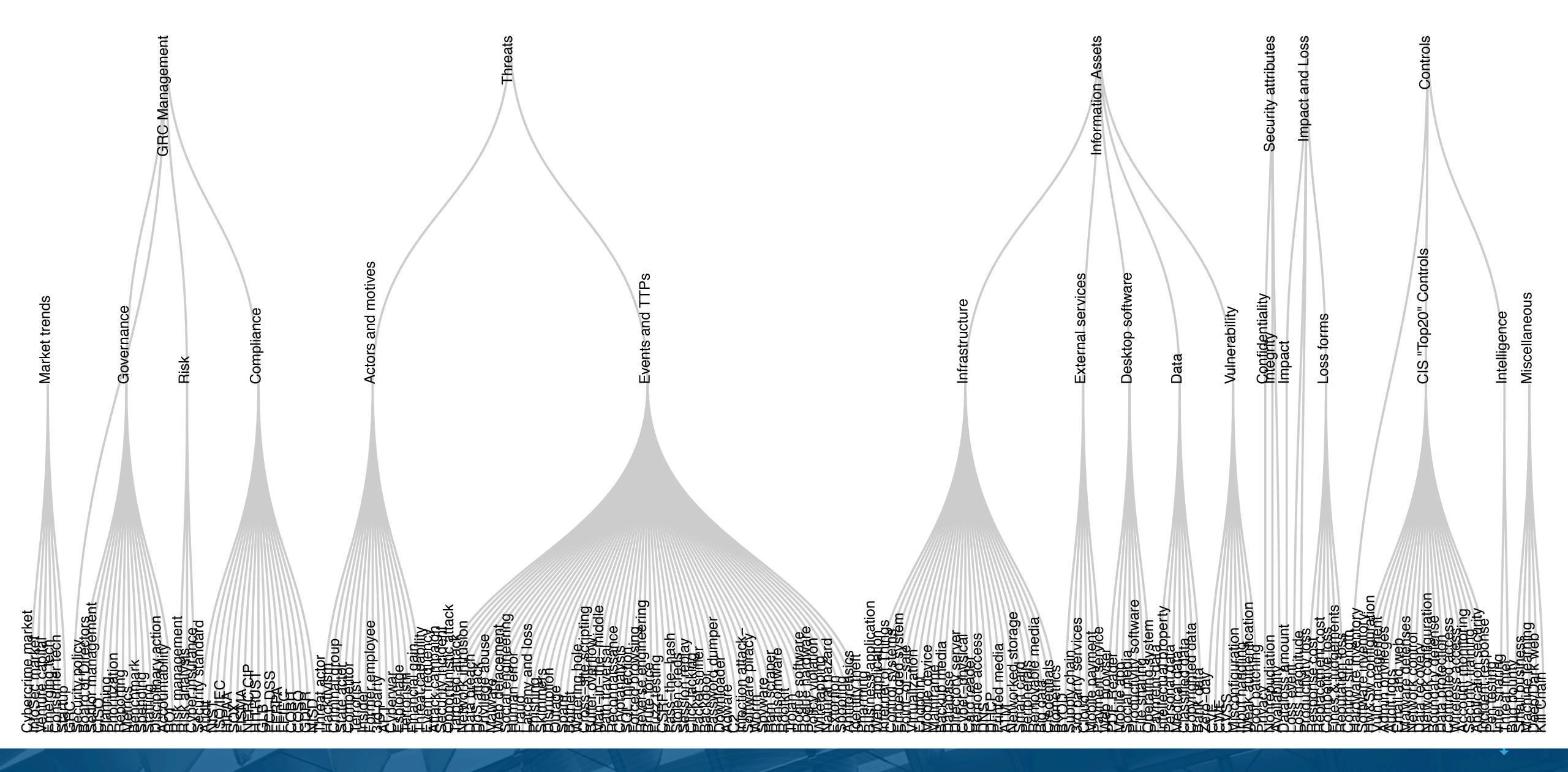
# Cyentia Library: Present and Future



# INSTITUTE



# Cyentia Library Tagging



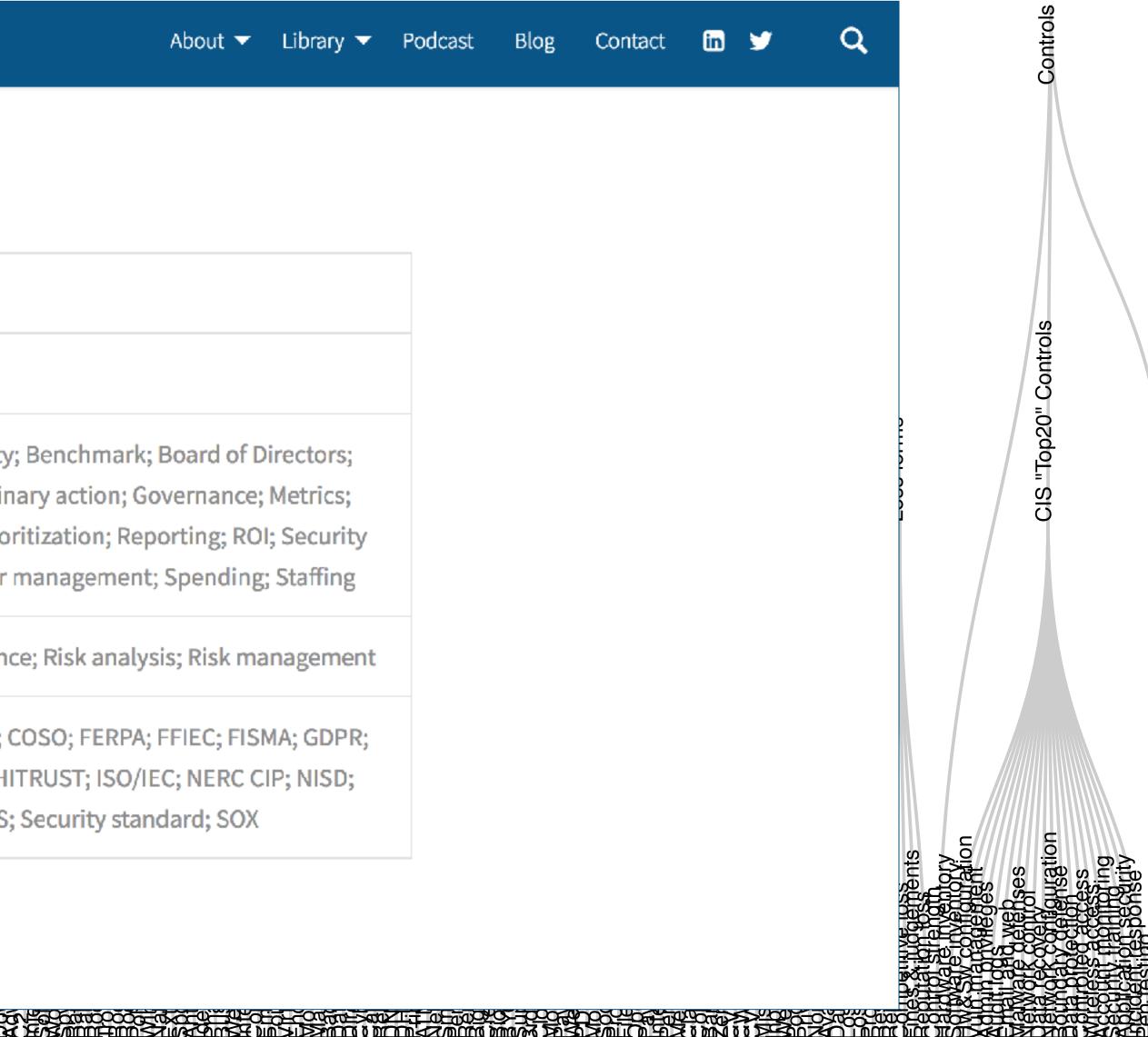
# INSTITUTE



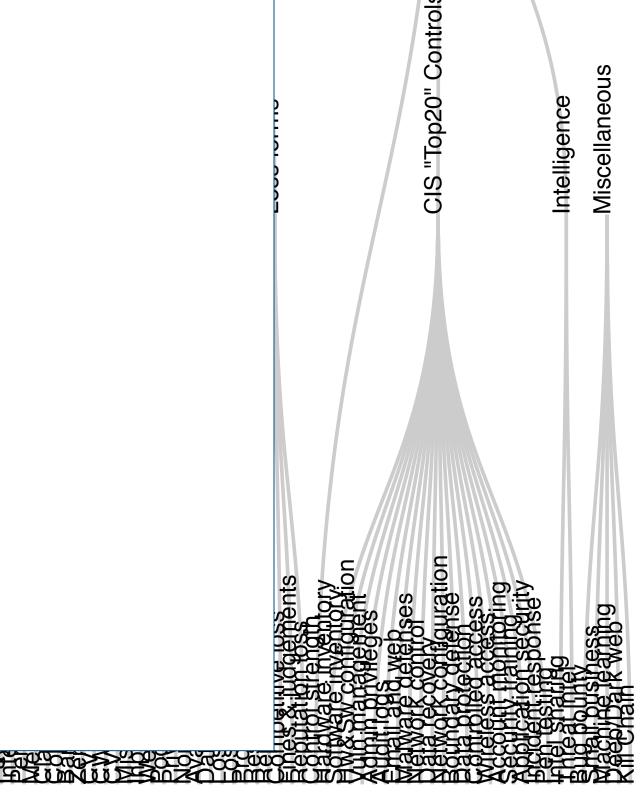
# Cyentia Library Tagging

GRC Management	CYENTIA INSTITUTE Topic: GRC Manageme	ent
	SUBTOPICS	TAGS
<u>v</u>	GRC	GRC
Governance	Governance	Accountability CISO; Disciplin Planning; Prio policy; Senior
	Risk	Cyber insuran
	Compliance	Audit; COBIT; GLB; HIPAA; H NIST; PCI-DSS
secrime market blangregenent dity policytation dity policytation dity on dity on di dity on dity on dity on dity on di	ଦ୍ୱ SEARCH LIBRARY	

# INSTITUTE

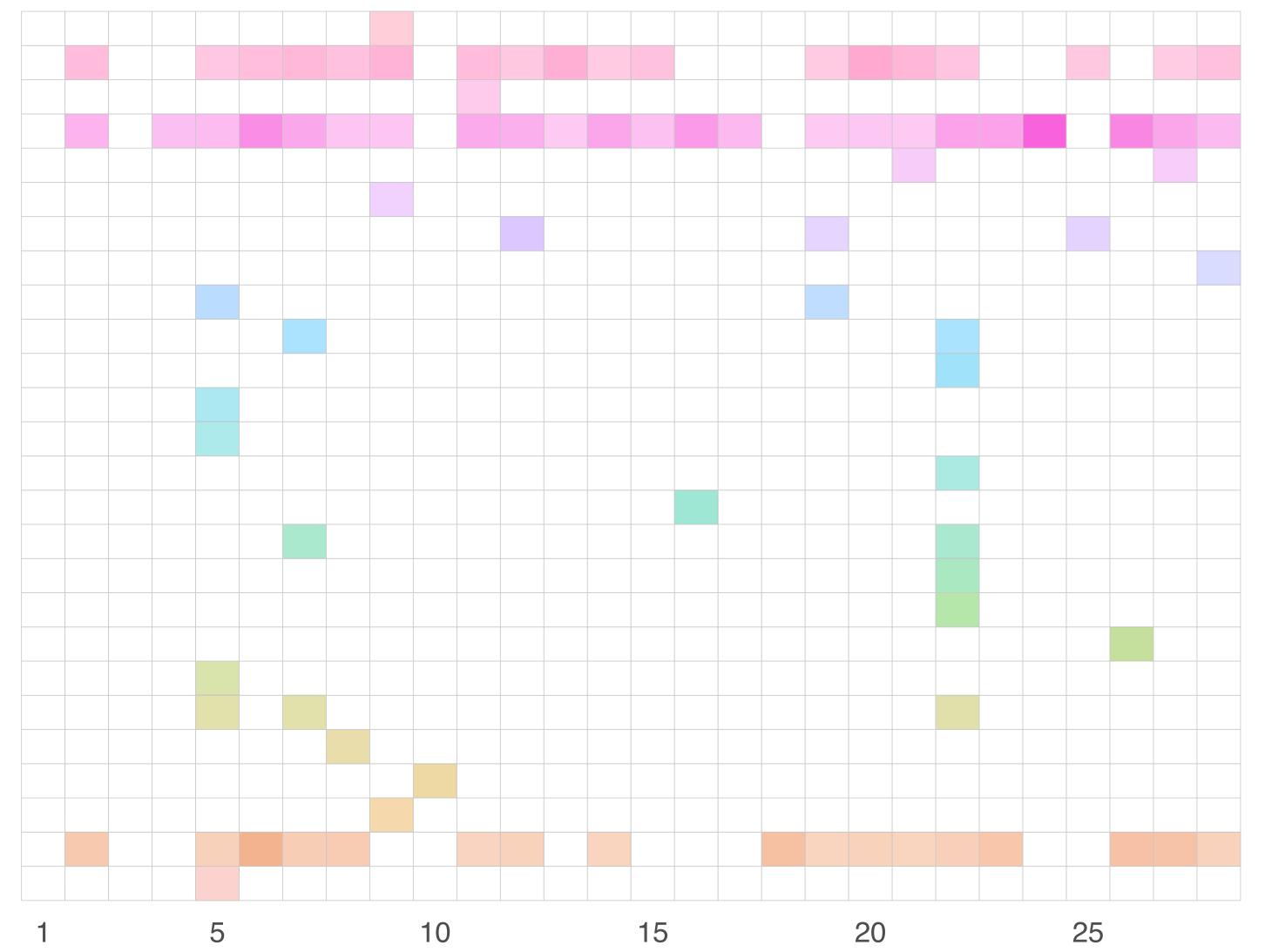






## Report: Hacker One

Web browser Vulnerability Vuln management Threat actor Staffing SQL injection Social media Senior management Security incident Personal data Pen testing Outage Operating system Mobile app Malware Intellectual property Impact Identity theft Financial gain Extortion Data breach Cyber-physical CSRF Cross-site scripting Bug bounty Availability



PDF Page

# CYENT A INSTITUTE



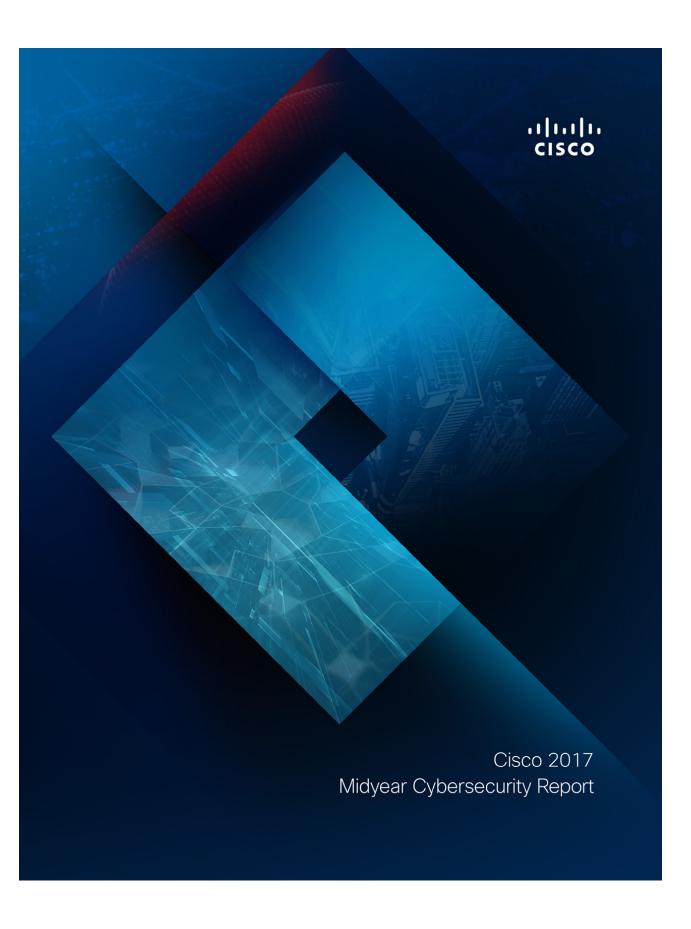


# Report: Cisco Mid-year Report 2017



Zero-day Web browser Vulnerability Vuln managemen Threat intel Threat actor Targeted attack Supply chain Staffing Spyware Spam Social media Security standard Security policy Security incident Risk management Ransomware Productivity loss Planning Phishinč Personal dată Outage Operating system Network intrusion Mobile device Malware defenses Malware Integrity InfoSec market Fraud Extortion Event frequency Endpoin Emerging tech Downloader DoS attack Disruption Database Data breach Criminal group Credentials Control systems Competitor Cloud Boundary defense Botnet Board of Directors Audit logs Audit APT 3rd party services

# NSTITUTE

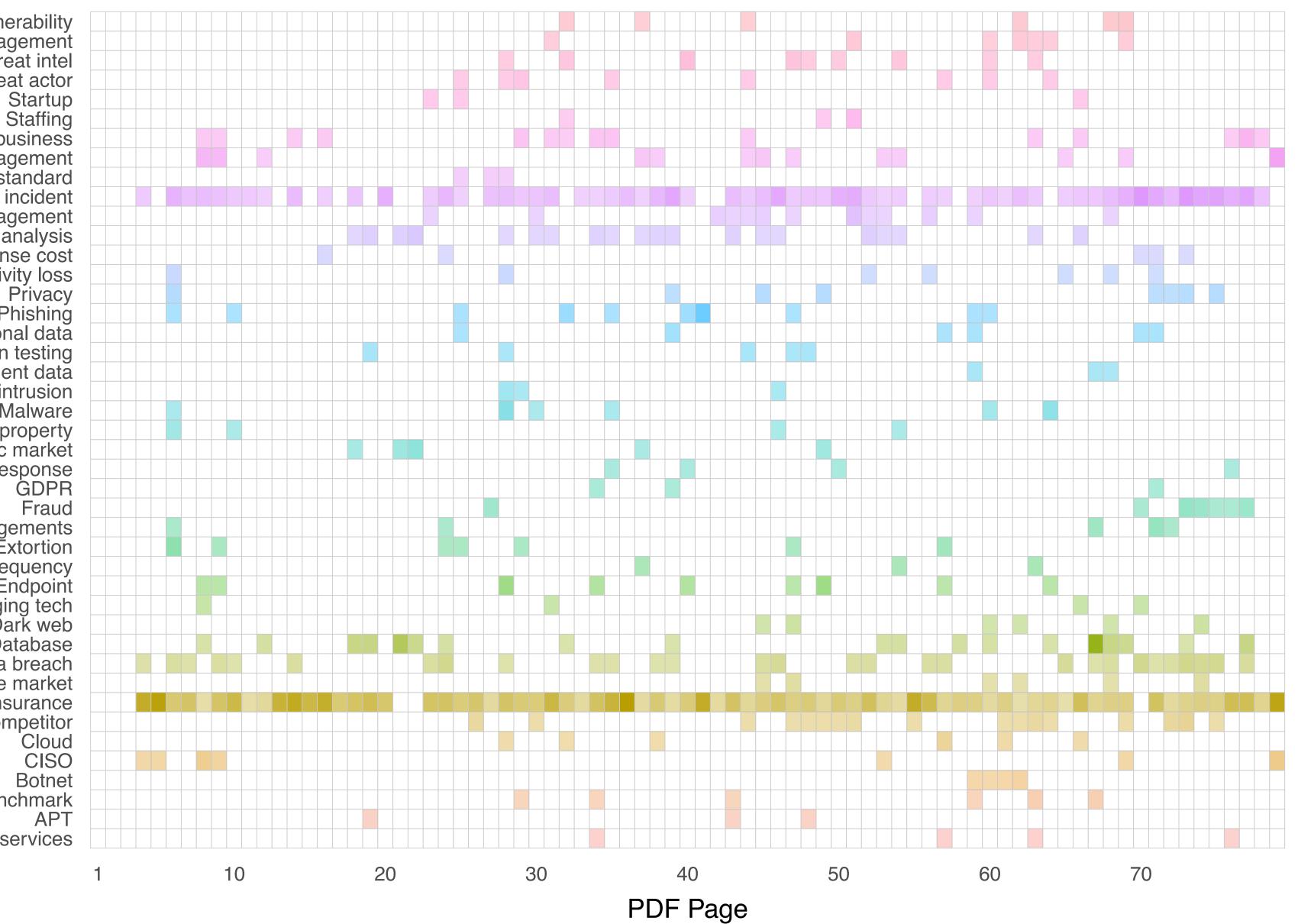






# Aite: Cyber Insurance

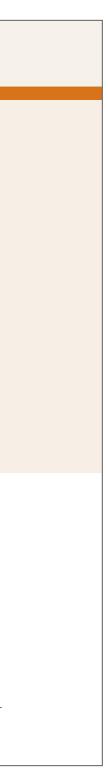
Vulnerability Vuln management Threat intel Threat actor Startup Staffing Small business Senior management Security standard Security incident Risk management Risk analysis Response cost Productivity loss Privacy Phishing Personal data Pen testing Payment data Network intrusion Malware Intellectual property InfoSec market Incident response GDPR Fraud Fines & judgements Extortion Event frequency Endpoint Emerging tech Deep/Dark web Database Data breach Cybercrime market Cyber insurance Competitor Cloud CISO Botnet Benchmark APT 3rd party services

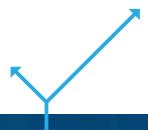


### $Y E N T \parallel A^*$ INSTITUTE

Aite partner. advisor. catalyst.
Cyber Insurance and Cybersecurity: The Convergence
June 2016
Gwenn Bézard
© 2016 Aite Group LLC. All rights reserved. Reproduction of this report by any means is strictly prohibited. Photocopying or electronic distribut is document or any of its contents without prior written consent of the publisher violates U.S. copyright law, and is punishable by statutory of up to US\$150,000 per infringement, plus attorneys? Fees (17 USC 504 et seq.). Without advance permission, illegal copyrigh includes regular

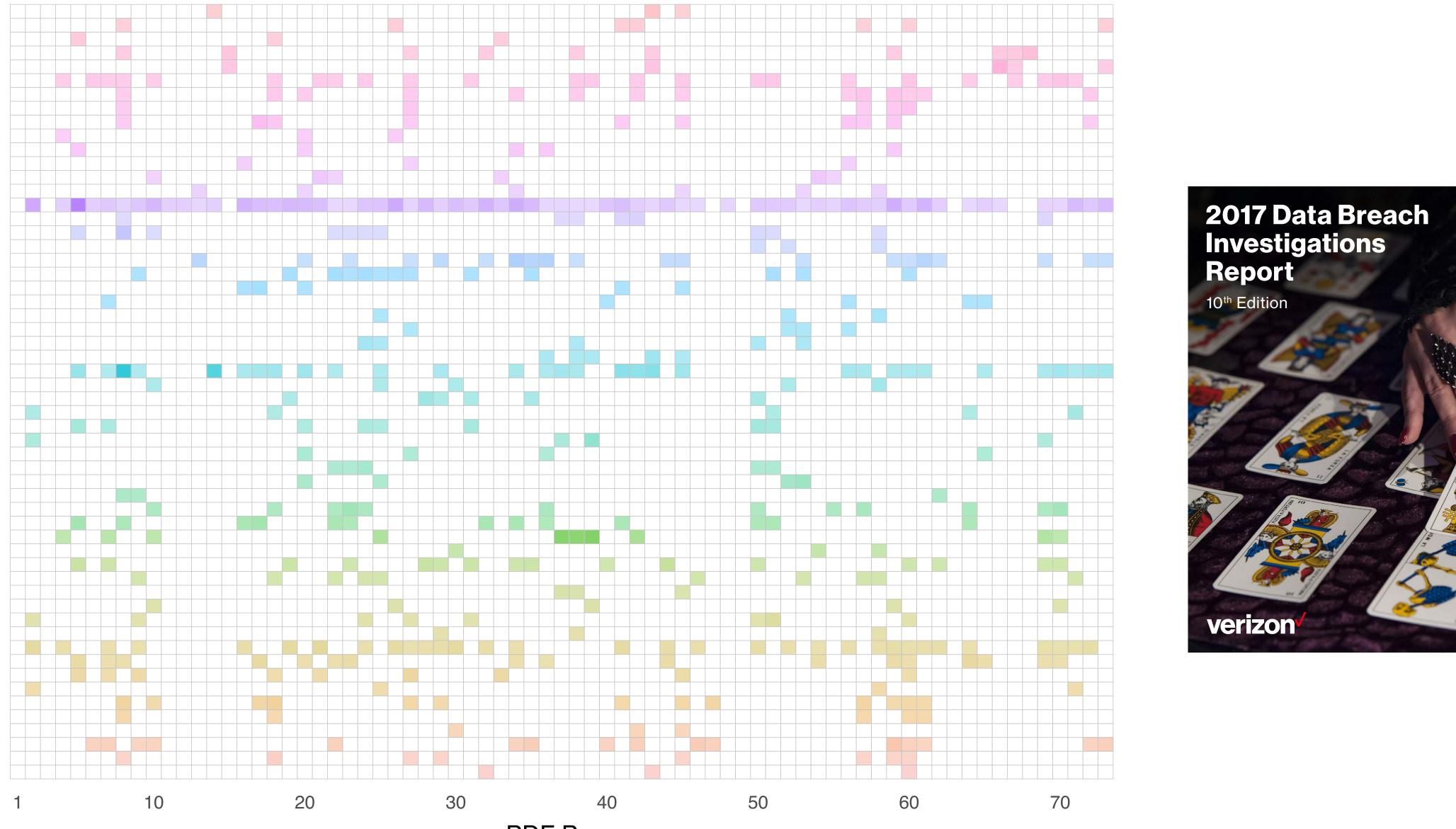






## Verzion DBIR 2017

Web browser Web application Weak authentication Vulnerability Vuln management Threat actor Stolen creds SQL injection Social engineering Small business Security training Security training Security incident Ransomware Privilege abuse Printed media Phishing Personal data Opportunistic attack Network intrusion Mobile device Misconfiguration Mobile device Misconfiguration Malware Larceny and loss Intellectual property Integrity InfoSec market Incident response Identity theft Human error Hacktivism Fraud Financial gain Fraud Financial gain Extortion Event frequency Espionage Endpoint Emerging tech Disruption Database Data protection Data breach Criminal group Credentials Confidentiality C2 Brute force Boundary defense Botnet Backdoor Application security



PDF Page

# INSTITUTE



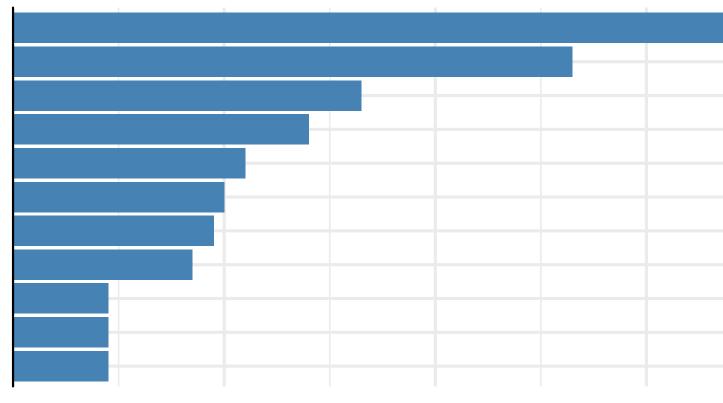




# Topic/Tagging

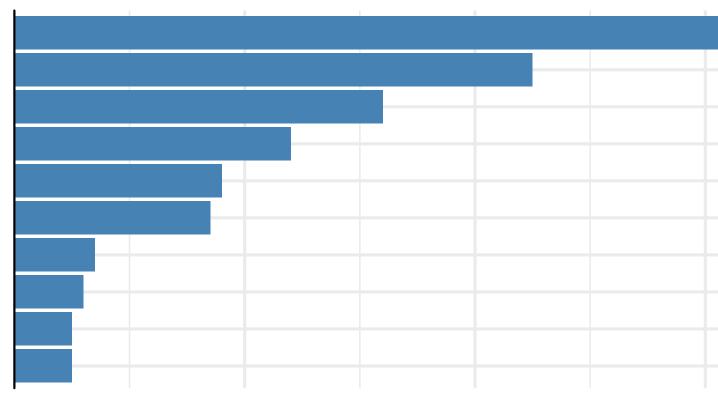
### Cisco Midyear 2017

Threats > Events and TTPs Threats > Actors and motives Information Assets > Infrastructure Controls > CIS "Top20" Controls Information Assets > Vulnerability GRC Management > Governance Information Assets > External services Information Assets > Desktop software Market trends > Emerging tech Information Assets > Data GRC Management > Compliance



### Verizon DBIR 2017

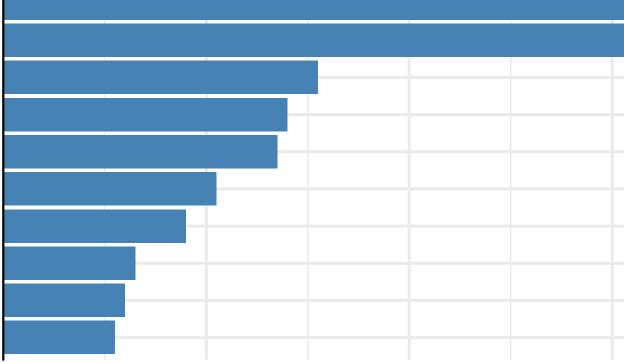
Threats > Events and TTPs Threats > Actors and motives Information Assets > Infrastructure Controls > CIS "Top20" Controls Information Assets > Data Information Assets > Vulnerability Information Assets > Desktop software GRC Management > Governance Security attributes > Integrity Security attributes > Confidentiality



INSTITUTE

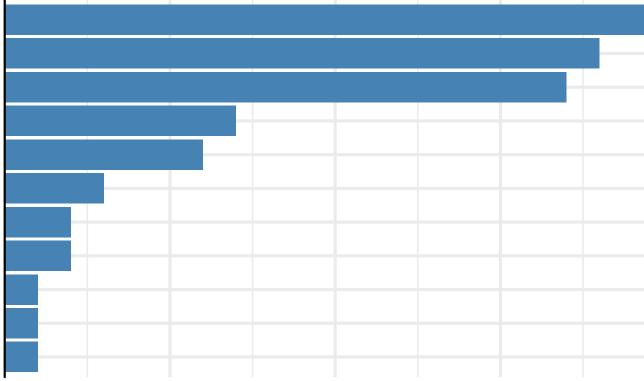
### Aite: Cyber Insurance

GRC Management > Risk Threats > Events and TTPs Information Assets > Infrastructure Threats > Actors and motives GRC Management > Governance Miscellaneous > NA Controls > CIS "Top20" Controls Impact and Loss > Loss forms Information Assets > Data Information Assets > External services



Threats > Actors and motives Information Assets > Vulnerability Miscellaneous > Misc Threats > Events and TTPs Information Assets > Desktop software GRC Management > Governance Information Assets > Data Controls > CIS "Top20" Controls Security attributes > Availability Information Assets > Infrastructure Impact and Loss > Impact

### HackerOne: Bug Bounty





1	



# Parsing PDFs: Text Extraction

### **TOP 3 LOCATIONS WHERE** DATA IS AT RISK IN VOLUME:

- Databases (49%)
- File Servers (39%)
- Cloud (36%)

### Corporate servers and databases pose the highest risk, yet spending remains stubbornly focused on endpoint and mobile

The top three locations by volume where companysensitive data is stored and must be protected are: databases (49%), file servers (39%), and the rapid growth area for cloud service environments (36%). The position is fairly consistent across most major geographies and mainstream verticals including financial services, healthcare, and the retail sector.

Along with the ubiquitous use of databases and servers, cloud and more recently big data take-up levels now force a stronger protection case to be made. Growing data volumes, when put alongside worries about a lack of control over third-party access; the use of third-party admins; and data

locational issues when foreign intervention and legal sovereignty come into play, make the case for improving cloud-services data protection. Also, as more data needs to transition between on-premise systems and cloud and big data environments, organizations need to make use of more inclusive data protection facilities to control and protect their data as it moves between corporate systems.

Another discussion that should take place revolves around the perception of risk that mobile devices and user mobility bring to the table. By comparison only 20% of sensitive company data is held on mobile devices and, of that 20%, a large proportion is being held on company-owned laptops and other company-protected mobile devices. In our opinion the discussion isn't really about the data volumes involved, and if it were, 20% is still significant enough to cause anxiety. But the real concern for the 70% of IT Decision Makers who were worried about mobile device protection is firmly about the lack of control over the mobile devices that are in use. It is also about not having enough information to know what data has been copied to those devices and not having the controls in place to stop copies of company-sensitive data being made.

Good quality monitoring and access control technology provide part of the answer. Irrespective of where the data is being held, it is important to know and be able to control who gets access and what they can do with that access. This provides the ability to highlight and report on misuse that could otherwise put company-sensitive data at risk.

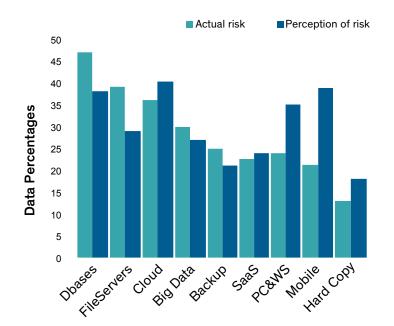


Figure 3: Data risks based on actual volumes of sensitive data stored in each location compared to the perception of risk

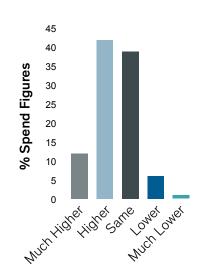


Figure 4: Global spending on security solutions during the next 12 months

### INSTITUTE

TOP 3 LOCATIONS WHERE DATA IS AT RISK IN VOLUME:

⢠Databases (49%) â□¢File Servers (39%) â□¢Cloud (36%)

Corporate servers and databases pose the highest risk, yet spending remains stubbornly focused on endpoint and mobile

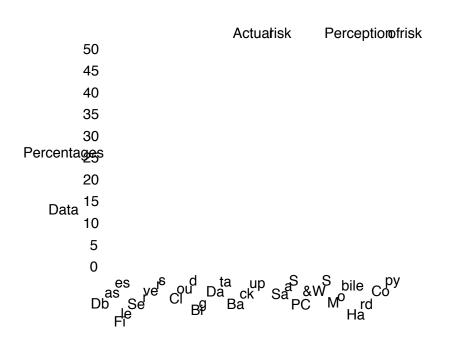
The top three locations by volume where companysensitive data is stored and must be protected are: databases (49%), file servers (39%), and the rapid growth area for cloud service environments (36%). The position is fairly consistent across most major geographies and mainstream verticals including financial services, healthcare, and the retail sector.

Along with the ubiquitous use of databases and servers, cloud and more recently big data take-up levels now force a stronger protection case to be made. Growing data volumes, when put alongside worries about a lack of control over third-party access; the use of third-party admins; and data

locational issues when foreign intervention and legal sovereignty come into play, make the case for improving cloud-services data protection. Also, as more data needs to transition between on-premise systems and cloud and big data environments, organizations need to make use of more inclusive data protection facilities to control and protect their data as it moves between corporate systems.

Another discussion that should take place revolves around the perception of risk that mobile devices and user mobility bring to the table. By comparison only 20% of sensitive company data is held on mobile devices and, of that 20%, a large proportion is being held on company-owned laptops and other company-protected mobile devices. In our opinion the discussionisnâ about the data volumes involved, and if it were, 20% is still significant enough to cause anxiety. But the real concern for the 70% of IT Decision Makers who were worried about mobile device protection is firmly about the lack of control over the mobile devices that are in use. It is also about not having enough information to know what data has been copied to those devices and not having the controls in place to stop copies of company-sensitive data being made.

Good quality monitoring and access control technology provide part of the answer. Irrespective of where the data is being held, it is important to know and be able to control who gets access and what they can do with that access. This provides the ability to highlight and report on misuse that could otherwise put company-sensitive data at risk.



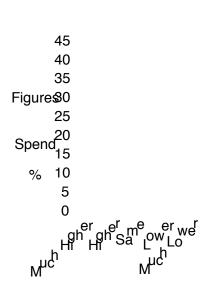


Figure 3: Data risks based on actual volumes of sensitive data stored in each location compared to the perception of risk

Figure 4: Global spending on security solutions during the next 12 months



# Parsing PDFs: early attempt

### **TOP 3 LOCATIONS WHERE** DATA IS AT RISK IN VOLUME:

- Databases (49%)
- File Servers (39%)
- Cloud (36%)

### Corporate servers and databases pose the highest risk, yet spending remains stubbornly focused on endpoint and mobile

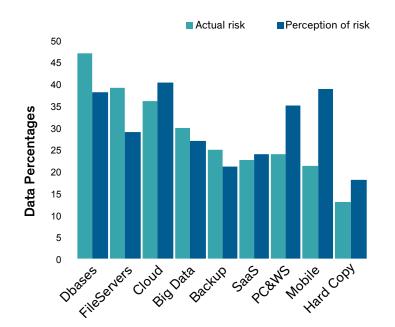
The top three locations by volume where companysensitive data is stored and must be protected are: databases (49%), file servers (39%), and the rapid growth area for cloud service environments (36%). The position is fairly consistent across most major geographies and mainstream verticals including financial services, healthcare, and the retail sector.

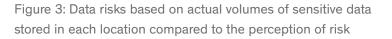
Along with the ubiquitous use of databases and servers, cloud and more recently big data take-up levels now force a stronger protection case to be made. Growing data volumes, when put alongside worries about a lack of control over third-party access; the use of third-party admins; and data

locational issues when foreign intervention and legal sovereignty come into play, make the case for improving cloud-services data protection. Also, as more data needs to transition between on-premise systems and cloud and big data environments, organizations need to make use of more inclusive data protection facilities to control and protect their data as it moves between corporate systems.

Another discussion that should take place revolves around the perception of risk that mobile devices and user mobility bring to the table. By comparison only 20% of sensitive company data is held on mobile devices and, of that 20%, a large proportion is being held on company-owned laptops and other company-protected mobile devices. In our opinion the discussion isn't really about the data volumes involved, and if it were, 20% is still significant enough to cause anxiety. But the real concern for the 70% of IT Decision Makers who were worried about mobile device protection is firmly about the lack of control over the mobile devices that are in use. It is also about not having enough information to know what data has been copied to those devices and not having the controls in place to stop copies of company-sensitive data being made.

Good quality monitoring and access control technology provide part of the answer. Irrespective of where the data is being held, it is important to know and be able to control who gets access and what they can do with that access. This provides the ability to highlight and report on misuse that could otherwise put company-sensitive data at risk.





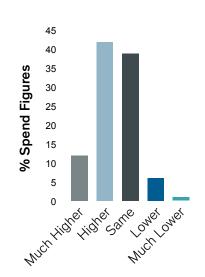


Figure 4: Global spending on security solutions during the next 12 months



### INSTITUTE

TOP 3 LOCATIONS WHERE DATA IS AT RISK IN VOLUME: ⢠Databases (49%) â□¢File Servers (39%) â□¢Cloud (36%)

Corporate servers and databases pose the highest risk, yet spending remains stubbornly focused on endpoint and mobile

The top three locations by volume where companysensitive data is stored and must be protected are: databases (49%), file servers (39%), and the rapid growth area for cloud service environments (36%). The position is fairly consistent across most major geographies and mainstream verticals including financial services, healthcare, and the retail sector.

Along with the ubiquitous use of databases and servers, cloud and more recently big data take-up levels now force a stronger protection case to be made. Growing data volumes, when put alongside worries about a lack of dentrol over third-party access; the use of third-party admins; and data

locational issues when foreign intervention and legal sovereignty come into play, make the case for improving cloud-services data protection. Also, as more data needs to transition between on-premise systems and cloud and big data environments, organizations need to make use of more inclusive data protection facilities to control and protect their data as it moves between corporate systems.

Another discussion that should take place revolves around the perception of risk that mobile devices and user mobility bring to the table. By comparison only 20% of sensitive company data is held on mobile devices and, of that 20%, a large proportion is being held on company-owned laptops and other company-protected mobile devices. In our opinion the discussionisnâl about the data volumes involved, and if it were, 20% is still significant enough to cause anxiety. But the real concern for the 70% of IT Decision Makers who were worried about mobile device protection is firmly about the lack of control over the mobile devices that are in use. It is also about not having enough information to know what that has been copied to those devices and not having the controls in place to stop copies of company-sensitive data being made

Good quality monitoring and access control technology provide part of the answer. Irrespective of where the data is being held, it is important to know and be able to control who gets access and what they can do with that access. This provides the ability to highlight and report on misuse that could otherwise put company-sensitive data at risk.

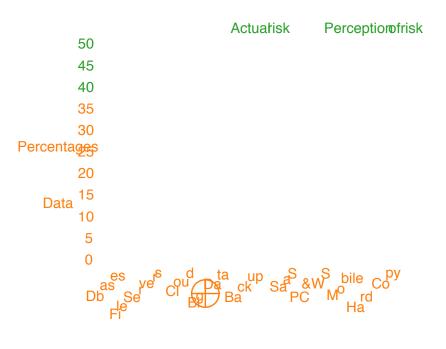


Figure 3: Data risks based on actual volumes of sensitive data stored in each location compared to the perception of risk

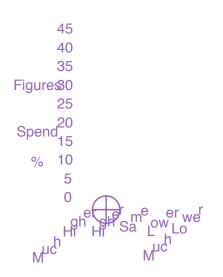


Figure 4: Global spending on security solutions during the next 12 months



# Parsing PDFs Spatially

### **TOP 3 LOCATIONS WHERE** DATA IS AT RISK IN VOLUME:

- Databases (49%)
- File Servers (39%)
- Cloud (36%)

### Corporate servers and databases pose the highest risk, yet spending remains stubbornly focused on endpoint and mobile

The top three locations by volume where companysensitive data is stored and must be protected are: databases (49%), file servers (39%), and the rapid growth area for cloud service environments (36%). The position is fairly consistent across most major geographies and mainstream verticals including financial services, healthcare, and the retail sector.

Along with the ubiquitous use of databases and servers, cloud and more recently big data take-up levels now force a stronger protection case to be made. Growing data volumes, when put alongside worries about a lack of control over third-party access; the use of third-party admins; and data

locational issues when foreign intervention and legal sovereignty come into play, make the case for improving cloud-services data protection. Also, as more data needs to transition between on-premise systems and cloud and big data environments, organizations need to make use of more inclusive data protection facilities to control and protect their data as it moves between corporate systems.

Another discussion that should take place revolves around the perception of risk that mobile devices and user mobility bring to the table. By comparison only 20% of sensitive company data is held on mobile devices and, of that 20%, a large proportion is being held on company-owned laptops and other company-protected mobile devices. In our opinion the discussion isn't really about the data volumes involved, and if it were, 20% is still significant enough to cause anxiety. But the real concern for the 70% of IT Decision Makers who were worried about mobile device protection is firmly about the lack of control over the mobile devices that are in use. It is also about not having enough information to know what data has been copied to those devices and not having the controls in place to stop copies of company-sensitive data being made.

Good quality monitoring and access control technology provide part of the answer. Irrespective of where the data is being held, it is important to know and be able to control who gets access and what they can do with that access. This provides the ability to highlight and report on misuse that could otherwise put company-sensitive data at risk.

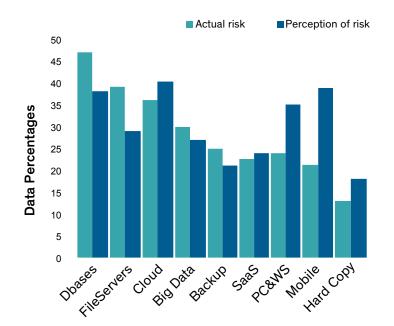


Figure 3: Data risks based on actual volumes of sensitive data stored in each location compared to the perception of risk

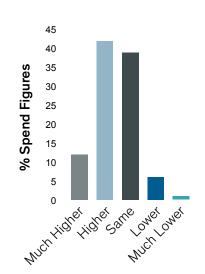
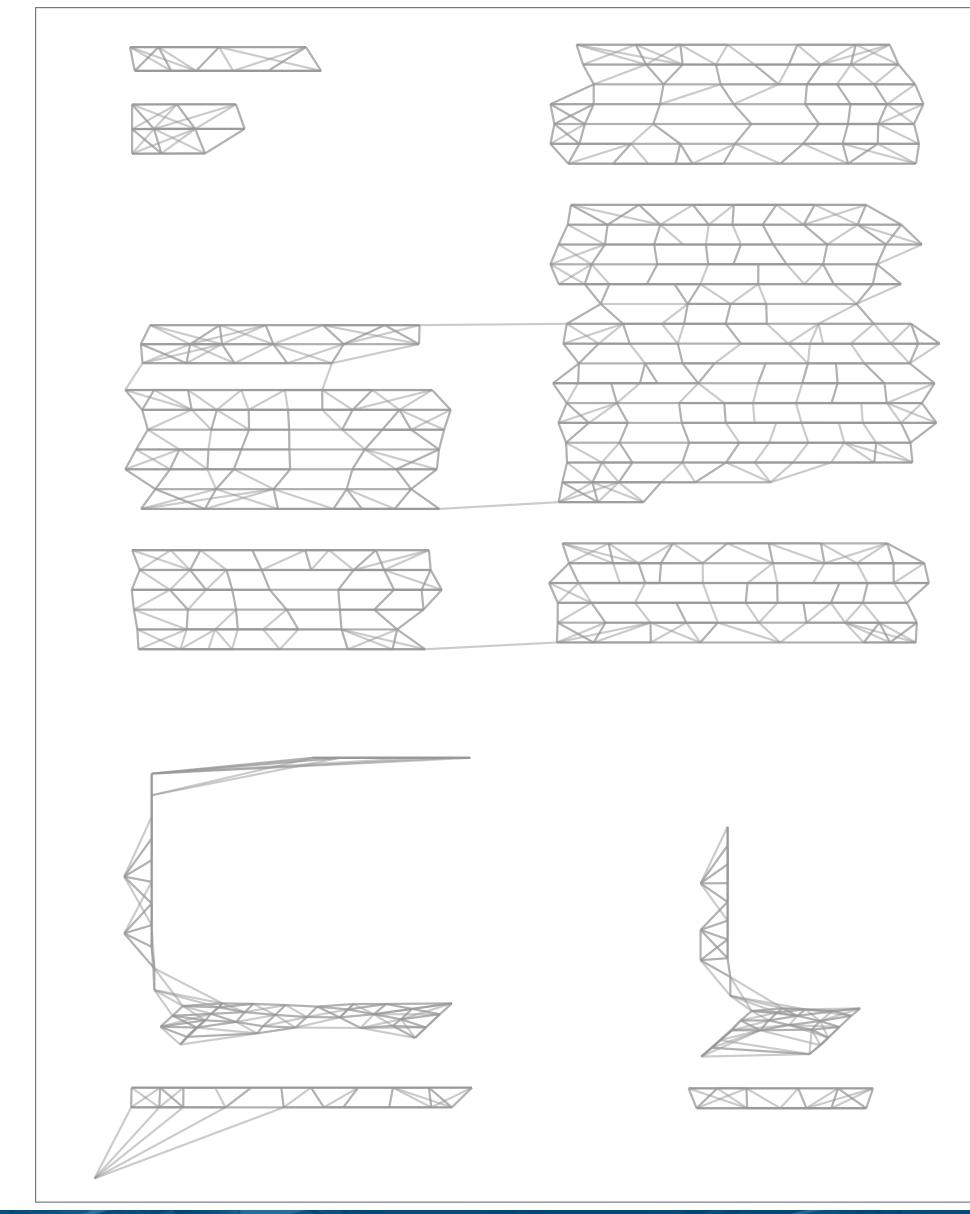


Figure 4: Global spending on security solutions during the next 12 months

### INSTITUTE





## Parsing PDFs

### **TOP 3 LOCATIONS WHERE** DATA IS AT RISK IN VOLUME:

- Databases (49%)
- File Servers (39%)
- Cloud (36%)

### Corporate servers and databases pose the highest risk, yet spending remains stubbornly focused on endpoint and mobile

The top three locations by volume where companysensitive data is stored and must be protected are: databases (49%), file servers (39%), and the rapid growth area for cloud service environments (36%). The position is fairly consistent across most major geographies and mainstream verticals including financial services, healthcare, and the retail sector.

Along with the ubiquitous use of databases and servers, cloud and more recently big data take-up levels now force a stronger protection case to be made. Growing data volumes, when put alongside worries about a lack of control over third-party access; the use of third-party admins; and data

locational issues when foreign intervention and legal sovereignty come into play, make the case for improving cloud-services data protection. Also, as more data needs to transition between on-premise systems and cloud and big data environments, organizations need to make use of more inclusive data protection facilities to control and protect their data as it moves between corporate systems.

Another discussion that should take place revolves around the perception of risk that mobile devices and user mobility bring to the table. By comparison only 20% of sensitive company data is held on mobile devices and, of that 20%, a large proportion is being held on company-owned laptops and other company-protected mobile devices. In our opinion the discussion isn't really about the data volumes involved, and if it were, 20% is still significant enough to cause anxiety. But the real concern for the 70% of IT Decision Makers who were worried about mobile device protection is firmly about the lack of control over the mobile devices that are in use. It is also about not having enough information to know what data has been copied to those devices and not having the controls in place to stop copies of company-sensitive data being made.

Good quality monitoring and access control technology provide part of the answer. Irrespective of where the data is being held, it is important to know and be able to control who gets access and what they can do with that access. This provides the ability to highlight and report on misuse that could otherwise put company-sensitive data at risk.

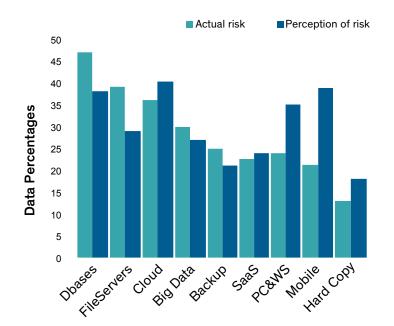


Figure 3: Data risks based on actual volumes of sensitive data stored in each location compared to the perception of risk

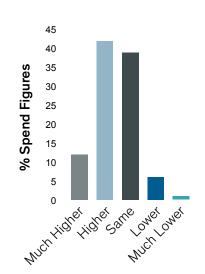


Figure 4: Global spending on security solutions during the next 12 months

### CYENT A INSTITUTE

### TOP 3LOCATION SVHERE DATAISATRISKINVOLUME:

â @atabase49%) â⊡¢Eil&erver\$39%) â⊡@lou(36%)

Corporateerversandiatabasesposethe highestrisk, yetspendingemainstubbornly focusedonendpointandmobile

Thetopthreacations/wolume/hereompanysensitivelataisstoredandmustbeprotectedare: database(\$19%)fileserver(\$39%)andtherapid growthareaforcloudervicenvironment86%). Theositionisfairleonsisteratcrosenostnajor geographieanchainstreamerticaliacluding financialervicelsealthcareanctheretailector.

Alongviththebiguitoussectlatabaseand serverscloudandmoreecenthbigdataake-up levelsnowforceastrongerrotectionasetobe madeGrowinglataolumesyherpuslongside worrieaboutalackofcontrobvethird-party accesstheuseofhird-partedminsanddata

location ad sue sy heriore igin terventio and egal sovereigntyomeintoplaymakethecasefoimproving cloud-servicedatarotectionAlsoasmoredatareeds toransitionetweem-premiseystemandcloudand bigdatanvironmentaganizationseedtomakeuse ofmorenclusivelaterotectiofacilitiescontroland protectheidatasimovebetweecorporataystems.

Anotheliscussiothatshouldakeplaceevolves aroundheerceptionsfriskthatnobildeviceand usemobiliteringtothetableBgomparisoonly20% osensitiveompanglataisheldormobildevices and,ofthat20%,alargeroportionsbeingheldon company-owneeptopsandotheompany-protected mobildevicesInouopiniorthdiscussiona really abouthedataolumeavolvedandifitwere20%isstill significantnoughtccausenxietyButthereatoncern forthe70% of Decision/lakers/hovers/orried/bout mobilelevicerotectionsfirmlyabouthelackofcontrol overthemobile/evices/hatareinuse.ttisals@abouthot havingnoughformation/know/hatlatahasbeen copied othoselevice and no having hecontrols in placetostopopieso6mpany-sensitive databeinomade.

Goodualityonitoringnotccessontrotechnology provid@artofth@nswdrrespectivefwherdhedata isbeinmeld it is moortantok now and beablet control whogetaccesandwhatheycandowiththataccess. Thiprovidesheabilitytchighlighandreportormisuse thabouldtherwisecompany-sensitivlataatrisk.



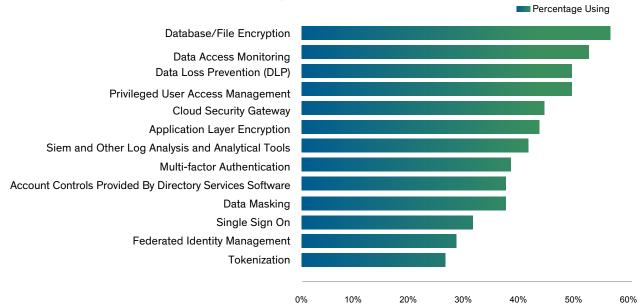
FigureGlobsplendingsecurity solutionsuringhenext2 nonths



## Parsing PDFs

### 2015 VORMETRIC INSIDER REPORT / GLOBAL EDITION

The most effective data protection technologies and the ones most frequently deployed by enterprise organizations were database and file encryption products, data access monitoring solutions, and data loss prevention technologies. As shown below, these topped a long list of protection solutions and were considered by enterprise respondents to offer the most effective protection against insider threats. Surprisingly tokenization, which has compliance-related uses, came bottom of the list. This may be due to restricted knowledge about the specific benefits the technology has. For example, if organizations need to protect data for specific purposes such as fulfilling payment card industry data security standard (PCI DSS) compliance, tokenization has scoping advantages over other forms of encryption that ensure the scope of audit requirements is reduced, as well as enabling the data to be used by other systems without compromising security.



**Security Protection Levels** 

Figure 7: Protection solutions used by enterprise organizations against insider threats

### THE MOST EFFECTIVE DATA **PROTECTION TECHNOLOGIES:**

• Database and file encryption

• Data Access Monitoring

### INSTITUTE

Themoseffectivedataprotectiotechnologieandthe onesmostrequentleployedbyenterpriserganizations werelatabasendfileencryptioproductsdata accessionitoringolutionsanddatalosprevention technologiesAsshow/below/thesetoppedalonglist of protectiosolutionsandwer@onsideredbyenterprise respondentsoofferthemoseffectiverotection againstnsidethreatsSurprisingtekenizationwhich hasompliance-relatedsescamebottomofthelist.This maybeduetorestrictellnowledgeboutthespecific benefitstheechnologyas.Foexampleiorganizations needtoprotectdataforspecifipurposesuchasfulfilling paymenbardndustrydatæecuritstandardPCDSS) complianctokenizationascopingdvantagesver otherformsofencryptionthatensurethescopeofaudit requirementisreduced as wella senabling hed at a to be usedbyothesystemswithoutompromisingecurity.

Securi®rotectiobevels

Percentageing

Database/Elleryption Dataccelstonitoring DataoBreventicDLP Privileged/seccelsanagement Cloublecurlinateway Application Application Sieran@thelco@malysian&halyticabols Multi-factothentication Accouldentre sovide Direct Servic Seftware Dataasking SinglSigrOn Federatetentitenagement Tokenization

0% 10% 20% 30% 40% 50% 60%

FigureProtectionalutionasedsynterpriseganizatioagainstsidehreats

THEMOSTEFFECTIVEDATA PROTECTIONTECHNOLOGIES:

â Batabasendfilencryption â⊡¢DataAccestionitoring



## Parsing PDFs

### 2015 VORMETRIC INSIDER REPORT / GLOBAL EDITION

### ANALYST PROFILE-ANDREW KELLETT, PRINCIPAL ANALYST SOFTWARE-IT SOLUTIONS, OVUM

Andrew enjoys the challenge of working with state-of-the-art technology. As lead analyst in the Ovum IT security team, he has the opportunity to evaluate, provide opinion, and drive the Ovum security agenda, including its focus on the latest security trends. He is responsible for research on the key technologies used to protect public and private sector organizations, their operational systems, and their users. The role provides a balanced opportunity to promote the need for good business protection and, at the same time, to research the latest threat approaches.

### HARRIS POLL-SOURCE/METHODOLOGY

Vormetric's 2015 Insider Threat Report was conducted online by Harris Poll on behalf of Vormetric from September 22-October 16, 2014, among 818 adults ages 18 and older, who work full-time as an IT professional in a company and have at least a major influence in decision making for IT. In the U.S., 408 ITDMs were surveyed among companies with at least \$200 million in revenue with 102 from the health care industries, 102 from financial industries, 102 from retail industries and 102 from other industries. Roughly 100 ITDMs were interviewed in the UK (103), Germany (102), Japan (102), and ASEAN (103) from companies that have at least \$100 million in revenue. ASEAN countries were defined as Singapore, Malaysia, Indonesia, Thailand, and the Philippines. This online survey is not based on a probability sample and therefore no estimate of theoretical sampling error can be calculated.



Andrew Kellett Principal Analyst Software IT Solutions, Ovum



### ABOUT VORMETRIC

Vormetric (@Vormetric) is the industry leader in data security solutions that protect data-at-rest across physical, big data and cloud environments. Vormetric helps over 1500 customers, including 17 of the Fortune 30, to meet compliance requirements and protect what matters—their sensitive data-from both internal and external threats. The company's scalable Vormetric Data Security Platform protects any file, any database and any application's data-anywhere it resides-with a high performance, market-leading solution set.

### INSTITUTE

### ANALYST PROFILE†ANDREMELLETT, PRINCIPAL ANALYST SOFTWAREâ€IT SOLUTIONS, OVUM

Andrew enjoys the challenge of working with state-of-the-art technology. As lead analyst in the Ovum IT security team, he has the opportunity to evaluate, provide opinion, and drive the Ovum security agenda, including its focus on the latest security trends. He is responsible for research on the key technologies used to protect public and private sector organizations, their operational systems, and their users. The role provides a balanced opportunity to promote the need for good business protection and, at the same time, to research the latest threat approaches.

### HARRIS

### POLL†SOURCE/METHODOLOGY

Vormetricâ€s2015 Insider Threat Report was conducted online by Harris Poll on behalf of Vormetric from September 22-October 16, 2014, among 818 adults ages 18 and older, who work full-time as an IT professional in a company and have at least a major influence in decision making for IT. In the U.S., 408 ITDMs were surveyed among companies with at least \$200 million in revenue with 102 from the health care industries, 102 from financial industries, 102 from retail industries and 102 from other industries. Roughly 100 ITDMs were interviewed in the UK (103), Germany (102), Japan (102), and ASEAN (103) from companies that have at least \$100 million in revenue. ASEAN countries were defined as Singapore, Malaysia, Indonesia, Thailand, and the Philippines. This online survey is not based on a probability sample and therefore no estimate of theoretical sampling error can be calculated.

Andrew Kellett Principal Analyst Software IT Solutions, Ovum

### ABOUT VORMETRIC

Vormetric (@Vormetric) is the industry leader in data security solutions that protect data-at-rest across physical, big data and cloud environments. Vormetric helps over 1500 customers, including 17 of the Fortune 30, to meet compliance requirements and protect what mattersâ€theisensitive dataâ€frdmoth internal and external threats. The companyâ€s scalable Vormetric Data Security Platform protects any file, any database and any applicationâ€s dataâ€anywiheræsidesâ€wita high performance, market-leading solution set.



# Data is Everywhere

- Security Industry has hundreds if not thousands of research reports released each year.
- **Meta-Analysis** is a promising approach (ransomware)
  - Research question > Identify Sources > Assess Quality > Synthesize Results
- Lots of opportunities to improve quality of research
- Discovery of publications is a challenge
  - Lower effort with better text extraction and NLP









# Data is Everywhere



Jay Jacobs jay@cyentia.com

# INSTITUTE

