Digital Forensics Association

The Leaking Vault 2011

Six Years of Data Breaches

A study conducted by Suzanne Widup

Published by the Digital Forensics Association

Publication Date: August 2011

The Leaking Vault 2011
Six Years of Data Breaches

Presented by the
Digital Forensics Association

Author: Suzanne Widup

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF FIGURES (Cont.)

LIST OF TABLES

EXECUTIVE SUMMARY

3,765 incidents

806.2 million records

$156.7 billion in data breach costs

Laptop theft the most common vector

Hacking exposed the most records

Outsiders caused the most damage

Findings

The Leaking Vault 2011 presents data gathered from studying 3,765 publicly disclosed data breach incidents, and is the largest study of its kind to date. Information was gleaned from the organizations that track these events, as well as government sources. Data breaches from 33 countries were included, as well as those from the United States.

This study covers incidents from 2005 through 2010, and includes over 806.2 million known records disclosed. On average, these organizations lost over 388,000 records per day/15,000 records per hour every single day for the past six years.

The estimated cost for these breaches comes to more than $156 billion to the organizations experiencing these incidents. This figure does not include the costs that the organizations downstream or upstream may incur, nor that of the data subject victims. Further, it is a low estimate of the cost, due to the fact that 35% of the incidents did not name a figure for records lost.

The Laptop vector remains the leader in incidents, but the Documents vector (printed material) is fast growing and demonstrates the need to manage both electronic data assets as well as printed documents. This vector has been trending upward for several years and is a potential contender for the incident leader if it continues.

The Hacking vector remains the records loss leader, responsible for 48% of the records disclosed in the study. The Drive/Media vector is in second place with the Web vector in third.

Outsiders continue to pose the largest risk in terms of both incidents and records disclosed. When the threat actor is an insider, the incident is significantly more likely to be accidental in nature. While accidental incidents are more prevalent, they also cause the most harm of the insider incidents in terms of records disclosed.

In 65% of the cases, the data disclosed included the data subject's name, address and Social Security Number. In contrast, only 15% of the incidents disclosed Credit Card Numbers, and 16% disclosed medical information. Medical disclosures saw a significant increase with the addition of the 2010 data. This is more likely due to the reporting requirement of existing regulations going into effect than any actual increase of incidents. The incidents where criminal use of the data was confirmed increased by 58% from the prior report. The two vectors most likely to show criminal use were the Fraud-SE and Hack vectors.

Recommendations

Data Mapping

- Know where your data is from inception to disposal. If you do not know where it enters the organization, where it is transformed, stored, shared with outside parties, archived, and finally disposed of—you cannot hope to keep it secure.
- Trace each data type from when it is created to when it is disposed and all the places it is used in between. Without making these types of data flow maps, organizations are operating on only a portion of the risk picture.

Data on the Move

- Laptops should come with a set of rules for the custodian of the device to follow. Include direction for maintaining physical control offsite (i.e., not to leave it in a vehicle, etc.) and onsite (i.e., lock it to their work surface).
- Don't neglect non-standard devices such as USB flash drives or even paper documents when looking at risk to your data.
- Manage the data lifecycle from creation to disposal and any transit point in between.

Early Detection is Critical

- Controls that make tampering evident upon inspection--or better still, alert those monitoring the systems automatically--should replace those that fail to notify when malicious behavior has occured.
- Preventative controls are ideal, but detective controls will limit the scope and damage of an event by allowing the discovery to happen sooner.
- Do not focus on a specific threat actor to the exclusion of all others. Focus on the highest risk to your environment first, and prioritize accordingly.

INTRODUCTION

Organizations seem to be in the news on a daily basis for disclosing data inappropriately. Hundreds of millions of people's personal private information has been lost, stolen or otherwise shared with unauthorized parties. The problem of data breaches is one that potentially impacts the economic health of the victim organizations, upstream or downstream partners, and the data subjects who face direct financial consequences.

There remain data breach disclosure laws on the books of 46 states, plus the District of Columbia, Puerto Rico and the Virgin Islands [17]. The states that do not currently have laws are New Mexico, Alabama, Kentucky, and South Dakota. Federal legislation has been proposed multiple times, but nothing has passed yet. Companies responding to incidents continue to have to navigate a maze of state requirements.

The Health Information Technology for Economic and Clinical Health Act (HITECH Act), which augmented the Privacy and Security sections of the Health Insurance Portability and Accountability Act (HIPAA) and had a breach notification requirement, took effect in November 2009. This means that breaches of health information became reportable for organizations that had not been covered under existing laws.

Here is a small sampling of the incidents from the study to put a personal face on the statistics:

- Three servers from a well-known chain restaurant were charged with using skimming devices to make more than $117,000 in fraudulent charges to customer credit card accounts.
- A restaurant employee stole customer credit card information and used it to purchase $200,000 of Walmart gift cards.
- In the span of six months, nine employees of a telecommunications company inappropriately accessed confidential customer account information and used it to make cloned cell phones. Over $15 million of unauthorized phone calls resulted from this scheme.
- An executive turned himself into authorities after being accused of selling customer information to identity thieves in exchange for sports tickets and gift cards.
- The owner of a medical equipment business used Medicare client information to obtain approximately $1.6 million worth of fraudulent claims.
- The owner of a farm equipment store pled guilty to federal charges, admitting she stole the identities of customers to obtain more than 80 loans worth $1.7 million.

The first two incidents show a common trend—wait staff using skimmers to obtain credit card data. Customer cards are then charged fraudulently, and in a growing number of cases, they are used to purchase gift cards and then these cards are either used for merchandise or sold for a fraction of their face value.

The remaining cases illustrate the damage an insider can wreak given sufficient access. Health care fraud was a significant trend in the 2010 data—in fact, since the HIPAA reporting requirement took effect, the number of medical data breaches have risen significantly. A percentage of those cases are perpetrated for medical insurance fraud.

One thing is clear—organizations must do more to ensure compliance with regulatory requirements for data protection and be proactive in securing the data they hold in their care. Those that focus only on compliance are more likely to suffer a breach than those who realize that they are the stewards of this data, and they have an obligation to protect it [13].

METHODOLOGY

The Leaking Vault (TLV) was published in July of 2010, and included data breach incidents from 2005 through 2009. This report (TLV2011) is a continuation of that work, and expands on the metrics developed in the prior publication. In this report, incidents from 2010 are incorporated and presented along with the changes that new data makes to the overall data set. The first TLV covered 2,807 incidents. Over the past year, an additional 181 breaches have been reported for those first five years, bringing the 2005 through 2009 total to 2,988. For 2010, there were a total of 777 breaches reported. This brings the number of cases in the study to 3,765 covering six years of data [1, 7, 9, 10, 11, 16, 18].

The sources of these breach reports expanded over the past year as more government agencies began publishing documentation in response to reports received by organizations complying with their data breach disclosure laws. While the three major data sources from the last report (The Open Security Foundation, Privacy Rights Clearinghouse and the Identity Theft Resource Center) continue to provide the bulk of the incidents, increasingly government sites factor in as incident sources [7, 11, 16]. For example, the Attorney General of Maryland's site lists original notification letters from organizations [9]. For health related incidents, the U.S. Department of Health and Human Services now maintains a database of medical data breaches as part of their material dedicated to HIPAA [18]. Unfortunately, the data from the HIPAA site is less detailed than some of the other sources, which hampered the gathering of certain metrics. However, as a source of medical breach incidents, it is a starting point for further research into the particulars of each case.

As with TLV, this report includes cases where sufficient detail exists about a breach to determine the organization that lost the data. The data involved must fall into the category of personal private data that the data subject did not consent to disclose. While individual elements of the data disclosed might be public, the combination of data is found in the wording of the data breach disclosure laws that all but four states have on their books.

ANALYSIS AND FINDINGS

The data was analyzed from several perspectives—frequency of incidents, number of records disclosed, breach vectors, organizational and data types. Also explored was the relationship between the data subjects and the organizations and how the data subject victims were treated. This remains consistent with the analysis from last year's report. The data is analyzed from both an overall perspective—with the full six years of data to determine if the additional cases changed any of the previous risk rankings—and the 2010 data as a separate data set to see if there are any new trends that deserve attention.

Frequency of Incidents

This section provides a view of the changes that the addition of these new incidents make to the overall study, and specific information about the 2010 incidents. Figure 1 shows the additional incidents in the first five years, as well as those from 2010. If you recall from TLV, there was a significant drop in incidents between 2008 and 2009. The decrease between the years was almost 200 cases, and there was discussion of whether this would be a trend. With the additional incidents that were reported after the publication of TLV, that gap has become less pronounced, and the 2010 incidents show a gradual decline from there. The more years we have in the study, the more actual long term trends should become visible.

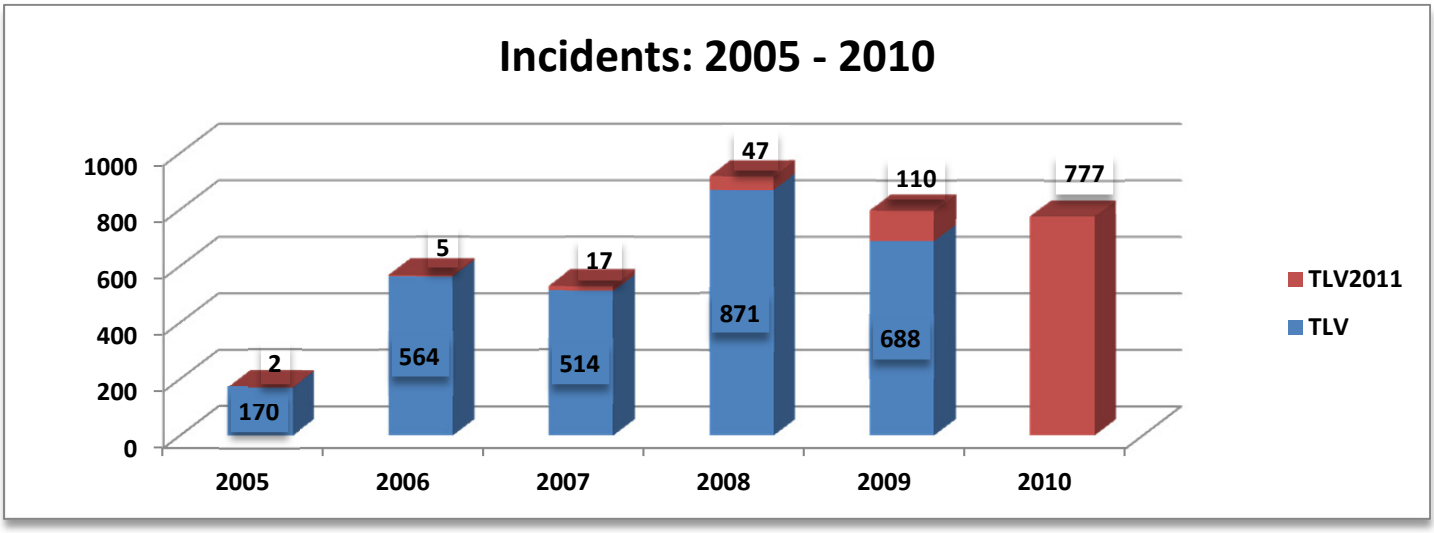
Overall Incidents (2005 – 2010)

As mentioned, the combined incidents total 3,765. As you can see from Figure 1, 2009 had the largest gain of incidents relative to the other years. The lead year is still 2008 with the most incidents reported, but 2009 no longer shows such a sharp drop off of events.

The Open Security Foundation continues to submit and process the results from Freedom of Information Act requests to obtain the original breach notification documents from government agencies. There were fewer source documents

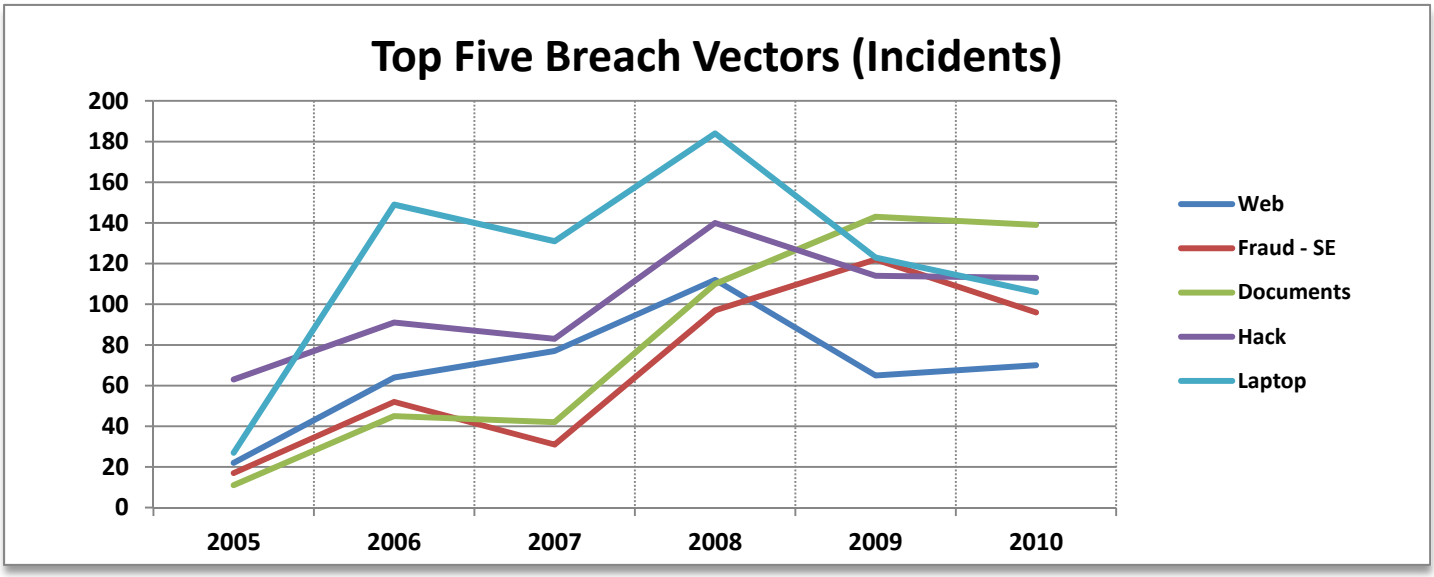
available in the new material, but since some state agencies are posting the data, these documents were sometimes obtainable from other sources. Government agencies posting original breach documents make it possible to obtain the best data on the incident, since it comes directly from the breached organization. Media reports may not contain as much detail as these source documents, and so there is a lower incidence of certain metrics in the 2010 data. Specifically, the only reliable data on whether the company offered credit card monitoring services to the data subjects is from these documents. Other metrics such as where a laptop was stolen from (home, office, vehicle, etc.) can sometimes be found in the media reports.

Figure 1: Data Breach Incidents (2005 - 2010)



As shown in Figure 2, the top vectors remain constant from this report to the last. The Laptop vector was the highest with 720 incidents (19%) over the course of the study. The Hacking vector took second place with 604 (16%), and the Documents vector moved into third place with 490 incidents (13%), displacing the Web vector.

Figure 2: Top Five Incident Breach Vectors (2005 – 2010)



Between 2009 and 2010, the number of incidents declined from 798 to 777. All except two vectors showed a decrease during that time period. Figure 3 shows the two that increased.

Figure 3: Increasing Vectors (2005 – 2010)

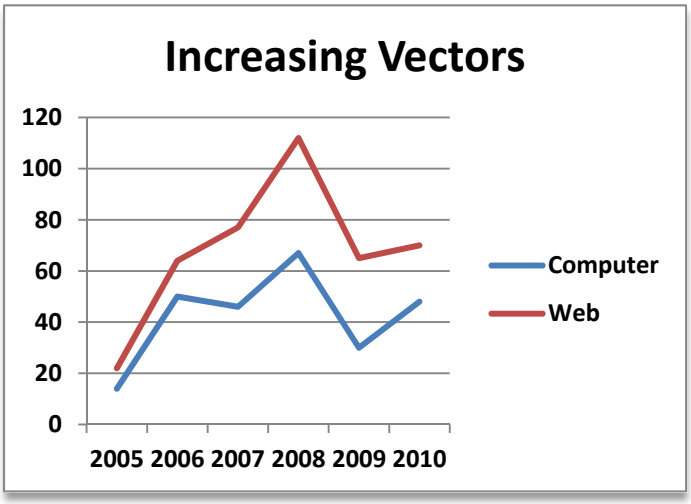
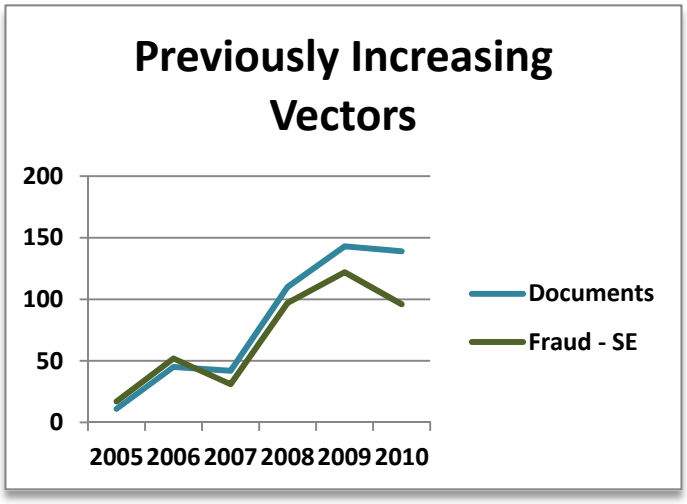


Figure 4: Previously Increasing Vectors (2005 – 2010)

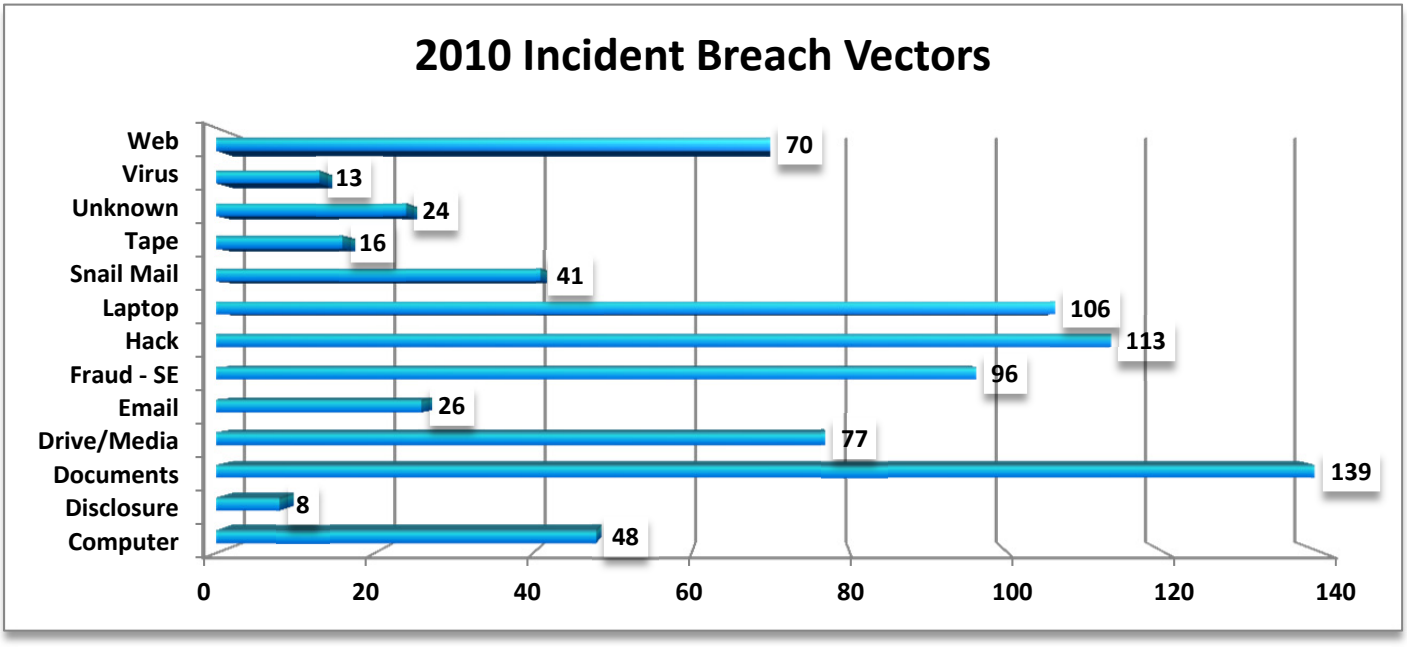


The last report showed the Documents and Fraud-SE vectors as the only two that increased in the drop between 2008 and 2009. As you can see from Figure 4, they both show a decrease with the addition of the 2010 data.

2010 Incidents

The breach vectors for 2010 are broken down in Figure 5. The Documents vector took the lead by a wide margin this year, with Hack and Laptop running close second and third respectively. The overall leader for the study remains the Laptop vector, but the Documents vector is a potential challenger.

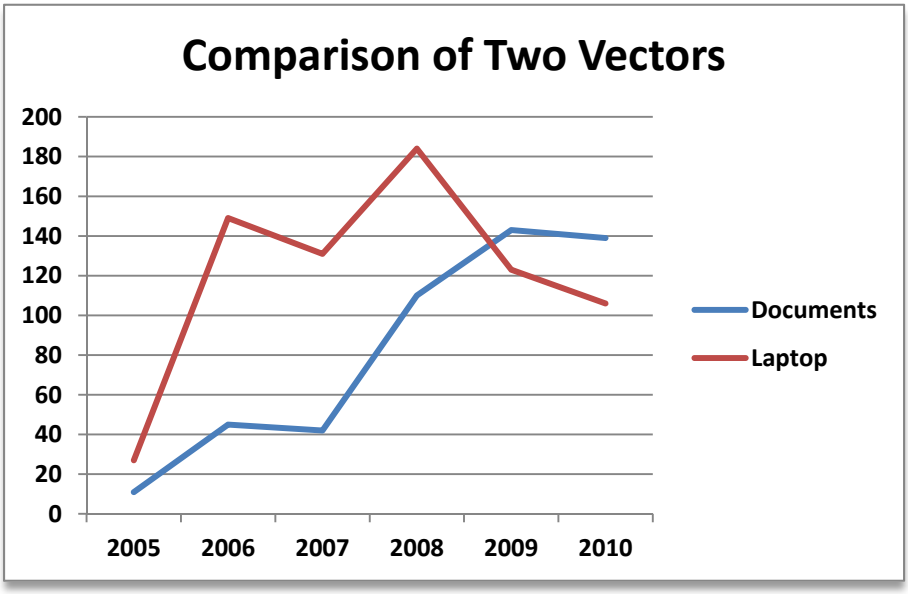
Figure 5: 2010 Incident Breach Vectors



2010 Trends

The breach vectors for 2010 are interesting, because the Documents vector seems to be trending towards overtaking the lead vector for incidents. In 2009, Documents had 143 incidents, where Laptop had only 123. This continued in the 2010 data when Documents had 139 and Laptop 106. Because the Laptop vector has been the incident leader longer, however, it still retains the overall lead. In total, there are 490 Document incidents versus 720 Laptop incidents. While Documents has a long way to go to get to the top spot, (with only 208 incidents for the first four years of data), it will be interesting to see if this trend continues in the future. Figure 6 shows the trend of the two vectors, with Laptop's downward trend for the last 2 years.

Figure 6: Comparison of Documents and Laptop Incident Vectors



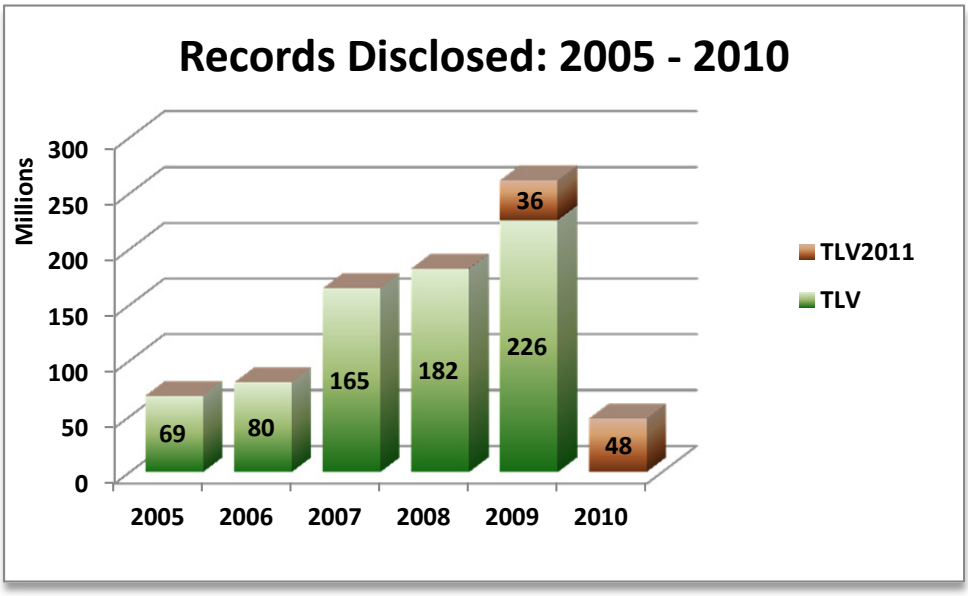
Another trend that was apparent in the 2010 data is the incidents of medical/dental offices having to declare breaches under HIPAA/Hitech when the practice suffers a burglary where the computer is stolen. These systems store their patient databases, and without encryption, most states require a breach notification. There were 87 incidents, accounting for just over 1 million records in 2010. This is in contrast to 150 incidents, accounting for 1.99 million records in the entire study. Whether this growth is due to the reporting requirements for HIPAA/Hitech going into effect remains to be seen in future years. A new metric that was added in 2010 is the time between the organization discovering the incident, and the breach being reported by the organization. One of the criticisms leveled at these organizations is the amount of time they took to notify the data subject victims. In 396 of the 777 cases in 2010, the time between when the incident's discovery and the reporting date were available. For those cases, the median time to notify was 45 days, and the mean was 64 days. The min was 1day and the max was 1,075 days—nearly 3 years.

The SAFE Data Act would specify how long organizations have to notify in the event of a breach. Specifically, 48 hours after the event, notification must be provided. This is a very short time, and there has been research showing that a month is a more optimal timeframe to require, based on cost and other factors. If companies are forced to report before they know the actual extent of the data affected, then the likelihood is that many people will start receiving notifications when their data was not actually at risk, because organizations are erring on the side of caution [4].

Records Disclosed

Figure 7 shows the additional records disclosed from the incidents added between TLV and TLV2011. It also illustrates the dramatic drop in the number of records disclosed. Only 48 million records were listed as disclosed in 2010. The number of incidents where the value for records disclosed was "unknown" and thus marked as zero in the database was 272, or 35%. This undisclosed total figure is not the highest in the study (which was 2009 at 47%), but it was sufficient to raise the overall average up by 1% for total unknown record incidents. More can be found about this in Table 3. While the first four years increased so minimally as to barely register in the chart, note the 2009 increase of 36 million records.

Figure 7: Records Disclosed (2005 – 2010)

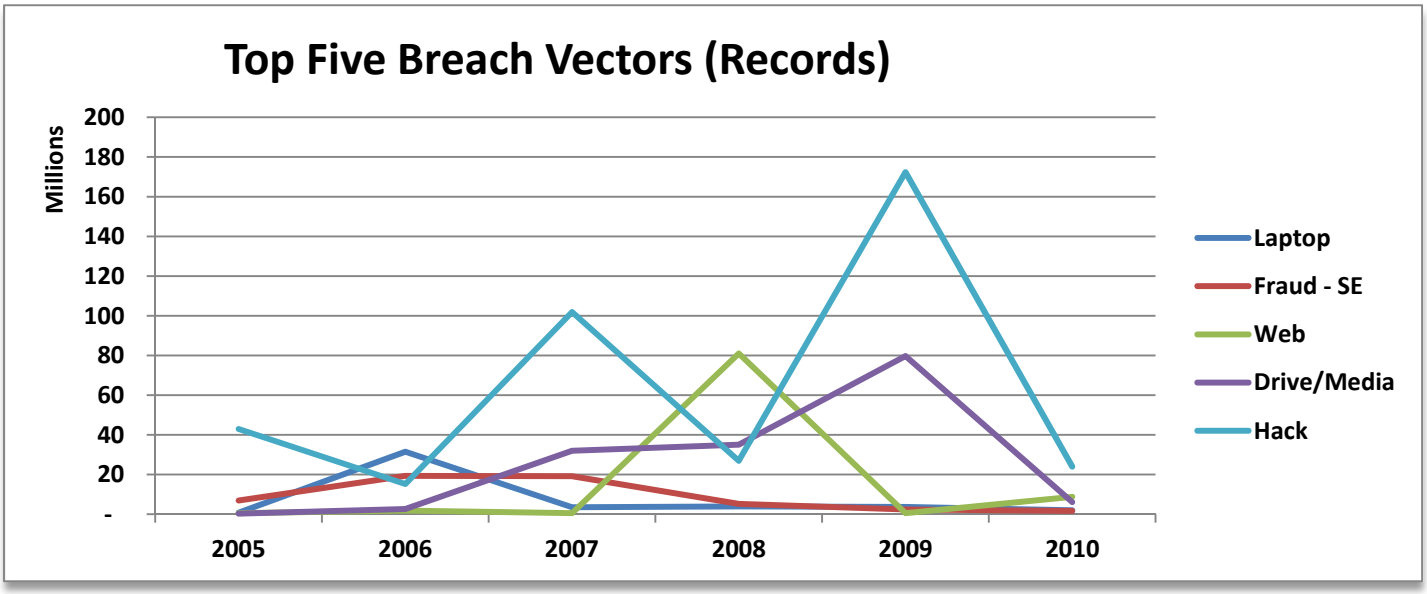


Easily the most glaring feature of Figure 7 is the drop in known records figure for 2010, ending the steady upward trend for records disclosure. In fact, 2010 saw the lowest number of disclosed records since the study began. This may be an anomaly, or it may be that the "unknown" records values are masking some very large breaches, and we just don't have data for them. Given that there were 777 breach incidents, which is not a correspondingly smaller number compared to other years in the study, the cause of the decrease is not clear.

Baker, Hutton & Hylender et.al speculate that the cause for the drop in records disclosed may be related to several high profile prosecutions of the figures who were responsible for some of the largest breaches in the study. Their other assertion is that these have had a deterring factor on the remaining actors to refrain from the larger incidents and avoid the attention they garner [2]. The reason remains uncertain, and it will be monitored to see if this represents the start of a trend, or is just a fluke.

Figure 7 shows the top five vectors in terms of records disclosed. The Hacking vector retains the top spot with over 383.3 million records (with 48% of records). In second place is the Drive/Media vector, responsible for over 155.8 million records (19%). The Web vector accounted for 93.4 million (12%), with the Fraud-SE vector at 54.8 million (7%) and the Laptop vector at 45.5 million (6%).

Figure 7: Top Five Breach Vectors for Records Disclosed (2005 - 2010)



Several of these vectors have high variability. The Hacking vector in particular shows significant fluctuation year after year. The Web vector had a spike in 2008 and has returned to its low trend after that; while the Laptop vector showed a spike in 2006 and has never reached that high a level since. The remaining vectors were consistently so low as to require an adjustment to the scale of the chart to display, and have been dropped for clarity.

Statistical Measures

To describe how the data is distributed, here are the standard statistical measures, and how they have changed from the prior report:

Table 1: Statistics on Number of Records Disclosed*

	2005 - 2010
Mean	327,629
Median	2,000
Min	1
Max	130,000,000
Standard Deviation	4,205,647

*For those incidents with finite numbers reported (2,461 total)

The mean has decreased from 387,926 to 327,629 with the additional data since the last report was published. The median has also decreased by 100 records. The min and max remain constant between the two reports, as we have not seen a new data breach that topped the Heartland Payment Systems incident. The standard deviation figure for TLV was 4,764,314, so there has been a slight decrease.

Table 2: Number of Records Disclosed/Year*

Year	Records Disclosed	Change from TLV
2005	68,555,563	=
2006	80,377,865	+14,807
2007	164,813,878	+64,465
2008	182,707,769	+293,008
2009	261,759,494	+35,912,130
2010	48,080,863	+48,080,863
Total	**806,295,432**	**+84,365,273**

*For those incidents with finite numbers reported (2,461 total)

As shown in Table 2 above, the total number of known breached records is over 806 million. If we assume 2,191 days between 1/1/2005 and 12/31/2010, we have an average loss of 368,003 records per day every day having been disclosed by these organizations. This has decreased from the daily average of 395,362 in the prior report. Likely, this is largely due to the very small number of known records disclosed in 2010. This represents a figure of over 15,000 people's data compromised per hour.

As with the prior report, the number of incidents where organizations do not put a finite number of records disclosed continues to be a challenge to accuracy of the data. Table 3 below shows the extent of the problem.

Table 3: Unknown Number of Records Disclosed per Year

	# Incidents	# Listing Unknown	Total % Unknown
2005	172	26	15%
2006	569	184	32%
2007	531	149	28%
2008	918	294	32%
2009	798	379	47%
2010	777	272	35%
Total	**3,765**	**1,304**	**35%**

With the new incidents from prior years, there are some small changes. For instance, 2007 changes from 27% to 28%. Adding to this, the 2010 data had a 35% unknown records factor as well. The effect of these changes on the overall total percentage of unknown records was to take the value from 34 to 35 percent. The increase in the percentage of incidents where organizations do not put a measurable figure on the scope of the incident is disturbing, and in the event of a federal law requiring data breach reporting, this should be a requirement. It would be worth researching whether these organizations make sufficient effort to notify the data subject victims. If they have, they should have a finite figure of how many people they notified and be able to include that in the report.

Table 4: Mean and Median Number of Records/Breach*

Year	Known # Records Disclosed	Mean Records/ Breach	Median Records/ Breach
2005	68,555,563	469,559	9,450
2006	80,377,865	208,774	3,235
2007	164,813,878	431,450	3,231
2008	182,707,769	292,801	1,000
2009	261,759,494	624,724	1,160
2010	48,080,863	95,210	1,700
Overall	806,295,432	327,629	2,000

*For those incidents with finite numbers reported (2,461 total)

Table 4 shows the mean and median number of records per incident. This takes into account only those incidents where the number of records disclosed is a known quantity. The median figure, however, is the more accurate of the two, since the number of records per breach varies so widely in this data set as it did in past years. The median record figure is no longer decreasing over time, potentially as more data points with finite numbers of known disclosed records are identified; the data is becoming more accurate. The increase in the number of records with lower record loss per incident helps to negate the impact of those sensational, media-favored, high record loss incidents. Many of these critical incidents were found through the previously mentioned FOIA requests, and would otherwise have remained "unknown" [11].

The changes in the mean and median between the prior report and this one, apart from the number of records disclosed which has already been discussed, is in the following table:

Table 5: Change in Mean/Median Records/Breach Figures between TLV and TLV2011*

Year	Change in Mean TLV to TLV2011	Change in Median TLV to TLV2011
2005	66,291	0
2006	66,286	107
2007	110,926	(19)
2008	83,370	(151)
2009	296,457	84
Overall	(60,297)	(100)

*For those incidents with finite numbers reported (2,461 total)

As you can see, the overall mean figure decreased by over 60,000 records, and the overall median figure decreased by 100. This trend is expected to continue as additional incidents are added to the dataset over time and accuracy is increased. The addition of incidents with finite records of all sizes will help to balance out the weight of the large incidents versus the small.

Since 2010 has no data in the previous report, it has not been included here. Clearly the difference would be equal to the mean/median figures in Table 4 for that year.

As before, to get a granular estimate of the scope of the underreporting, the median records disclosed figure was calculated on a yearly basis per breach vector. This allows for a more accurate estimate of the total records disclosed if the records marked as "unknown" are calculated using these median figures, since we know the year and vector for each incident. This is the formula used:

known records disclosed per vector and year	+	median records per breach vector per year	X	number of incidents where records lost is "unknown"

While this is an estimate, it is the best data we have given the high percentage of uncertainty in self-reporting the number of records disclosed. Using the same example from the original TLV report—with the updated records disclosed figures—the Drive/Media vector in 2008 was responsible for the known disclosure of 35,026,807 records. The 2008 median figure per incident remained constant at 3,000, as did the 11 incidents listed with the number of records disclosed as unknown. To get the total of how many records were disclosed including the estimate, we use the median figure for those "unknown" values:

35,026,807		3,000		11		35,059,807
known records	+	median records	X	number of incidents	=	New Records Disclosed Total

Table 6 shows the median records per breach for each of the vectors between 2005 and 2010. This is a critical set of figures, as it allows us to calculate with increased accuracy an estimate of how many records may be hidden in the unknown figures, and provides a good basis for cost calculations based on actual median figures. These can then be used for risk calculations and return on investment decisions.

Table 6: Median Records Per Breach Type Per Year

	Median Records Per Breach Type Per Year						
	2005	**2006**	**2007**	**2008**	**2009**	**2010**	**Overall**
Computer	16,000	7,500	3,500	1,318	1,000	2,000	3,100
Disclosure	-	-	-	-	-	1,438	1,438
Documents	208	650	350	100	140	716	287
Drive/Media	130,000	3,500	7,500	3,000	4,450	3,793	3,900
Email	584	561	367	250	533	1,300	500
Fax	-	4	-	-	-	-	4
Fraud - SE	150,000	1,000	1,400	193	205	135	279
Hack	8,800	14,500	10,500	5,000	6,600	4,585	8,029
Laptop	15,000	3,020	3,000	2,500	2,000	2,949	2,950
Snail Mail	597	6,000	1,927	553	859	3,800	1,870
Tape	403,000	45,000	120,000	20,400	100,000	13,000	47,000
Unknown	4,103,000	4,327	4,143	214	1,000	300	1,000
Virus	-	-	-	91	1,000	900	900
Web	2,800	1,300	1,618	1,000	900	1,672	1,300

In TLV, an estimated 7.5 million additional records were exposed over what was previously reported, bringing that estimated count to over 729.5 million. Between the 2010 incidents and the additional records for the prior years, we have over 806 million known records disclosed. The changes to the estimated additional records are shown below in Table 7.

Table 7: Estimate of Records Disclosed

Year	# of Unknown Incidents	Known Records Disclosed	Estimated Additional Records Disclosed	Estimated Records Disclosed Totals
2005	26	68,555,563	885,432	69,440,995
2006	184	80,377,865	941,094	81,318,959
2007	149	164,813,878	1,391,386	166,205,264
2008	294	182,707,769	701,778	183,409,547
2009	379	261,759,494	1,283,854	263,043,348
2010	272	48,080,863	585,150	48,666,013
Total	**1,304**	**806,295,432**	**5,788,693**	**812,084,125**

As you can see, the new estimated total comes to over 812 million records in our arguably conservative estimate. The likelihood that a larger breach is hiding in these unknown records is always a concern when we see initial reports of what turn out to be very large breach incidents listing “unknown” figures. This has been the case for many of the largest incidents in the database and is in part due to the organization trying to get a good estimate before releasing it to the public. However, as time goes by if no corrected totals are given, speculation on the organizations that do not put a number on the breach becomes common.

In Sex, Lies and Cyber-crime Surveys, Florencio & Henley performed a meta-study on data breach reports (not including TLV). The researchers noted that the minority of large incidents tend to dominate the studies in the upper tail, particularly when the sample size is limited [6]. This is unavoidable when early in a research topic, we simply don't have sufficient data to get the order of accuracy that statistical analysts expect. These laws have been on the books less than ten years, so the number of data points is necessarily too small for generalization of the data. Thus, when interpreting this data, caution should be used, with the understanding that it may be a very long time before sufficient data exists to make truly accurate estimates. In the meantime, these reports provide trending information otherwise unavailable, and remain valuable for risk calculations.

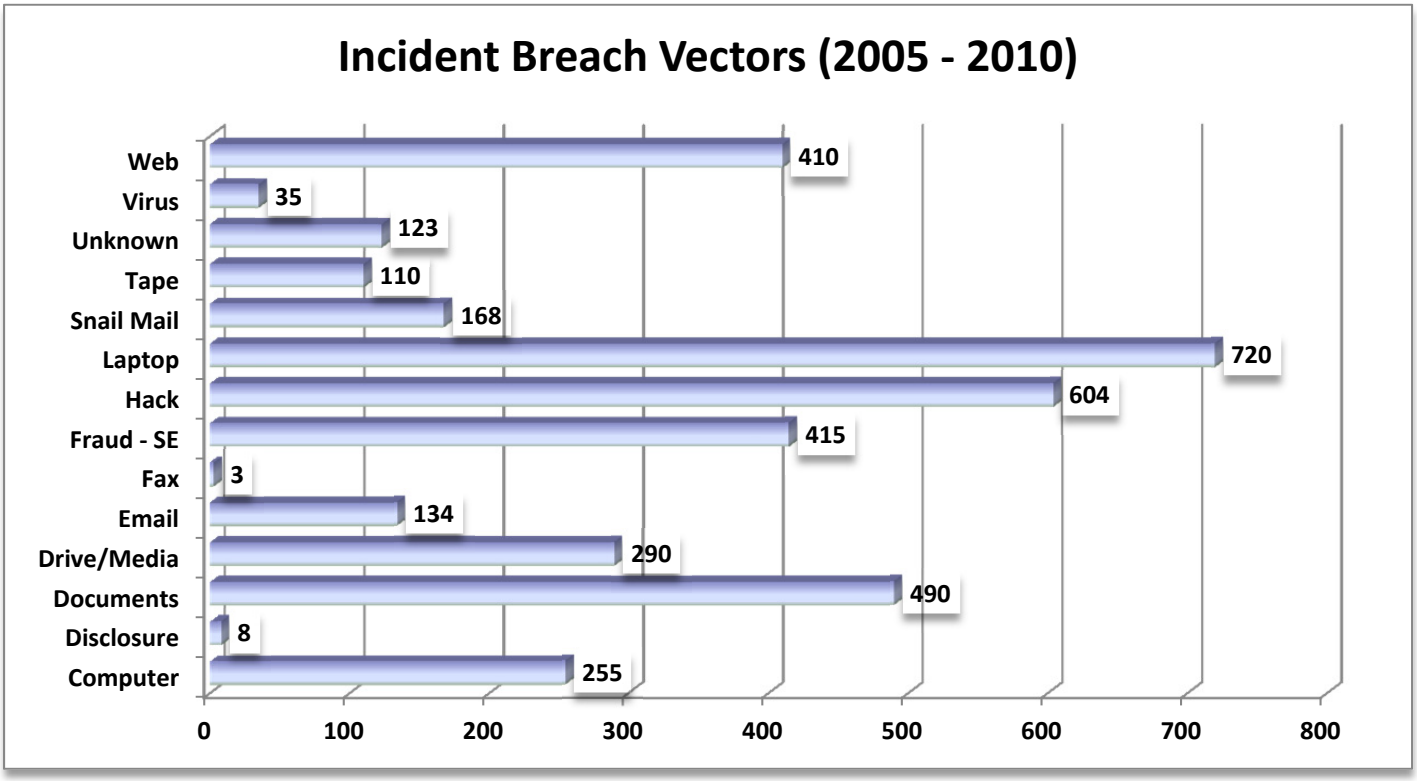
Breach Vectors

There has been a rise in snooping and other inappropriate disclosure where the confidentiality of the data is breached, but the data may not have left the control of the organization; or the act was done with the approval of the organization, but found later to be an inappropriate breach of confidentiality. In a recent case, UCLA Medical Center agreed to pay $865,000 to settle instances where employees snooped on the medical records of celebrities being treated at the facility. Another example is when the California Department of Health Care Services released confidential and identifying information about HIV positive MediCal recipients to a third party service provider. This was later deemed to be both illegal and unauthorized. To classify these types of cases, the new breach vector of Disclosure has been added to the study beginning with 2011.

Overall Breach Vectors (2005 – 2010)

The hacking vector is again the loss leader in the study with over 383 million records known. While it was not the top vector for incidents, it did come in second. The incident leader was Laptop, which came in fifth in records disclosed.

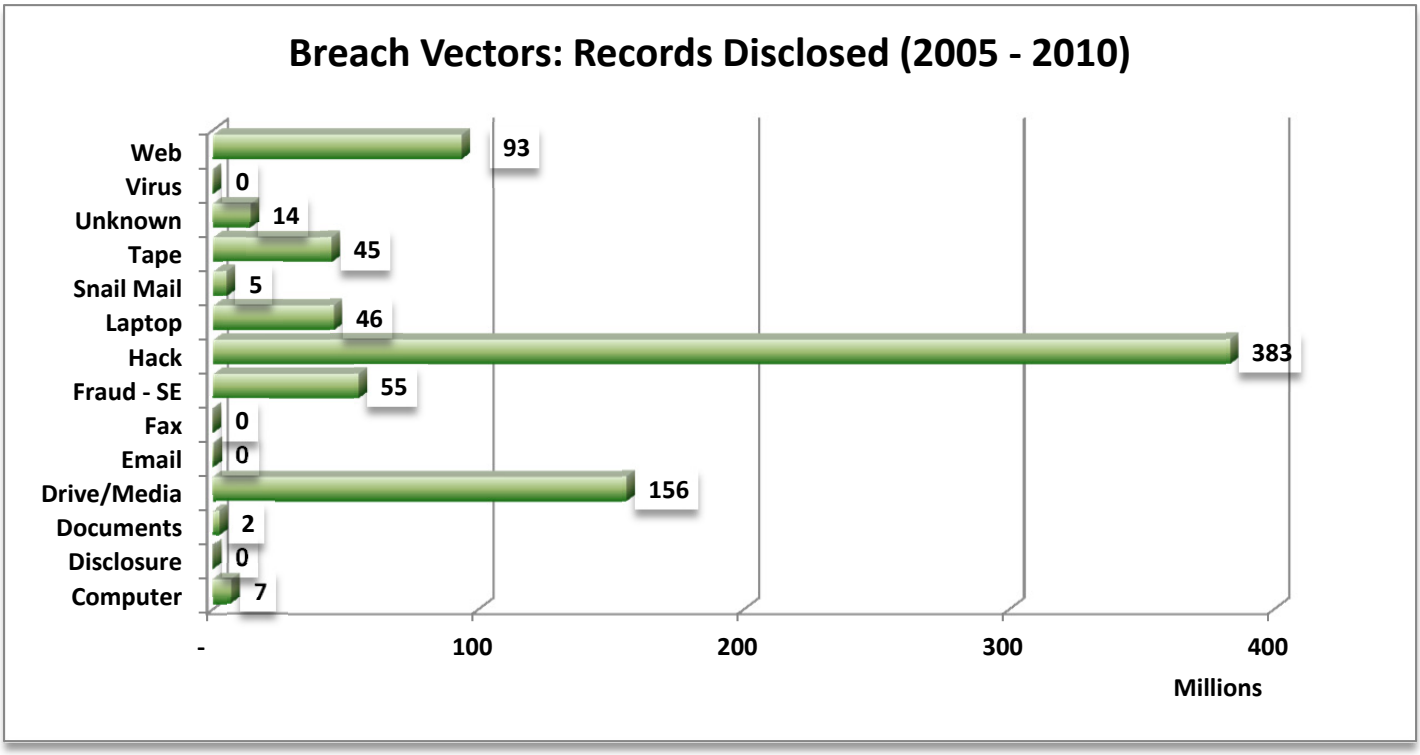
Figure 8: Breach Vectors - Incidents (2005 - 2010)



The Laptop vector retains the leadership spot in the incident breach vectors ranking—in TLV the total was 589 incidents, compared to the 720 we see now. Hacking was second, growing from 456 to 604, and the Documents vector has

overtaken the Web vector for third place. While Web grew from 325 to 410, Documents grew from 316 to 490—a significantly higher growth rate.

Figure 9: Breach Vectors - Records Disclosed (2005 - 2010)



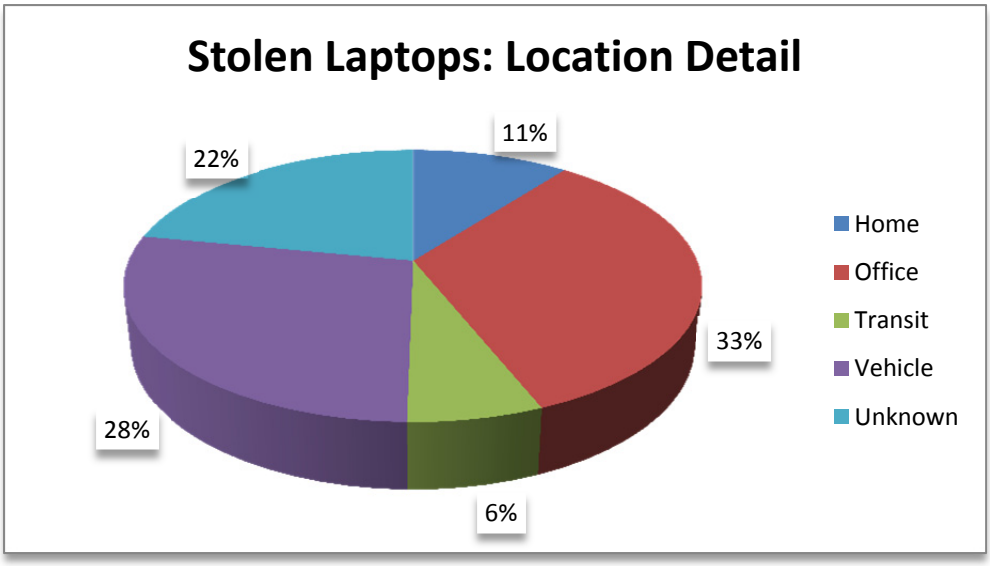
The Records disclosed retained the same rankings in the prior report, only the number of disclosed records has increased. For Hack, the records increased from 327 million to 383 million. For Drive/Media, the previous total was 149 million; and for Web, the prior total was 84 million.

The Laptop Vector

Laptops increasingly contain significant amounts of organizational data. They are frequently the sole computer employee's use, and come with a hard drive that can contain very large datasets. It is not uncommon for companies to find out after a breach incident that the individual assigned the asset had spreadsheets, and even whole databases containing sensitive data. When a laptop is issued to an individual, it should be accompanied by a set of rules for the custodian of the device to follow. This should include direction for maintaining physical control offsite (i.e., not to leave it in a vehicle, etc.) and onsite (i.e., lock it to their work surface), as well as controls for when these rules either are insufficient to keep the asset safe, or when the individual does not follow them. Potential controls include encrypting the device, remote wiping capability, tracking/recovery software, etc. The organization has a responsibility to the data subjects to take appropriate steps to ensure their data will not be at risk of disclosure when the unexpected happens.

Of the 3,765 incidents in the study, 719 involved laptops being improperly disposed of, getting stolen, or being lost. In 96% of these incidents, the laptops were stolen. Overall, the laptop vector accounted for 45,500,147 records in the study. Further, Figure 10 details where these devices are stolen from. The largest quantity of laptops were stolen from the office of the organization suffering the loss. This illustrates the need for locking mechanisms for the laptops when unattended at work. The second largest number of laptops were stolen from inside a vehicle. This is the most preventable, and represents 191 incidents over 4 million records.

Figure 10: Stolen Laptops - Location Detail (2005 - 2010)



Information Security professionals who can influence the contents of their organization's awareness training should be lobbying to have something included about laptops and vehicles. Ideally, laptops would never be left unattended. However, if they must be in a vehicle, they should be placed in the trunk at the start of the trip, not when the individual reaches their destination. Otherwise, they are sending up a red flag to any criminals who are looking for an easy score that is where a valuable is being stored. Even with this precaution, the vehicle may be broken into—in no case should the laptop be left overnight in the vehicle—particularly at the employee's residence. If they habitually leave their laptop in their vehicle overnight, it is only a matter of time before they are observed and the device stolen. Keep in mind, the laptop may be the vector for the initial breach, but there may also be other materials kept with the laptop that will put the organization at further risk. These include remote access security devices (tokens, etc.), paper notes with pin codes or passwords, or even sensitive data on printouts the employee brought with them.

A common thread in the data breach notification letters is that the laptops are password protected. This does not provide any real protection against accessing the data, and it is trivial to bypass this control with minimum effort or expertise. In fact, a quick Google search will yield numerous easy to use tools for the would-be cracker. In a few short minutes, the data is accessible, typically the time necessary to boot from a USB key or CD/DVD drive and the control is rendered ineffective. There are even YouTube videos on the step by step use of such tools [3]. Encrypting the data negates this attack and is an effective control when implemented properly.

The Hacking Vector

The 2010 data increasingly showed the prevalence of skimmer use. Skimmers are credit card readers that are typically hand held or installed in ATMs and point of sale devices to read the credit card track data and steal it. This was most commonly seen in retail establishments, and especially in restaurants. Anywhere the credit card is taken away from the customer's control; there is a higher risk that a skimmer might be used by the dishonest. However, this is not to say that the card data is safe when in the control of the customer. Another increasingly common incident is the skimmer installed inside the gas pump. In this case, there is either a skimmer on the outside of the pump (these are becoming very clever and difficult to spot), or there is a device inside the pump where the customer has no hope of detecting it, and it can be wirelessly unloaded by the criminals, posing minimal risk of being caught [2].

Figure 11 shows an example of a skimmer on an ATM machine. It looks like it belongs, and would fool most customers.

Figure 11: Skimmer on ATM Machine [19]

The portion with the dotted lines around it is the skimmer, placed over the top of the actual reader in this ATM. There is nothing immediately suspicious about this machine's appearance, and this is why these are so successful. The skimmer is frequently paired with a camera to capture the associated PINs as the cards are read.

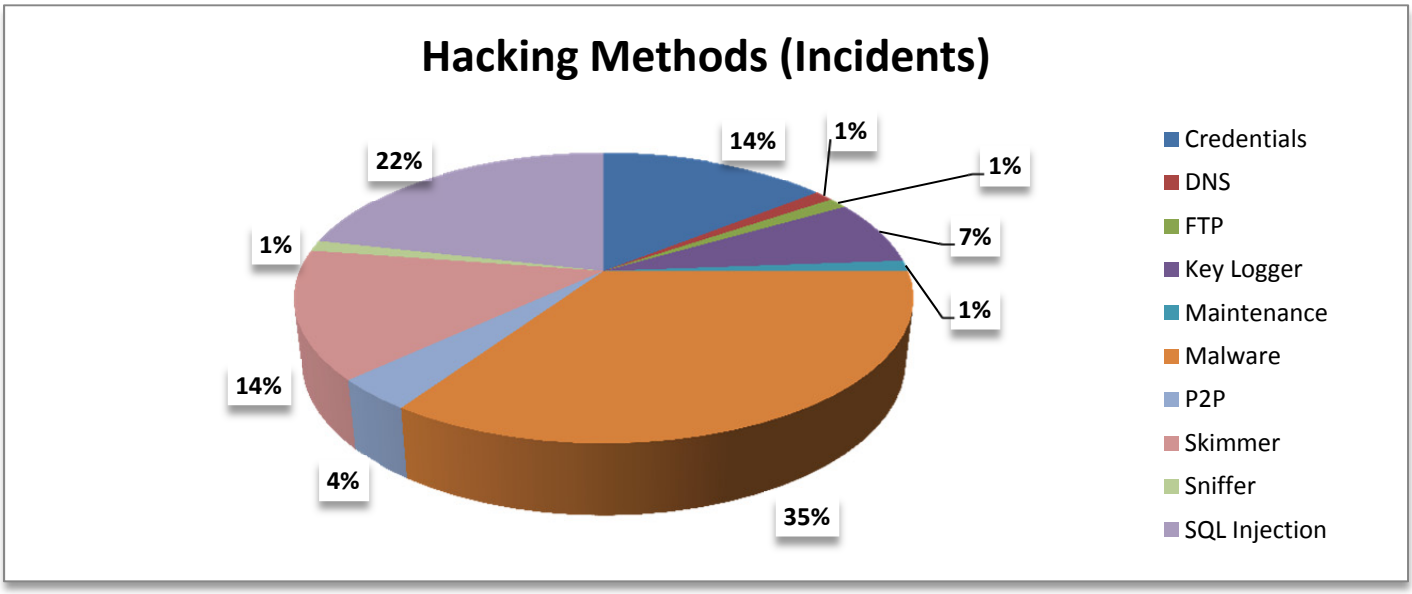
> Advances in data exfiltration techniques have included the use of Bluetooth technology within the skimmer to allow for wireless retrieval within a finite proximity. This, of course, reduces the risk of apprehension when attempting to retrieve the device, which may occur if the skimmer is discovered. Additionally, it allows the possibility of collecting data at various intervals, so if a device is removed by a bank employee or law enforcement not all of the captured data is lost. The latest evolution in data retrieval is the use of technology, again embedded in the skimmer, that utilizes GSM standards and will text captured data in real-time to the criminal's cell phone [2].

Organizations should either put controls in place that notify when a device is tampered with, or have regular inspections of their point of sale devices, gas pumps and ATMs to mitigate this risk [2]. Physical security controls should be in place so that the defense in depth principle is followed. If gas pumps use a common key (i.e., not using unique keys per location or per pump), this should be changed to make access to the pump's internals both more difficult and tamper evident.

The vast majority of the incident reports in the Hacking vector do not give details as to the method used to gain access. For those few that do give us this information (84 incidents), Figure 12 shows the methods used. Malware was the primary method of choice, with SQL Injection in second place. Using default/known credentials and the use of skimmers tied for third place. Malware is difficult to combat, as it can be a multi-layered attack. Many of these incidents begin in the form of an email—targeted to the recipient or not—enticing the person to click on a link, open a file, or perform some other action that will allow the malware to get a foothold on their computer. From there, the system is used as a springboard to other

systems in the organization. However, it doesn't always take this form—in one study, dropping malware infected USB sticks in parking lots or near the entrances of a company was an effective vector [5]. Employees pick up the devices and plug them into their computers—whether to see what is on there to identify the person who lost it, or just because they think they scored a free device. Either way, the controls are frequently not in place to stop the malware from being automatically run when the device is plugged into the system [15].

Figure 12: Hacking Methods Detail - Incidents (2005 - 2010)



SQL Injection remains a popular method for gaining access to database driven web applications and other online database engines. Mitigating controls include encrypting the data both at rest in the database, and in transit. There are also specialized firewall applications geared specifically towards mitigating this type of attack. Programming controls can also be effective, and illustrate the need for developer-specific security awareness training. Attention should also be given to the use of production data in test and development environments, since those environments typically have less stringent security controls in place.

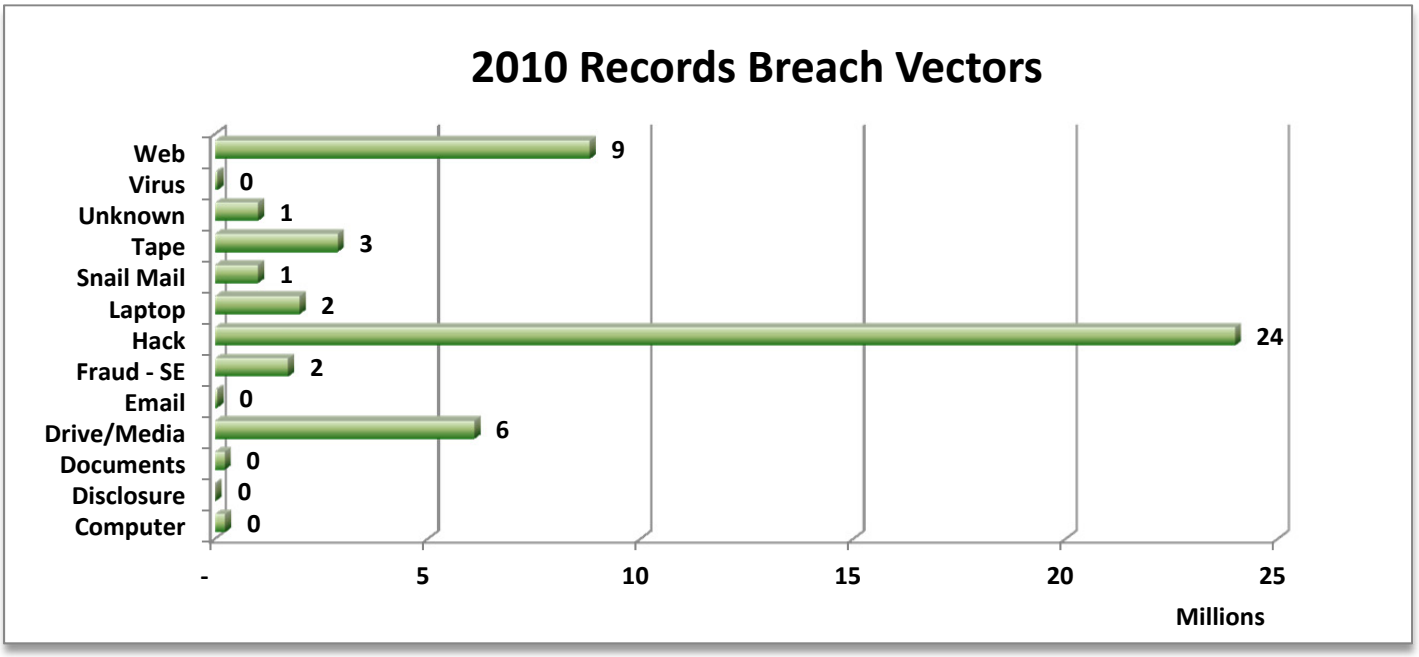
> Over 83 percent of companies use, real (live) customer or employee information in development and testing, and 51 percent of these companies admit they do not take appropriate steps to protect real data used in development and testing such as anonymization of data, masking, subsetting or other methods [14].

The rank of the methods used has not changed from TLV to TLV2011, at least for the top two. Skimmers came in third place in TLV, and share third place in TLV2011. The compromising of credentials was in fourth place in the previous report, and has gained several incidents to tie with Skimmers for third place.

2010 Breach Vectors

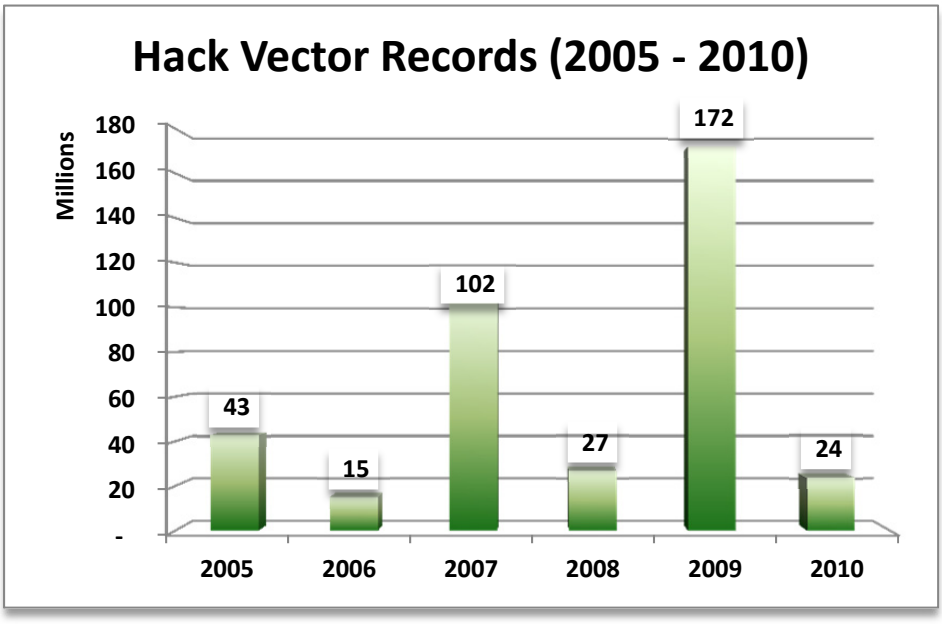
The Hacking vector is the loss leader in in 2010, with 24 million records. In contrast, the leader for incidents (Laptop) came in with only 2 million records.

Figure 13: 2010 Records Breach Vectors



The Hack vector has decreased from 172 million records in 2009 to only 24 million in 2010. While this seems like a huge decrease, keep in mind that all but two of the vectors show significant decreases from the prior year. The Hack vector fluctuates quite a lot from year to year as shown here:

Figure 14: Hack Vector Records Detail (2005 - 2010)



Given the variability in this vector and the overall drop in records for 2010, the current decrease is not as surprising. Since we are dealing with records disclosed, and many of the largest incidents in the database involve the Hack vector, it only take a few large incidents to drive the yearly total up to the heights we see in 2009. In fact, the 2009 total is that high primarily due to the Heartland Payment Systems breach, which accounted for 130 million records, and the RockYou breach, which accounted for 32 million records. You will see these two incidents later in the Large Incidents section. Since the Hack vector is the records leader for the study, it will show the most volatility when these types of large Hacking breaches are reported.

The most important recommendation for this vector is **know where your data is**. If you do not know where it enters the organization, where it is transformed, stored, shared with outside parties, archived, and finally disposed of—you cannot hope to keep it secure. This can be an enormous job for a large complex enterprise, but you must start somewhere. You may want to break this down by data type—customer data versus employee data for example. Where does your employee data originate? Usually this would be in a Human Resources department. Trace each data type from when it is created to when it is disposed of and all the places it is used in between. You will find places this data resides that you didn't know about, and processes that are not as secure as you may have thought. Without making these types of data flow maps, organizations are operating on only a portion of the risk picture.

Insider, Outsiders and Third Party Partners

In this section, we examine the actor responsible for the breach, and in some cases, their motivation. Insiders are employees and other individuals who have approved access to the organization's systems. Outsiders are individuals who do not have authorized access and are assumed to be malicious, given that they are deliberately accessing systems without permission. Third Parties are those companies where some form of access to the data has been granted (typically on a contractual basis) to facilitate some business process.

Organizations must "bake" security controls into contracts with third party partners. This means the Information Security personnel should be involved early in the selection and vetting of potential business partners where sensitive data is concerned. When the Business partners with Information Security, risk is reduced.

> Awareness of suppliers' and other business partners' security practices—including understanding the country's data protection regulations under which the organization operates and strictly monitoring how and when their data is used by providers and where such data is sent—is critical to verify proper practices are in place to protect sensitive data. Organizations also should ensure that provider's, as well as their own responsibility and accountability, are clearly understood [13].

Figure 15: Incidents by Actor (2005 - 2010)

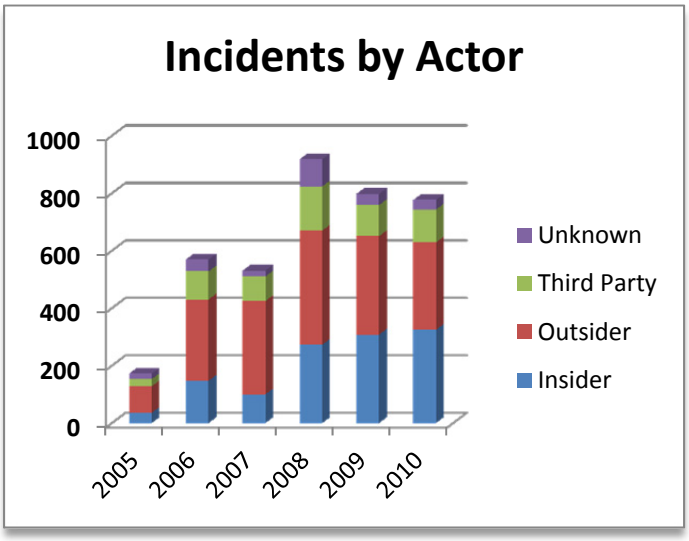
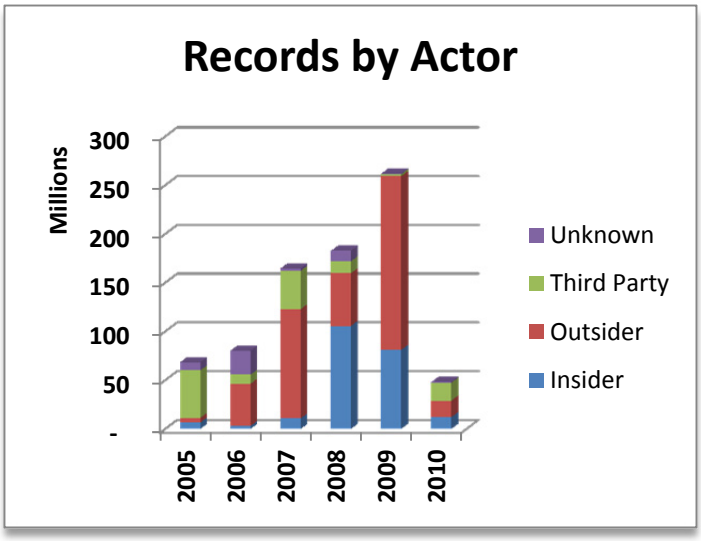


Figure 16: Records Disclosed by Actor (2005 - 2010)



As we saw in the previous report, the majority of the incidents are caused by Outsiders. Organizations should note, however, the increasing number of incidents caused by those with approved access to the organization's data. This trend has continued for the past several years and should be addressed with controls designed to prevent (in an ideal world) the insider from causing an incident, and at a minimum, detect when such an event has occurred.

The records lost are dominated by the Outsider actor, showing that when they get a foothold in an organization's network and systems, the damage can be much greater than when the actor is on the inside. Again, while prevention is ideal, detection is a must to minimize the damage of an incident, regardless of actor.

Figure 17: Insider Intent Detail (2005 - 2010)

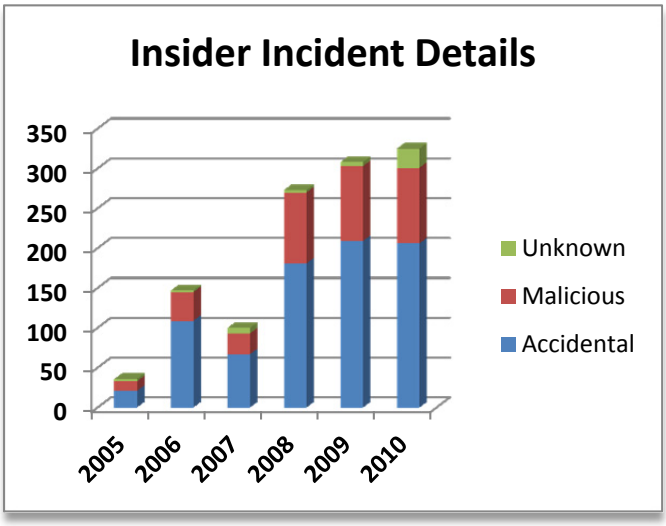
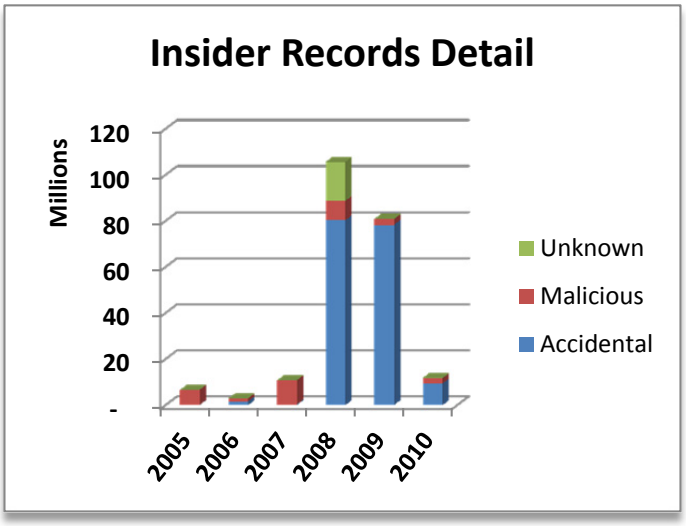


Figure 18: Insider Intent Records Detail (2005 - 2010)



The incidents of malicious insiders are increasing over the past several years. However, where an insider is involved, the motivation is much more likely to be accidental. While the damage is not intentional, the harm is significantly higher, as shown in Figure 18. Monitoring controls designed to catch the activity quickly will help ensure the damage is minimized.

Two vectors tied for first place for Insider/Accidental incidents. Both the Documents and Web vectors showed 241 incidents. It should be noted, however, that Laptop thefts are classified as Outsider incidents, given that the threat actor is not the person who the laptop was assigned to within the organization. With 720 laptop incidents, these remain a vector where insiders share some responsibility, given that they are so frequently left unattended and unprotected.

Table 8: Insiders, Outsiders & Third Party Partners Records Detail

	2005 - 2010
Insider	219,162,928
Outsider	409,465,576
Third Party	129,574,021
Unknown	48,092,907
Totals	806,295,432

Table 8 shows the damage in terms of records lost by actor. The Outsiders are responsible for almost twice the record loss of Insiders, and the disparity is even higher between them and the Third Party partners.

Table 9: Insiders Records Detail

Insider Intent	2005 - 2010
Accidental	170,040,878
Malicious	31,544,486
Unknown	17,577,564
Totals	219,162,928

Table 9 shows the Insider intent breakdown. When the incident is a result of an accidental action, the damage is over 3 times the size. Some recommended control measures include a change control process to have others vet the planned changes won't cause unintended consequences; code review for developers; and detective controls to catch these issues early before they become a huge problem.

Figure 19: Median Known Records Per Actor (2005 - 2010)

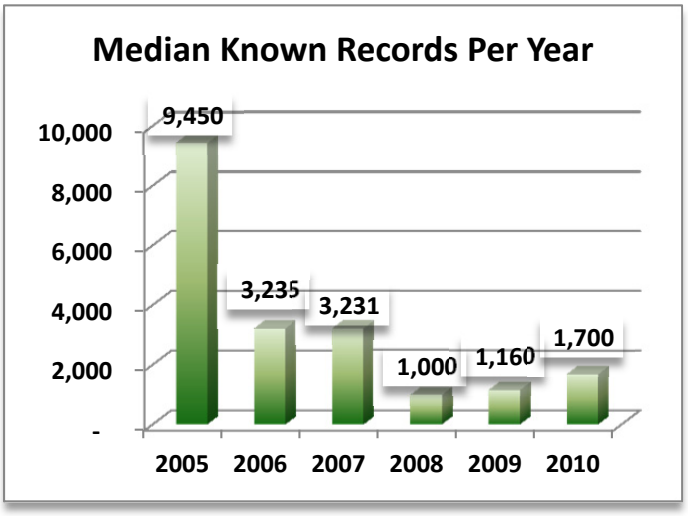
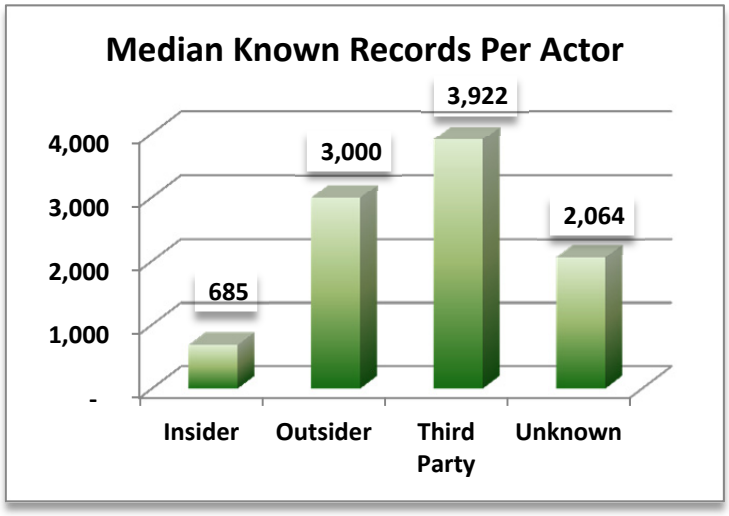


Figure 20: Median Known Records Per Actor (2005 - 2010)



The median known records figure has shown a slight increase between 2009 and 2010. There has also been some slight changes year over year with the addition of incidents from past years between TLV and TLV2011. The changes are not very large, and over time, the expectation is to see the level of accuracy increase with more incidents in the study. Only those incidents where a known number of records disclosed were included in this total to avoid skewing the figures inappropriately. However, given the level of uncertainly within the actual records disclosed, the data set can only get more accurate to a point. True precision is unattainable with the volume of unknown values in the records disclosed variable.

The median figures per actor have also seen some change. The table below shows the TLV and TLV2011 figures for reference.

Table 10: Median Records Change (TLV - TLV2011)

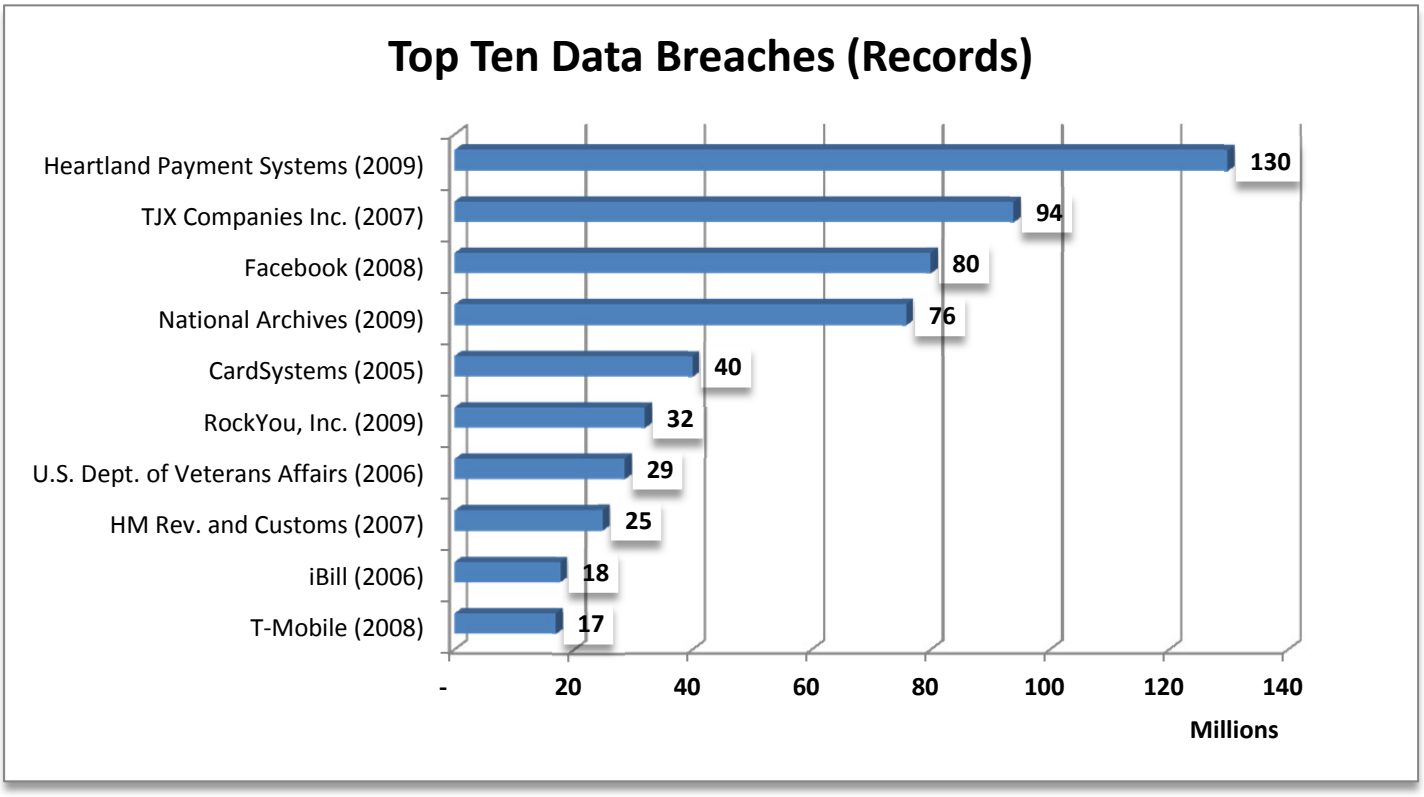
Actor	TLV	TLV2011
Insider	575	685
Outsider	3,450	3,000
Third Party	6,000	3,922
Unknown	2,000	2,064

The most significant change is in the Third Party figure. It went from 6,000 to 3,922, which is the largest decrease of all the actors. Still, the number of incidents listing “unknown” for their number of records disclosed figure is 36% over the course of the study for Third Party incidents. Only the Insider incident had a higher level of uncertainty, with 41% of the incidents not providing a finite value for number of records exposed. Outsider and Unknown are 31% and 27% respectively.

The Large Incidents (Involving over 1 Million Records)

Only 66 of 3765 incidents involved over 1 million records. However, those 2% of incidents made up 91% of the records disclosed over the study. The top vector for large incidents was the Hack vector, claiming 29% of the incidents. The Drive/Media vector took 22% of the incidents, with the Fraud – SE vector accounting for 17%.

Figure 21: Ten Largest Breaches (2005 - 2010)



As you can see from Figure 21 above, 2010 did not have a data breach incident with sufficient known records disclosed to make the top ten. In fact, the largest listed in 2010 was the DeviantArt (SilverPop) breach at 13 million records, which puts it at the eleventh largest breach in the study. Without several large incidents in the year, we saw a significant drop in the disclosed figure for 2010.

In terms of the vectors that drove these breaches, here is how they break down:

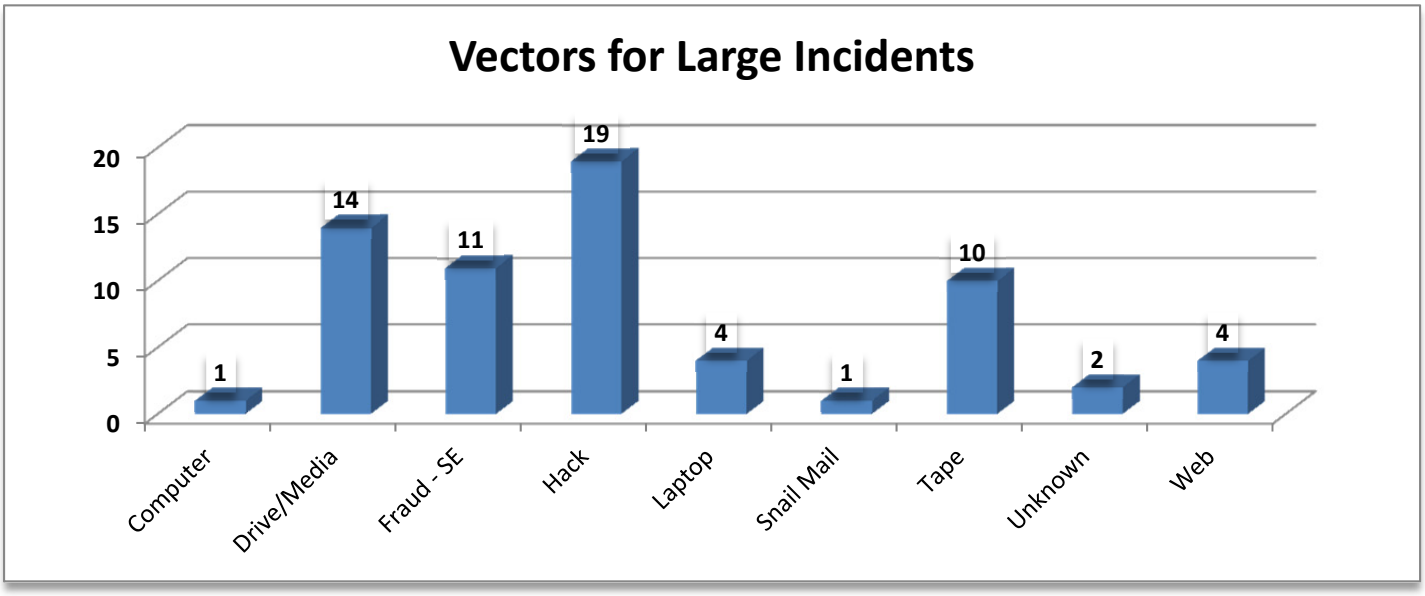
Table 11: Breach Vectors of the Ten Largest Incidents (2005 – 2010)

Organization	Records	Vector
Heartland Payment Systems	130,000,000	Hack
TJX Companies	94,000,000	Hack
Facebook	80,000,000	Web
National Archives	76,000,000	Drive/Media
Card Systems	40,000,000	Hack
RockYou, Inc.	32,000,000	Hack
U.S. Dept. of Veterans Affairs	28,600,000	Laptop
H.M. Revenue and Customs	25,000,000	Drive/Media
iBill	17,781,462	Fraud-SE
TMobile	17,000,000	Drive/Media

As you can see, in four of the top breaches in the study, the Hack vector is the culprit. The Drive/Media is in second place, and underlines the risk of portable media when controls are not in place to ensure that data placed on them is both authorized and protected against loss of the device.

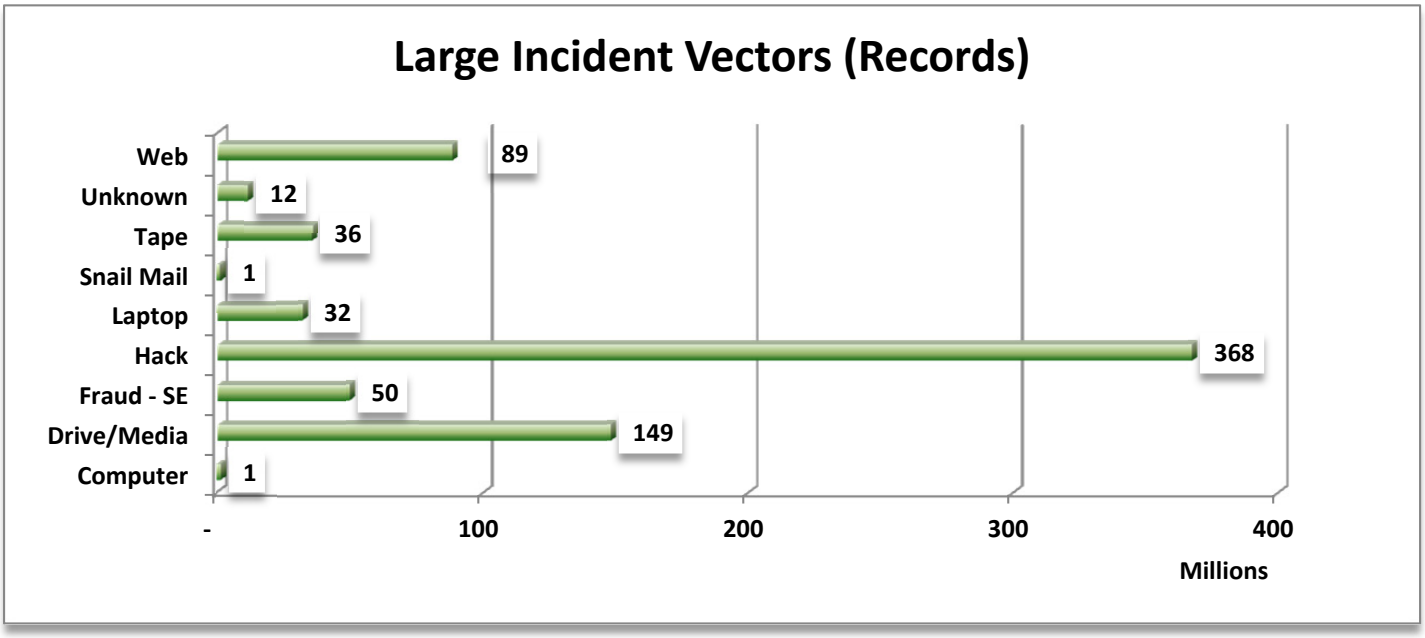
Looking closer at these 66 largest incidents, the breach vectors are shown below in Figure 22. The Hack vector was the lead, with Drive/Media and Fraud-SE coming in second and third for all the largest events.

Figure 22: Incident Breach Vectors for Large Incidents (2005 - 2010)



To get a better appreciation for the impact of the size of the incidents, the same vectors in terms of the records disclosed follows in Figure 23. Between these two charts, you can easily see that the Hacking vector not only is the leader in large incident occurrences, but it also accounts for the largest amount of records disclosed. In fact, the median records disclosed for the Hack vector among these large incidents is 5,379,909. The large incident Hack vector min was 1.2 million and the max was 130 million.

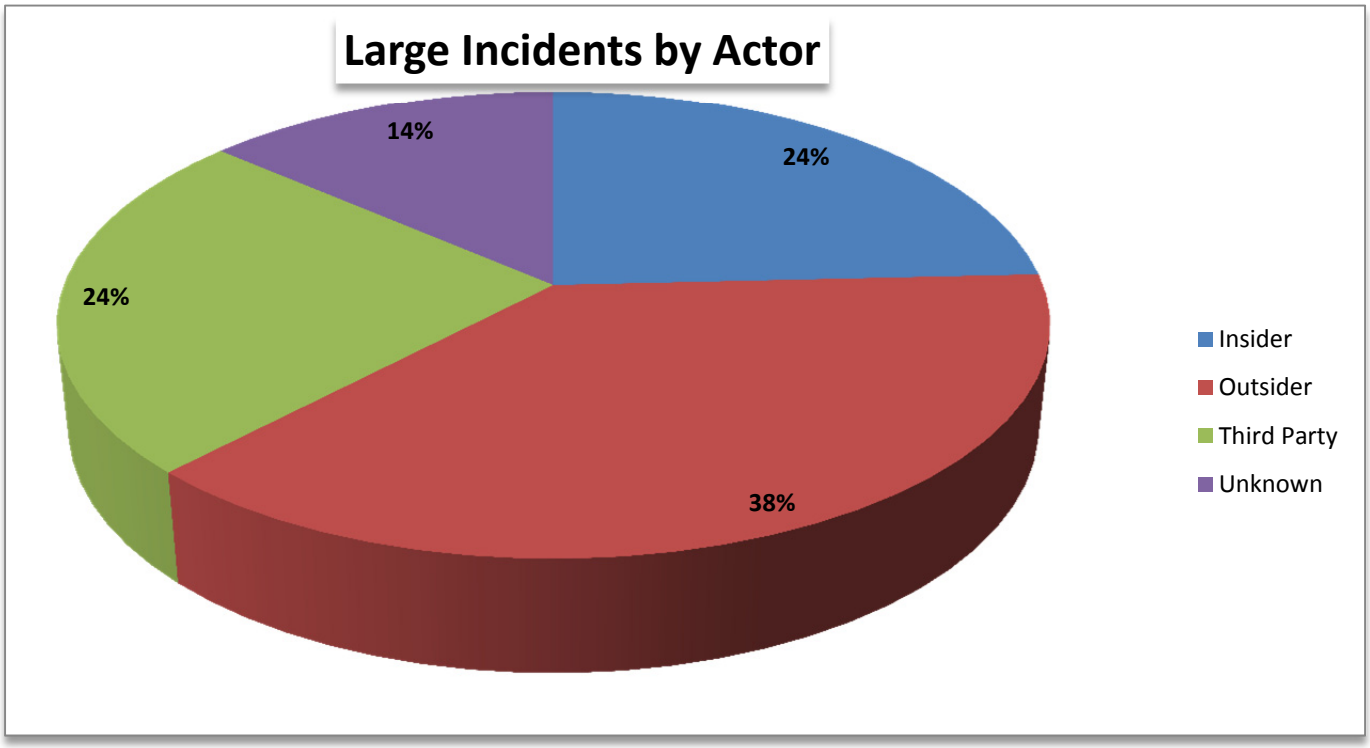
Figure 23: Records Breach Vectors for Large Incidents (2005 - 2010)



In thirteen (20%) of these large incidents, there is confirmed criminal use of the data, accounting for a total of over 203 million (28%) records. The overall confirmed use rate for the study was only 10.5% in contrast.

The majority of the large incidents were caused by Outsiders as shown below in Figure 24. This is expected with the Hacking vector being dominant.

Figure 24: Largest Incidents by Actor (2005 - 2010)



Criminal Use

Criminal or malicious motivation in attacks makes for more expensive breaches [12]. This is true both for the organizations who suffer them, and the people whose data is compromised.

Between 2005 and 2010, in 396 cases were confirmed to have been used for criminal activity. This is a difficult metric to track; since the criminal activity associated with breach activity shows that the data is commonly sold and resold [2]. The crime where the perpetrator has a direct connection to the victim is most frequently where the arrest is reported with the event. To that end, the Fraud-SE category is represented by a much higher margin than some of the vectors that have generated these large scale data disclosures.

Figure 25: Incidents of Confirmed Criminal Use of Data (2005 - 2010)

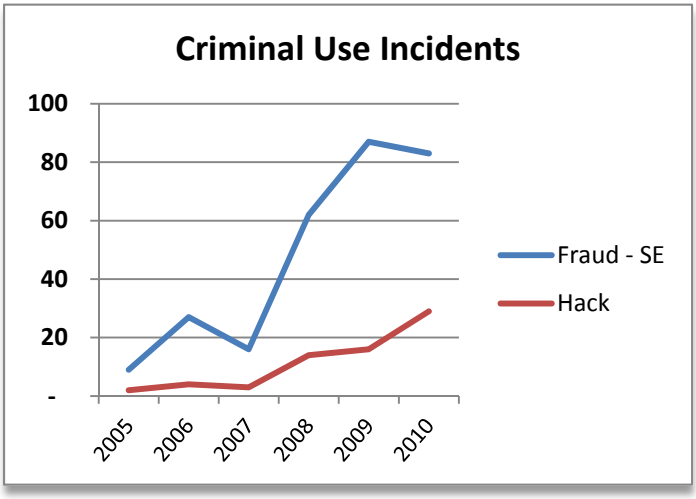
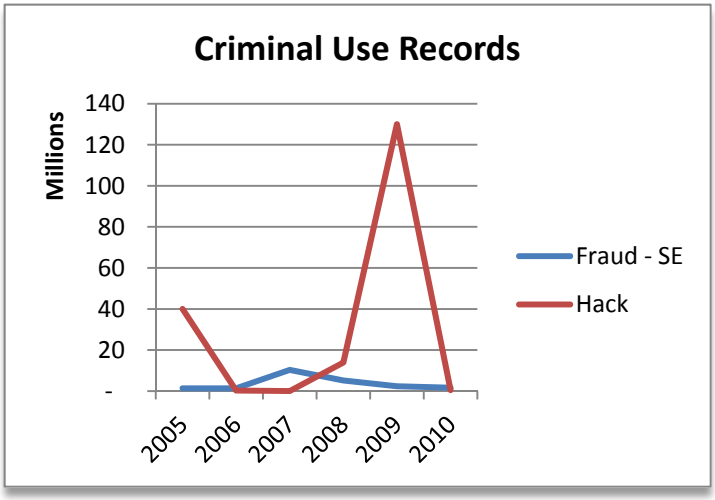


Figure 26: Records Disclosed with Confirmed Criminal Use of Data (2005 - 2010)



As you can see in Figures 25 and 26, the top two vectors where criminal use has been established remain Fraud-SE (the leader) and Hack (in second place), consistent with the TLV data. Other vectors where data was confirmed, but not consistent across the study, included Documents, Web, and Unknown. The records lost were similarly consistent, as would be expected.

While the Fraud-SE vector leads in incidents, clearly the Hack vector leads in records disclosed. In hacking incidents where criminal activity was confirmed, the use was most frequently put to credit card fraud. Reports of fraudulent charges on existing accounts, and of leveraging the data to open new accounts were also reported. The 130 million record spike in the Hack vector represents the Heartland Payment Systems breach.

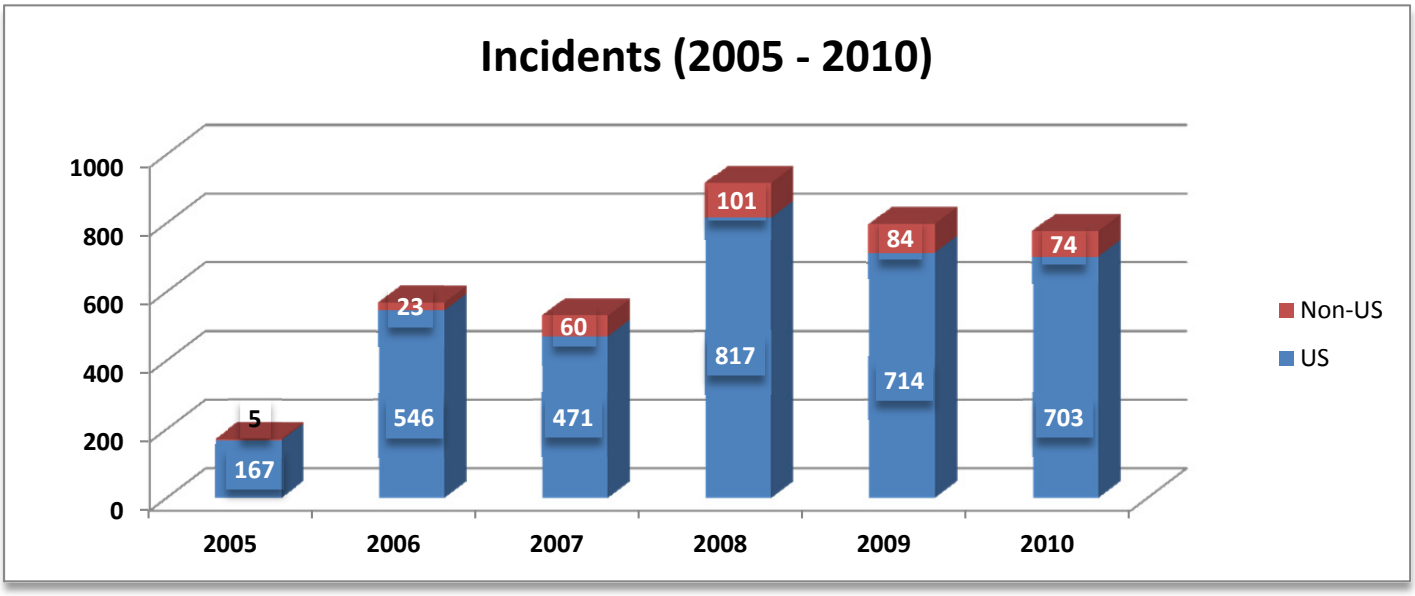
Geographic View

For organizations that are doing business across international boundaries, keeping up with the United State's laws is just the tip of the iceberg.

> Making matters worse is the fact that there are no common or consistent standards for dealing with data privacy and protection from country to country or even within individual countries. For example, in the United States alone, there are 49 different state laws that regulate notification of security breaches, as well as separate laws that govern the use of various types of data (such as financial and health data) [13].

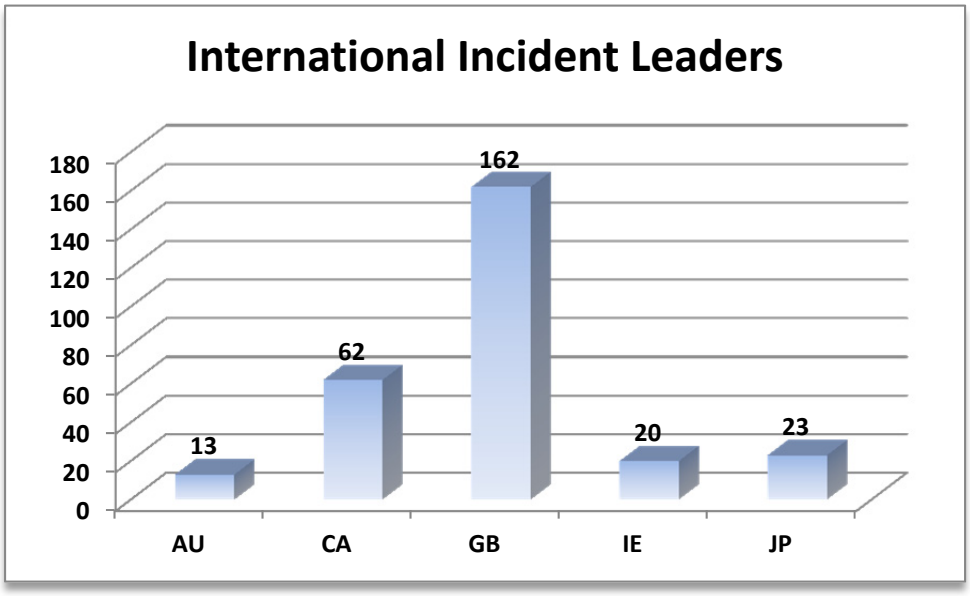
The United States again accounted for the vast majority of incidents (91%) and the records disclosed (88%). As with TLV, this is in part due to the focus of the reporting sources. It is also partially due to the lack of privacy breach laws in other countries. These incidents are likely occurring in every country, but we lack the visibility the breach reporting laws provide.

Figure 27: U.S. and International Incidents (2005 - 2010)



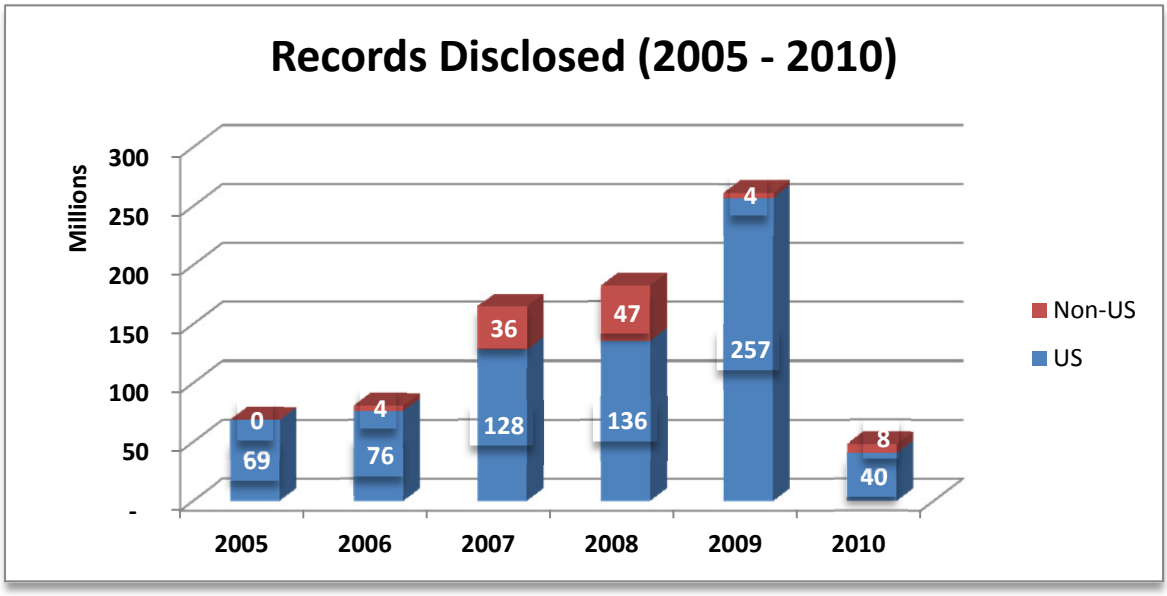
There were 34 countries reporting incidents in the study. The figure below indicates the five countries with the most incidents over the six year time period. As you can see, Great Britain is the leader by a significant margin, with Canada in second place. The remaining countries in the study not listed here were all in the single digit range for number of incidents.

Figure 28: Top Five Countries for Incidents (Non-U.S.) - (2005 - 2010)



The United States also dominated the record loss figures, as expected. As a percentage of loss, however, 2010 represents a change—primarily due to the sharp reduction in record loss reported in the United States rather than any real change in the number of records lost internationally.

Figure 29: U.S. and International Records Disclosed (2005 - 2010)



Overall, international reports disclose a finite figure for the number of records exposed much more frequently than do United States organizations. The rate of unknown records for domestic incidents is 36% while the same rate for international incidents is only 18% over the course of the study. The highest yearly percentage that international incidents reached were 30% unknown records in 2006. In contrast, the highest that domestic incidents reached was 51% in 2009.

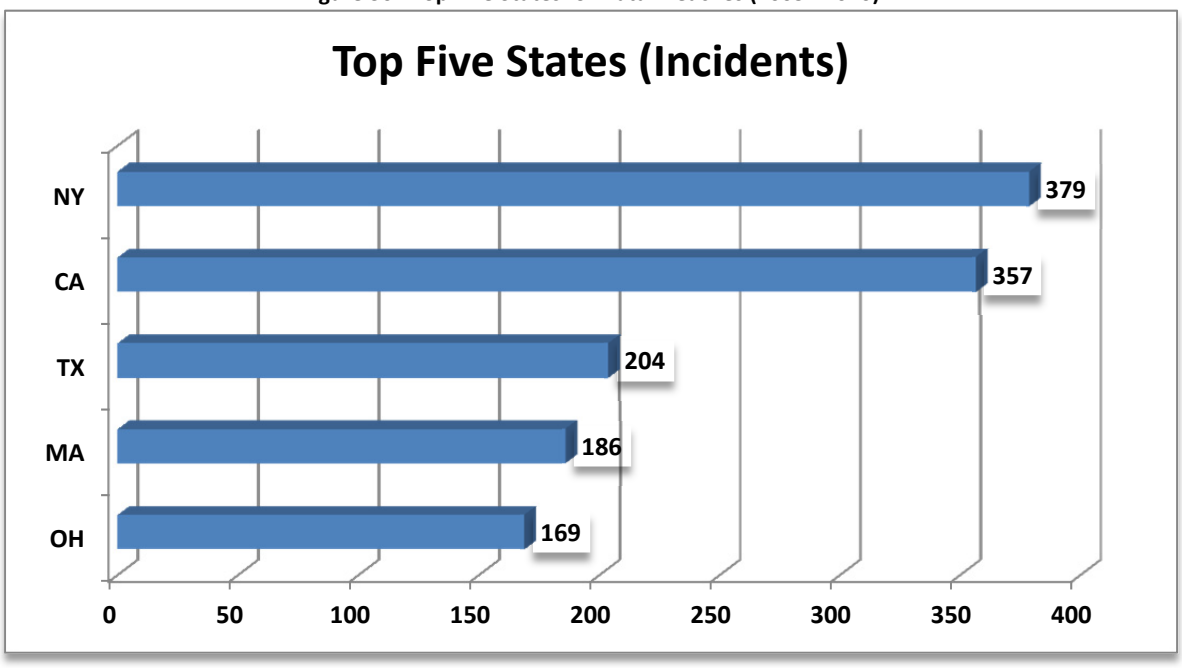
As mentioned before, the complexity induced in having multiple state data breach laws makes navigating compliance very difficult. When you add to this those organizations that straddle the boundaries of multiple countries and continents, the problem becomes even more challenging.

> Unfortunately, while data privacy regulations continue to multiply, such regulations generally are not anchored on a common global standard. Worse, they also have trouble keeping up with technology advances and business practices that are dramatically changing how data is created, shared and stored. The result is a maze of regulations and privacy laws that are often intricate and complex at best, and at worst are costly and contradictory, or fail to properly address changing business models, global data flows and technology advances [13].

An organization's activities regarding cross-border data sharing are particularly important, as data may be moved from a more secure environment to a less secure one without this being taken into the risk consideration. Legal remedies for organizations whose data has been breached in countries where this activity is not against the local laws may have no recourse. Contracts that involve data sharing in geographically dispersed areas should always contain language that deals with how the data will be protected, and how it will be handled in the event the contract is terminated.

Looking at the United States incidents again, Figure 30 shows us the top five states for data breach incidents. When looking at this, keep in mind the fact that not all states have laws on the books requiring organizations to report these events. Thus, the states with the longest tenure with these laws may show a larger total by virtue of having reporting requirements the longest.

Figure 30: Top Five States for Data Breaches (2005 - 2010)

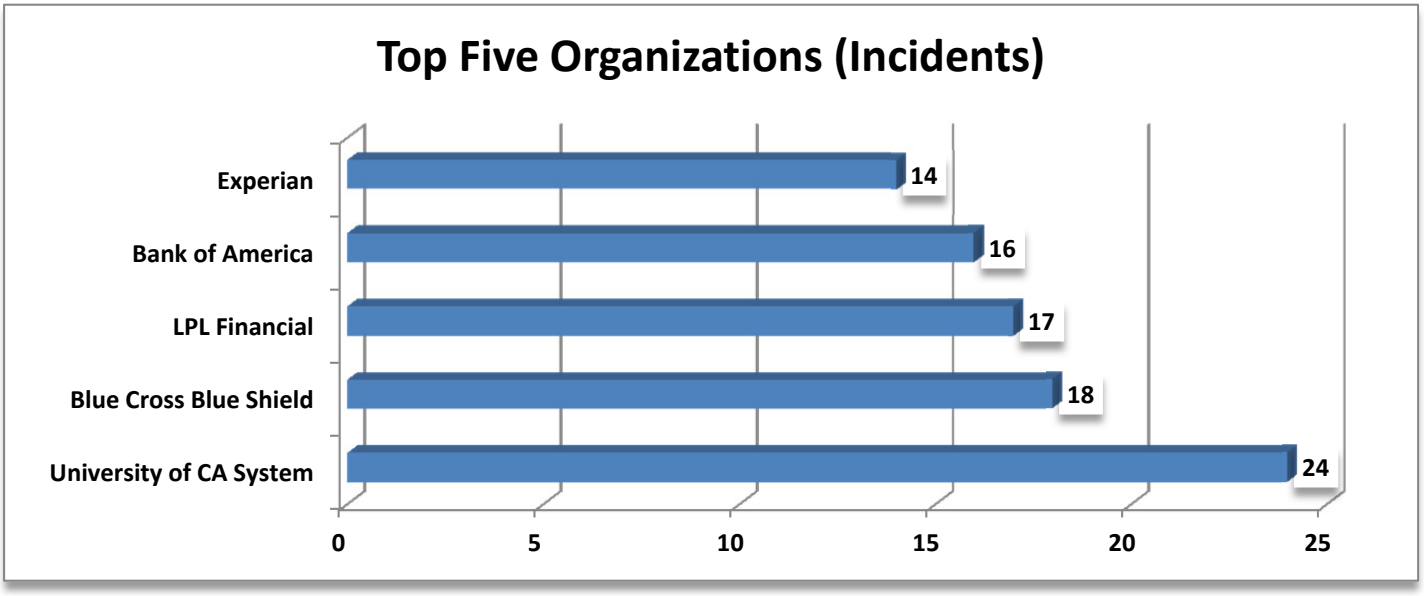


The states with no incidents in the study were Idaho, North Dakota, and South Dakota. Given that New Mexico, Alabama, Kentucky, and South Dakota have no data breach laws, it is surprising that two of the states listed have laws on the books. New Mexico listed 9 incidents, Alabama had 6 and Kentucky showed 12 incidents. These may have been reported based on where the data breach occurred, or where the organization that experienced the breach was located. Since it seems that the states with notification requirements show more incidents, it stands to reason that the people being notified based on these requirements are better able to monitor the situation than those in states where there are no such notification requirements. If the data subject victims do not know their data has been compromised, they may not be on the lookout for any subsequent financial consequences.

Organizational Sectors

TLV featured a graphic with the top five organizations with the highest number of incidents over the course of the study. Figure 31 shows these results with the subsequent additional incidents. The players in the list are the same, what has changed slightly is the order.

Figure 31: Organizations with the Most Incidents (2005 - 2010)



The University of California system maintained its lead, increasing from 21 to 24 incidents. The UC system's incidents are made up of multiple locations, which may handle security independently from each other. In any case, they would do well to pool their knowledge as an organization to work together to get out of first place.

LPL Financial and Blue Cross switched places. LPL had three more incidents, while Blue Cross had four. Experian and Bank of America also switched places. Experian gained one more incident, while Bank of America gained four.

The problem of multiple breaches is likely a combination of factors. These organizations are large and complex. As the complexity increases, the likelihood that one group knows about all the data flows across the organization becomes increasingly small. When this is paired with the potential monetary value of the data to thieves, you have the potential for multiple external threat actors targeting an organization. Combine this with the internal and third party threat, and organizations can become overwhelmed [13].

Figure 32: Incidents by Organizational Type (2005 - 2010)

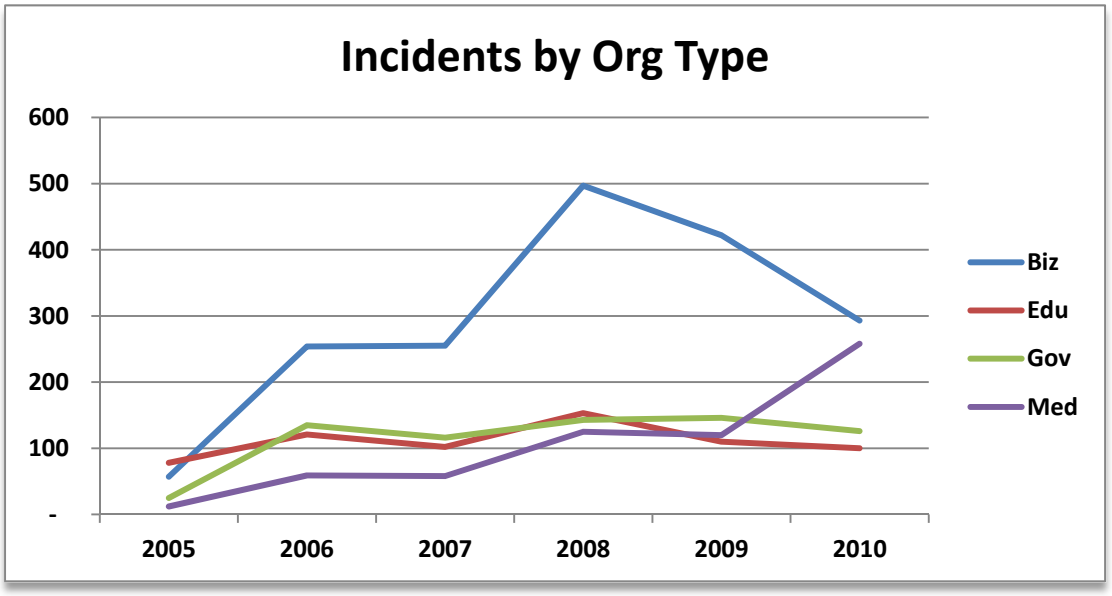


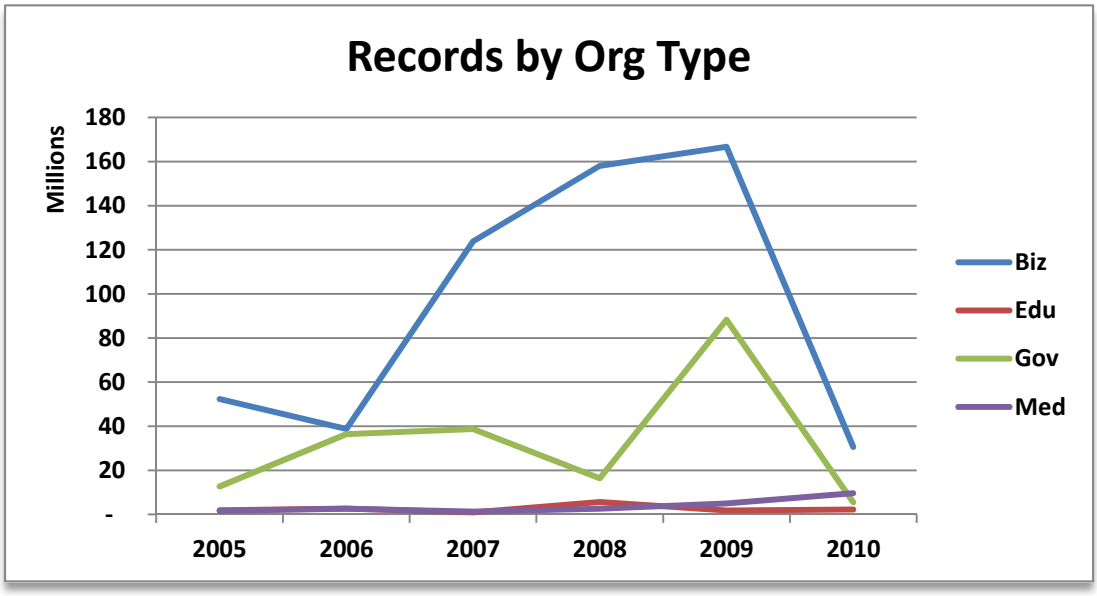
Figure 32 breaks out the incidents by organizational type. The Business sector remains the leader in incidents by a significant margin. The spike in Medical sector incidents, however, is new for 2010, and shows the effect of the previously mentioned reporting requirements for medical data loss. The Medical sector is trending upward steadily, however, despite the fluctuations in overall incidents. Table 12 shows the full breakdown of the incidents by sector and year.

Table 12: Number of Incidents by Organizational Type (2005 - 2010)

	Organizational Type				
Year	Biz	Edu	Gov	Med	Total
2005	57	78	25	12	**172**
2006	254	121	135	59	**569**
2007	255	102	116	58	**531**
2008	497	153	143	125	**918**
2009	422	110	146	120	**798**
2010	293	100	126	258	**777**
Total	**1,778**	**664**	**691**	**632**	**3,765**

The Business sector maintains the lead position in incidents. It is almost as many as the other three sectors combined. The Business sector also leads in the number of records disclosed, as you can see in Figure 33. Even with the sharp decline in number of records disclosed in 2010, the Business sector retains dominance. The Government sector in contrast, dips below the Medical sector in rank for records disclosed.

Figure 33: Records Disclosed by Organizational Type (2005 - 2010)



The Education sector, despite a small bump in 2008, has maintained the position of the least amount of records disclosed for almost every year in the study. This is remarkable, given that Educational institutions have a high number of highly desirable records from an identity thief's point of view. Social Security Numbers (SSNs) are frequently part of the student's records, and financial aid information almost always requires either an SSN or some other type of desirable data.

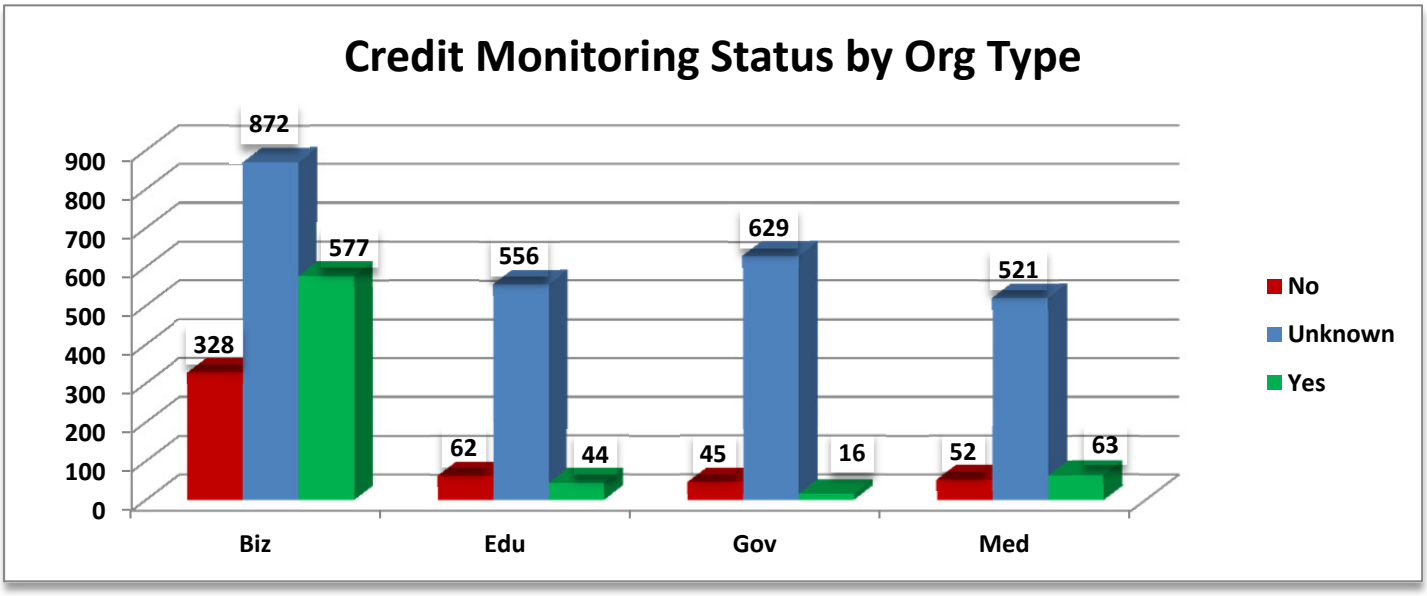
Table 13: Number of Records by Organizational Type (2005 - 2010)

Year	Organizational Type				Total
	Biz	Edu	Gov	Med	
2005	52,337,228	1,867,896	12,698,605	1,651,834	**68,555,563**
2006	38,704,234	2,666,803	36,414,622	2,592,210	**80,377,869**
2007	123,855,756	938,610	38,711,912	1,307,600	**164,813,878**
2008	158,107,514	5,635,558	16,390,863	2,573,834	**182,707,769**
2009	166,689,083	1,797,117	88,277,764	4,995,526	**261,759,490**
2010	30,646,732	2,243,820	5,535,352	9,654,959	**48,080,863**
Total	**570,340,547**	**15,149,804**	**198,029,118**	**22,775,963**	**806,295,432**

To see how the data subject victims are being treated, we look at whether or not the breached organizations are offering them basic credit monitoring services to deal with the increased risk of identity theft and financial fraud. Figure 34 shows this data by organizational type. For the majority of the incidents, this data simply is not included in the reports. For those where we do know the status, however, the changes from the prior report to this one are telling.

In TLV, the number of organizations offering credit monitoring was 505 in the Business sector. The Educational sector offered monitoring in 33 cases; Government in 14; and Medical in 56. The organizations that chose not to offer this service in TLV were 282 in the Business sector; 59 in Education; 39 in Government; and 41 in Medical. You can see below how this has changed.

Figure 34: Credit Monitoring Status by Organizational Type (2005 - 2010)



When the data stolen included the subject's Social Security Number, only 24% of the cases indicated credit monitoring services were offered. When the data stolen was the credit card information, only 14% of the victims were offered these services.

The Business Sector

In the prior report, the rank for incidents was Laptop, Hack and Fraud-SE. The ordered rank for records was Hack, Fraud-SE and Tape. As you can see, the incident vectors and ranking remain consistent, but there is a change to the third place from Tape to Web in the records lost vectors.

Figures 35: Business Sector Top Three Incident Vectors (2005 - 2010)

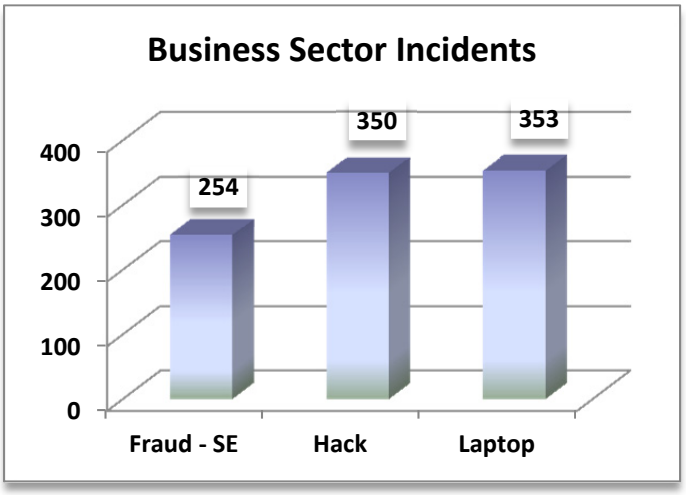
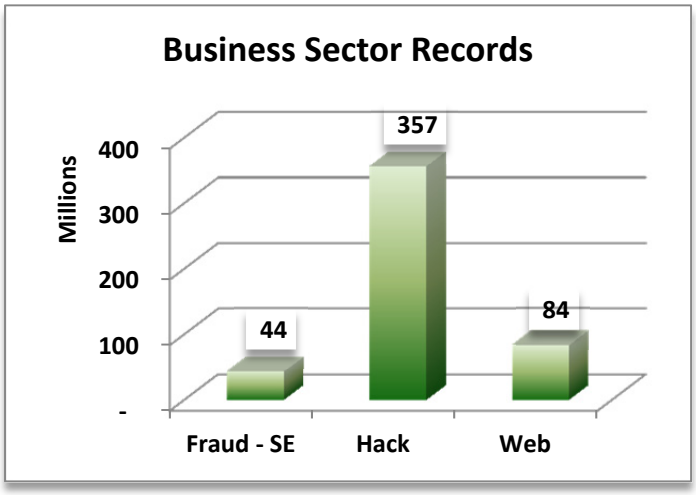


Figure 36: Top Three Records Vectors (2005 -2010)



There is a significant gap between the first and second place vectors for the Business sector. Hacking represents such a large percentage of the records for this vector that it accounts for 63% of the records divulged. In fact, all of the other vectors combined come to only 213 million records, compared to Hacking's 357 million.

By breaking the Business sector down into some of the larger SubSectors, we can see how certain industries are faring. The Retail SubSector is a new addition with TLV2011. While the data was in the study all along, it was not broken out

specifically. It represents a larger slice of the Business sector than even the Financial SubSector by incident count. By records count, however, Financial takes the lead.

Figure 37: Business SubSector Incidents (2005 - 2010)

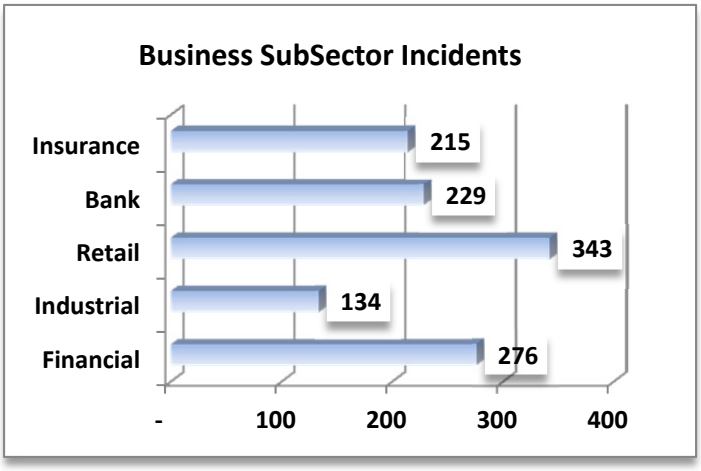
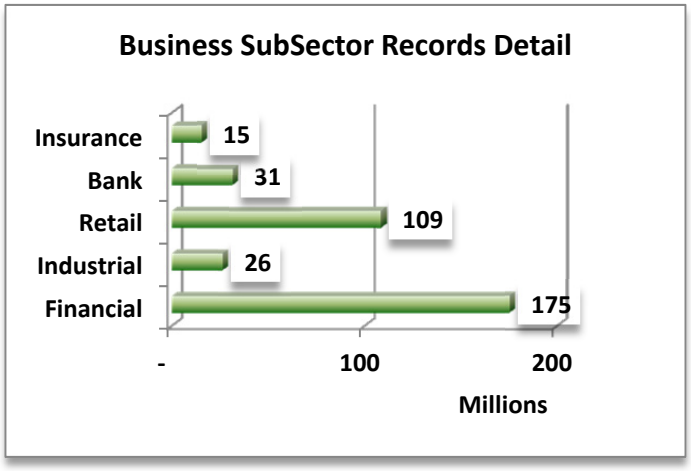


Figure 38: Business SubSector Records (2005 - 2010)



The Retail SubSector leads in incidents by a comfortable margin. Skimmers were a common tool used in these incidents where credit card fraud was the goal. The Financial SubSector was more likely to suffer from the hacking vector, but with banks, skimmers on ATM machines was also prevalent.

The Educational Sector

For TLV, the ranked order for Education incident vectors was Hack, Web and Laptop. The same data for records was Hack, Tape and Web. As you can see, the incident vectors have seen some significant changes, with the Laptop vector taking over second place, and the Computer vector new to the top three. The Hack vector increased from 158 incidents to 184. The Laptop vector had been at 78, while the Computer vector hadn't even made the list. The records vectors remained constant between the two studies including preserving the order of rank. The Tape vector saw the most significant increase, going from 2 million to 4.3 million. Tapes tend to have large numbers of records on them, so when they are lost or stolen, the breach size is usually quite significant. Encryption technology for tape backups has existed for many years, and should be explored as a mitigating control for this vector.

Figure 39: Education Sector Top Three Incident Vectors (2005 - 2010)

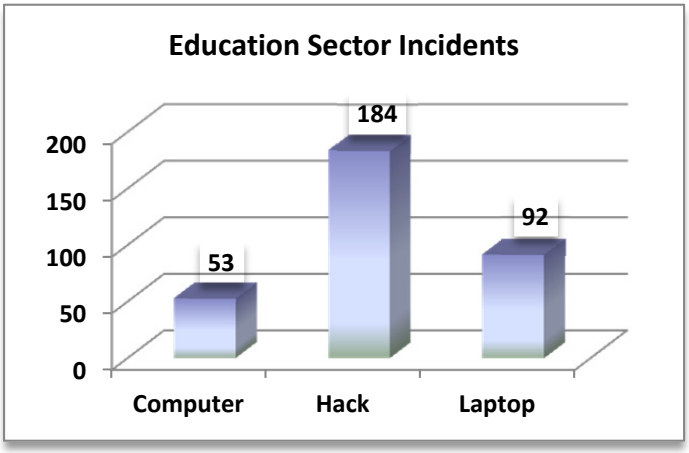
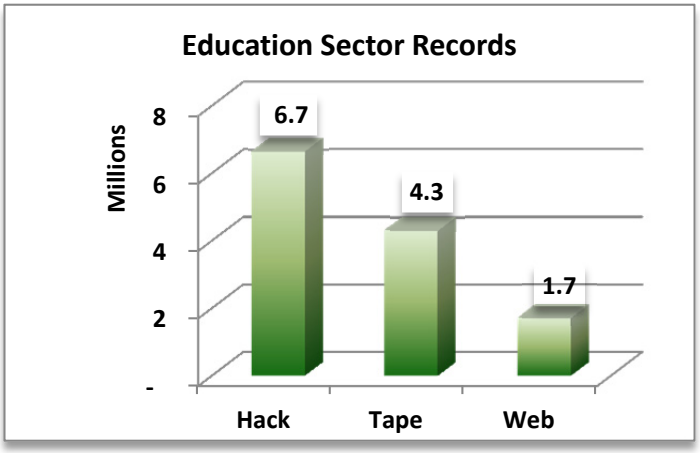


Figure 40: Education Sector Top Three Records Vectors (2005 - 2010)



Hacking was the most prevalent cause of both records disclosed and incidents for the Education sector. The median figure for records for this vector in the Education sector was 6,540. Universities remain the largest SubSector of Education, representing 83% incidents and accounting for 91% records.

The Government Sector

The TLV top three Government incident vectors were Laptop, Web and Documents in that order. The Document vector rose from 84 to 118 incidents to claim the top spot, while Laptop and Web each saw small increases. The Documents vector was clearly prominent in the Government sector incidents. These incidents most commonly involved inappropriate disposal, which demonstrates the need for managing the organization's paper as well as electronic information for their full lifecycles. The records vectors were Drive/Media, Laptop and Hack, so the rank remains the same. All three experienced increases of only about a million records each.

Figure 41: Government Sector Top Three Incident Vectors (2005 - 2010)

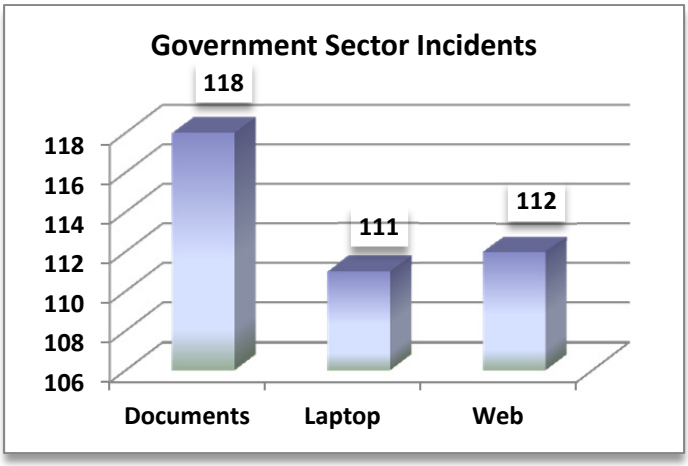
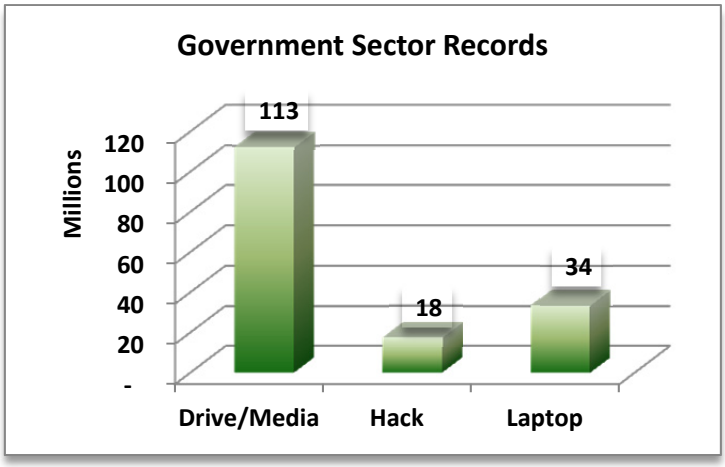


Figure 42: Government Sector Top Three Records Vectors (2005 - 2010)



One major subsector of the Government is the Military. The following two figures illustrate the top three vectors for both incidents and records lost in military organizations. This data is new for the current report and was not specifically called out in the prior paper.

Figure 43: Top Military Incident Breach Vectors (2005 - 2010)

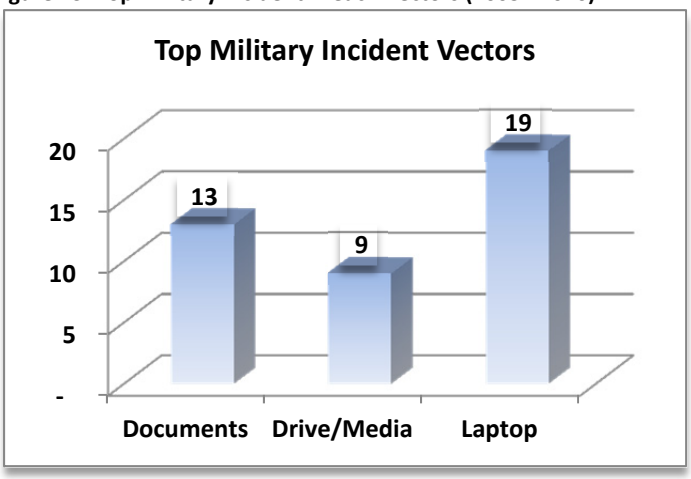
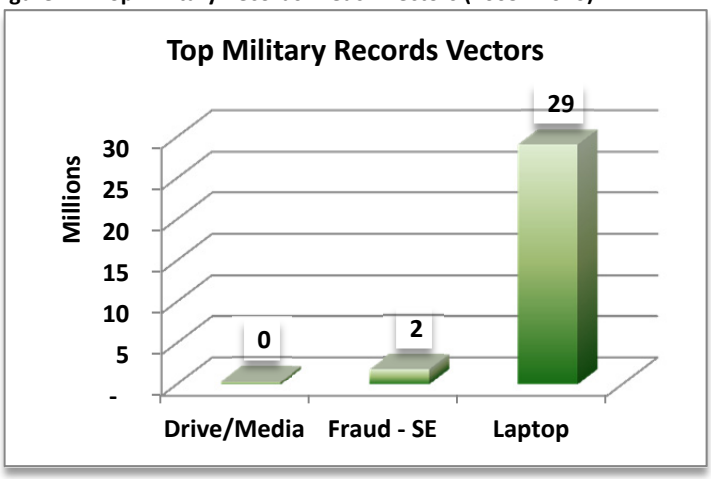


Figure 44: Top Military Records Breach Vectors (2005 - 2010)



The Military SubSector shares the problem with managing the paper document lifecycle with the rest of the Government. However, when looking at record losses, the laptop was the leader. Note that 28.6 million of those records were from once incident when a laptop was stolen from a Veteran's Affairs office in 2006.

The Drive/Media vector is in the top three for both of these records charts. This is of particular interest, given the 2008 Department of Defense incident where a USB stick containing malware from a foreign intelligence agency was plugged into a military laptop, causing infection of both classified and unclassified networks [8].

The Medical Sector

In TLV, the ranked order for Medical sector incidents was Laptop, Documents and Fraud-SE. The incident vectors all saw significant growth, with Laptop changing from 103 to 164; Document going from 48 to 118 and Fraud-SE increasing from 41 to 76. With Documents nearly tripling, it was able to overtake the Fraud-SE vector for second place. The TLV records vectors were Tape, Laptop and Drive/Media in that order. Laptop barely edged out the Drive/Media vector for the top spot this year, but Drive/Media had the largest change. That vector went from 2.5 million to 5.4 million, while Laptop went from 3 million to 5.7 million records. Tape saw a modest increase from 3.8 million to just over 4 million.

Figure 45: Medical Sector Top Three Incident Vectors (2005 - 2010)

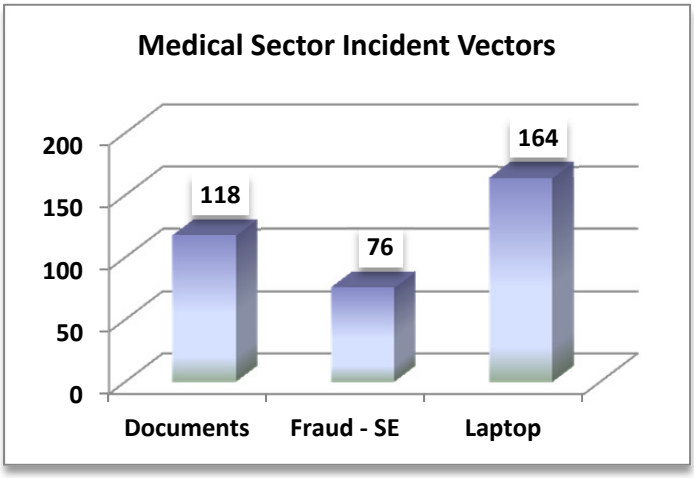
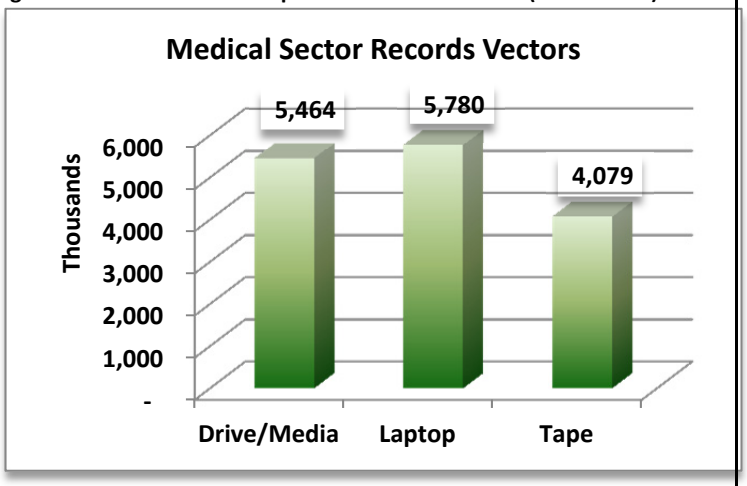


Figure 46: Medical Sector Top Three Records Vectors (2005 - 2010)



A related trend was the incidents of doctor and dental offices having to disclose incidents under the new requirements. To that end, here are the vectors for both incidents and records disclosed for that SubSector of the Medical category.

Figure 47: Medical Practitioner Incident Vectors (2005 - 2010)

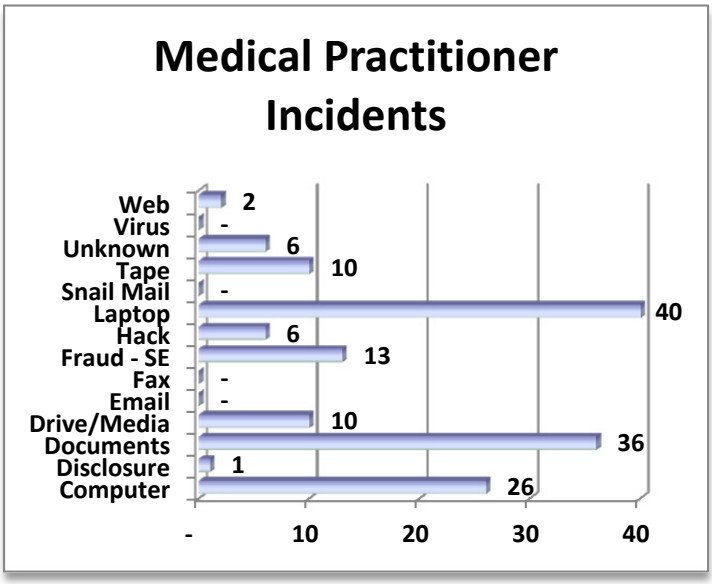
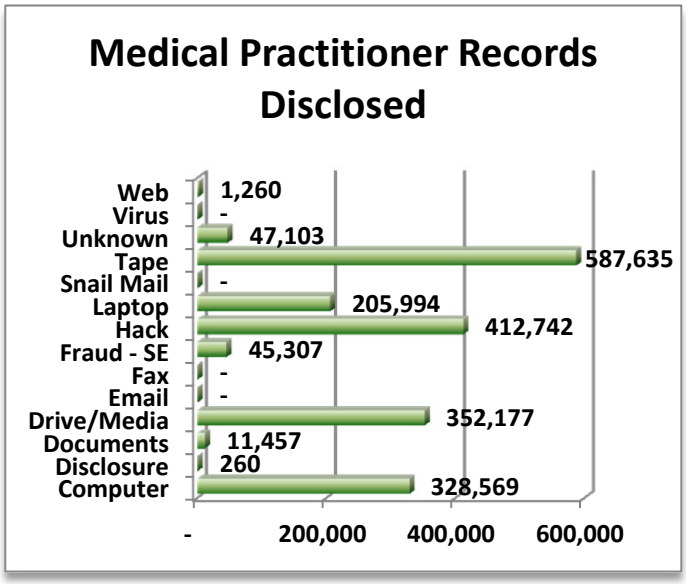


Figure 48: Medical Practitioner Records Disclosed Vectors (2005 - 2010)



The Medical Practitioner Records Vector values were so close together that in the interest of clarity, the actual values are shown with no rounding. Note, the scale on this chart has changed from millions to thousands.

As you can see, laptop vector is the lead in both incidents and records disclosed.

Data Types

Most commonly, the data breach disclosure laws have a data combination requirement to trigger a reportable event. Usually, it is a person's first initial and last name in combination with one of the following:

- Social Security Number (SSN)
- Driver's License or State ID
- Credit card or account number (CCN)
- Financial or account information
- Date of Birth
- Medial data

The media also takes note of disclosures of other data if the breach is large enough. For example, large scale disclosure of email addresses or phone numbers have made the news reports. With email, there is some merit to calling this a reportable breach, since recent cases have resulted in targeted phishing attacks designed to gain a foothold at the recipient's organization or location.

For an SSN to be useful, it must be associated with a person's name. Without that association, it is just a group of numbers that could have been randomly generated and its usefulness for committing fraud or identity theft is reduced. If organizations are still using SSNs as their unique identifier, however, they should be taking steps to eliminate them wherever possible. Reducing the locations where this highly sought after data element is stored will only help to reduce the risk of their disclosure. Data masking and encryption should be considered in cases where they must be stored and used.

It should be noted that some incidents disclosed multiple data types. The total number of records in this section will thus not equal 3,765 between these data types because of that overlap.

Social Security Numbers

Between 2005 and 2010, there were 2,438 incidents involving SSN data. It was by far the most common data type listed, but this is in part due to the prevalence of its mention in the breach disclosure laws as a trigger. For most of these laws, if an SSN is lost along with some variation on the person's name or initials, the organization is required to notify. Overall, there were over 276 million known records containing SSNs disclosed. The median records per breach was 2,249, but 835 (34%) of the incidents did not disclose how many records were lost. In TLV, 69% of the cases disclosed SSNs. In TLV2011, that figure has dropped to 65%.

Figure 49: Incidents of SSN Data Type by Organizational Type (2005 - 2010)

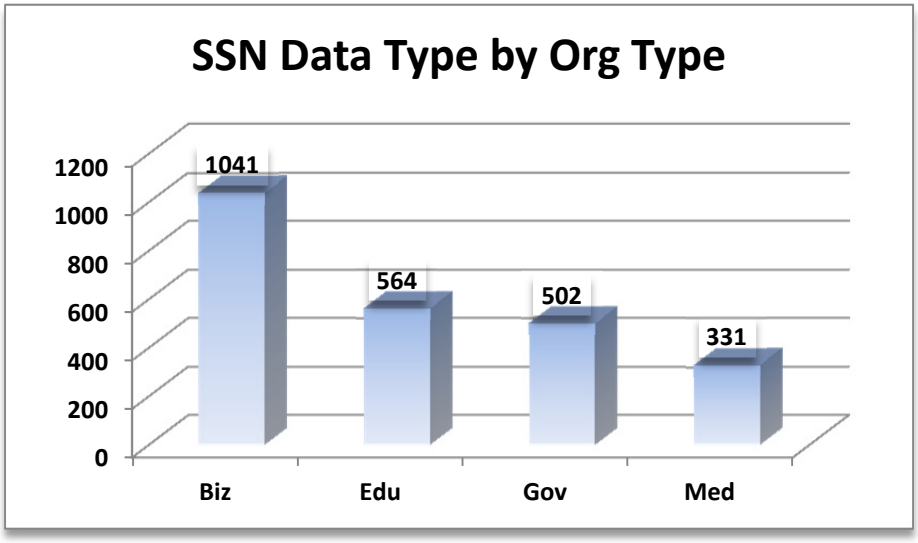
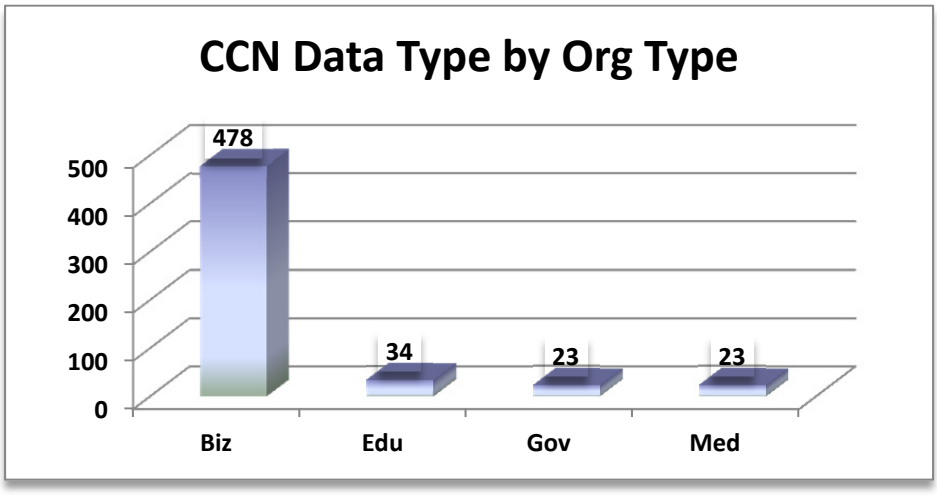


Figure 49 illustrates the breakdown of incidents by organizational type. The Business sector leads in SSN disclosure incidents, as they did in the prior report. The number has changed from 844 to 1041 incidents. The ranking of organizations remains the same between the two papers, and the same can be said for the CCN ranking.

Credit Cards

There were 558 incidents where CCN data was involved. They accounted for almost 330 million records. The median records disclosed was 1,000; and 45% of the incidents did not list how many records were disclosed. These records should fall under the Payment Card Industry's Data Security Standard (PCI-DSS), and the organizations that have experienced these incidents will have to undergo further scrutiny to prove they are compliant with this standard.

Figure 50: Incidents of CCN Data Type by Organizational Type (2005 - 2010)

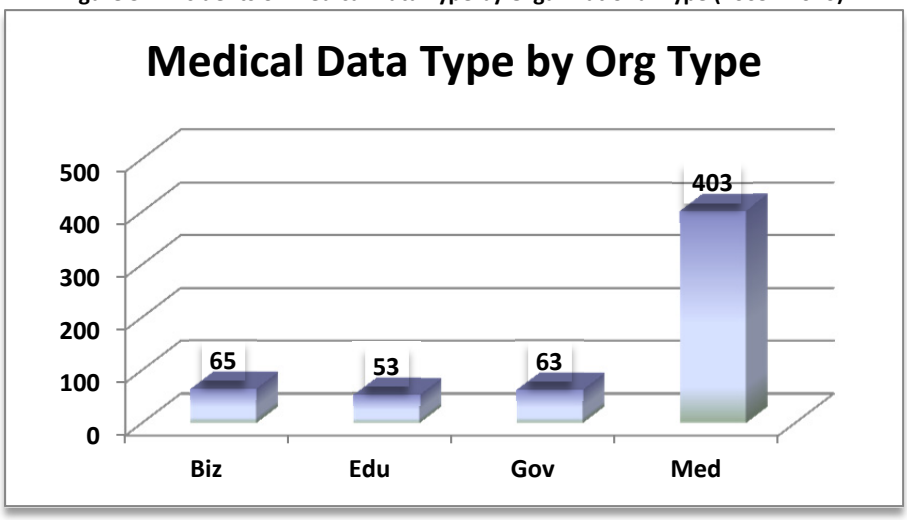


The Business sector has increased from 352 incidents in the prior report to 478, while the other organizational sectors remain in the same order and saw only proportionally small increases.

Medical Information

From 2005 to 2010, there were 584 incidents disclosing over 17 million records with Medical data. The median records disclosed for medical data was 2,000. There were 27% of incidents reporting "unknown" loss figures. These incidents are now required to be reported under the HIPAA/Hitech regulations.

Figure 51: Incidents of Medical Data Type by Organizational Type (2005 - 2010)



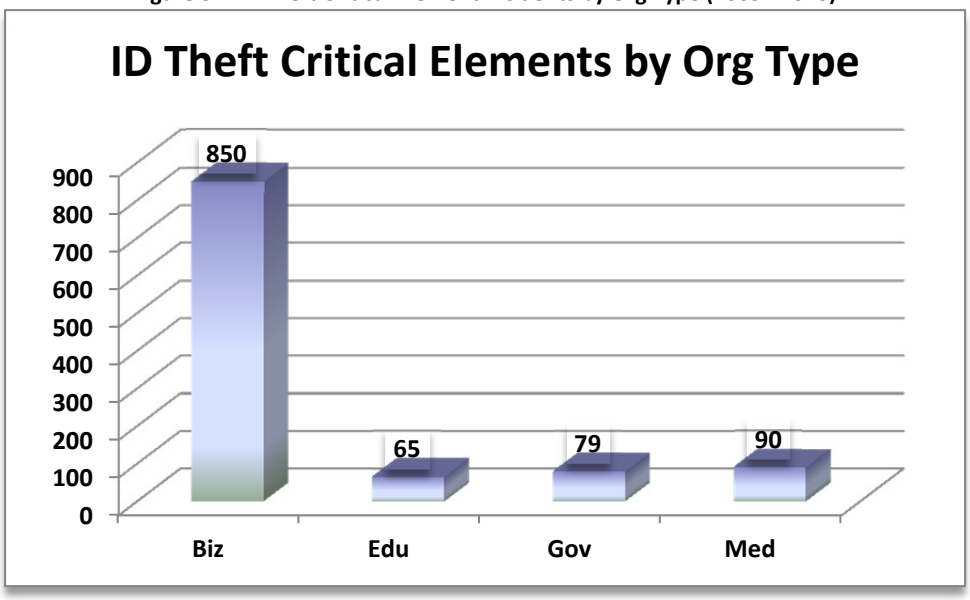
As expected, the Medical sector remains the lead in disclosing medical data. The increase from TLV to TLV2011 was significant. The Medical sector increased from 160 to 403 incidents—more than doubling the number of incidents in such a short time.

The ID Theft Critical Data Elements

The Identity Theft critical data elements are those that, in combination with the Name and Address, facilitate the commission of identity theft and financial fraud—namely the SSN and date of birth. In TLV, we looked at the incidents with these three data items all lost in the same event. At the time of that study, there were only 262 incidents that contained all three items. In contrast, there are now a total of 1,084.

As you can see in the figure below, the Business sector shows a substantial increase. It has gone from 168 incidents in the prior study to 850. However, in only 13% of these cases where the combination of data puts the subject victim into the worst position possible, are these organizations confirmed to have offered credit monitoring. Now, there are a large number of unknowns in this area as well—in the majority of the cases, the reports simply do not say one way or the other whether this service is offered. This is a metric primarily gleaned from the original data breach notification letters obtained through either FOIA requests or from those government entities that are directly posting the original documents as part of the event report. For instance, in the Business sector, 38 cases are confirmed that the service definitely is not offered. In the remaining 701 records, the credit monitoring status is not provided.

Figure 52: ID Theft Critical Element Incidents by Org Type (2005 - 2010)



In only 12% of these cases overall (regardless of organizational type), is there confirmation that the organization offered credit monitoring services.

The Educational sector came close to doubling between TLV and TLV2011. The prior report showed 39 cases where the ID Theft trio of elements was simultaneously disclosed. The Government sector tripled from 26 to 79 incidents. The Medical sector also saw a noteworthy increase, from 29 to 90.

Relationships

In this section, we examine the data from the perspective of the relationship between the organization losing the data and the data subject. Specifically, we look at who the data subject is to the organization. There is a cost for a lost customer, but there is also a cost for a lost employee if they leave in part because of the way the organization mishandled their data.

Organizations may tend to focus exclusively on customer data when this is not the only sensitive information they hold in their care.

Figure 53: Incidents by Data Subject Relationship (2005 - 2010)

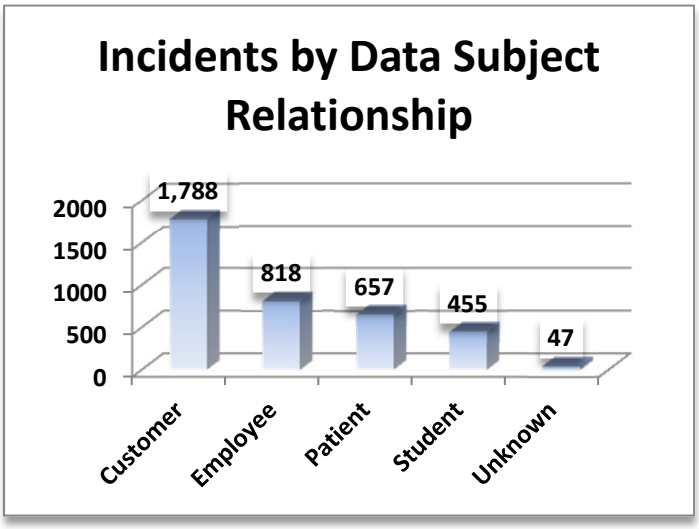
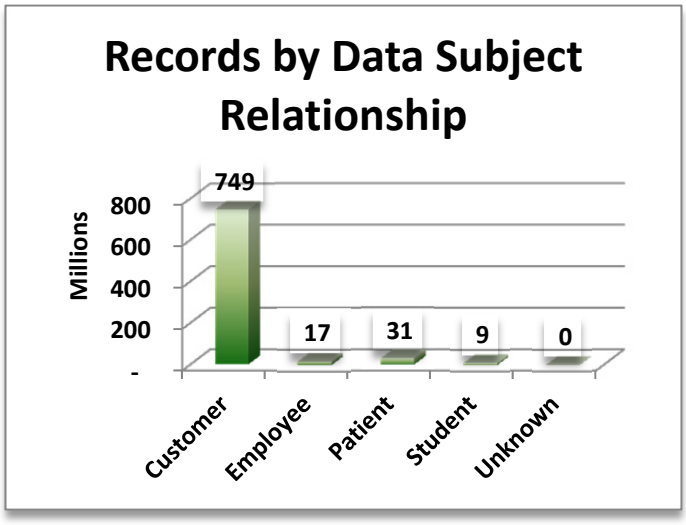


Figure 54: Records by Data Subject Relationship (2005 - 2010)



Looking at the above graphs, we can see that the Customer relationship is the most commonly compromised in data breaches, both in terms of incidents and records disclosed. The second most commonly jeopardized relationship is with Employees. Beyond that, the rank continues with both incident and records in the same order.

Customers

Figure 55: Customer Data Records Disclosed by Organizational Type (2005 - 2010)

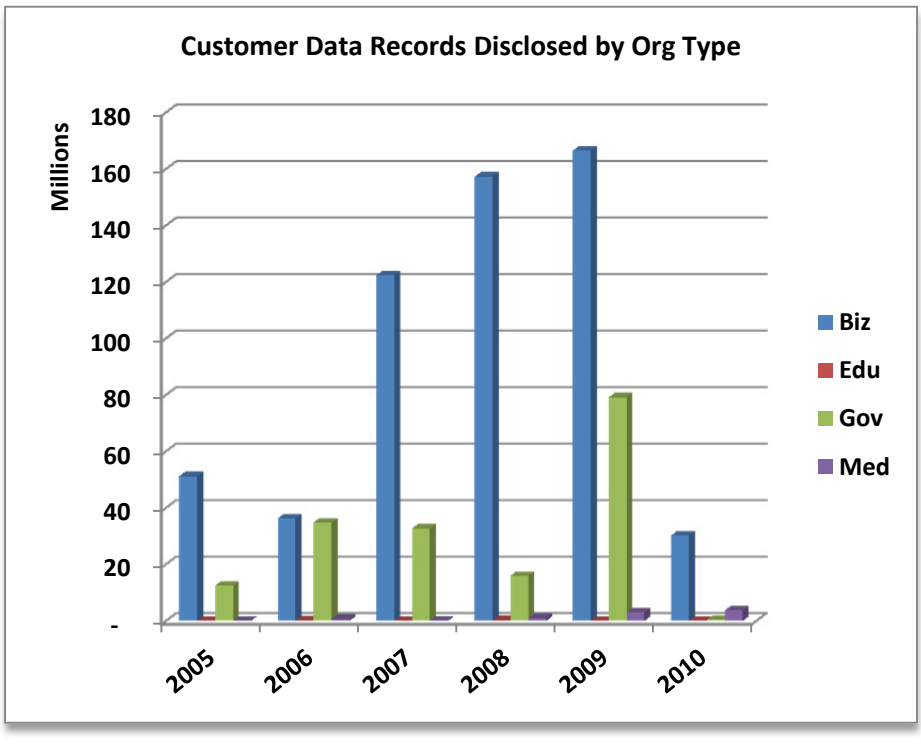


Figure 55 illustrates the trend in loss for each of the organizational sectors over the course of the study. The Business sector is the clear leader in losing customer data. In fact, this sector lost customer data 72% of the time. In second place is the Government sector, losing customer data 56% of the time. In 54% of these incidents, the data that was disclosed was the name and address in combination with the Social Security Number. In contrast, customer credit card data was lost only 28% of the time.

Looking at the top incident breach vectors for Customer data, the Hack vector increased from 275 in TLV to 380. The Fraud-SE vector moved from third place to second place in incidents, with an increase of 75 incidents. Laptop dropped to third with only an increase of 31 incidents.

Figure 56: Customer Incident Breach Vectors (2005 - 2010)

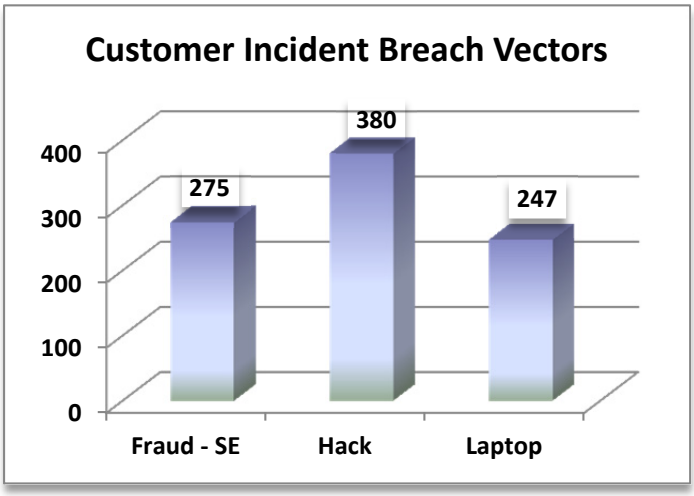
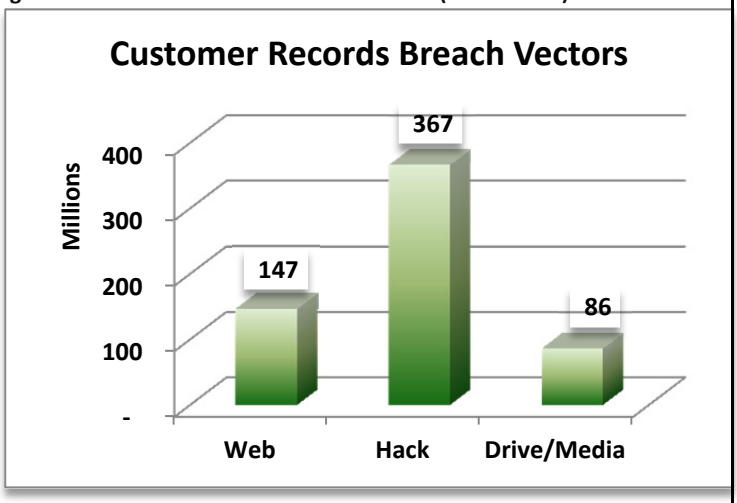


Figure 57: Customer Records Breach Vectors (2005 - 2010)



The records breach vectors saw increases as well. The Hack vector increased by 54 million records and retains the lead. The Web vector jumped from third place to second with an increase of 65 million records. The Drive/Media vector increased by only 4 million records, but managed to stay in the top three.

The median records per breach was 2,000 for Customer data. There were 1,788 incidents accounting for nearly 750 million Customer records during this study.

Employees

Figure 58: Employee Data Records Disclosed by Organizational Type (2005 - 2010)

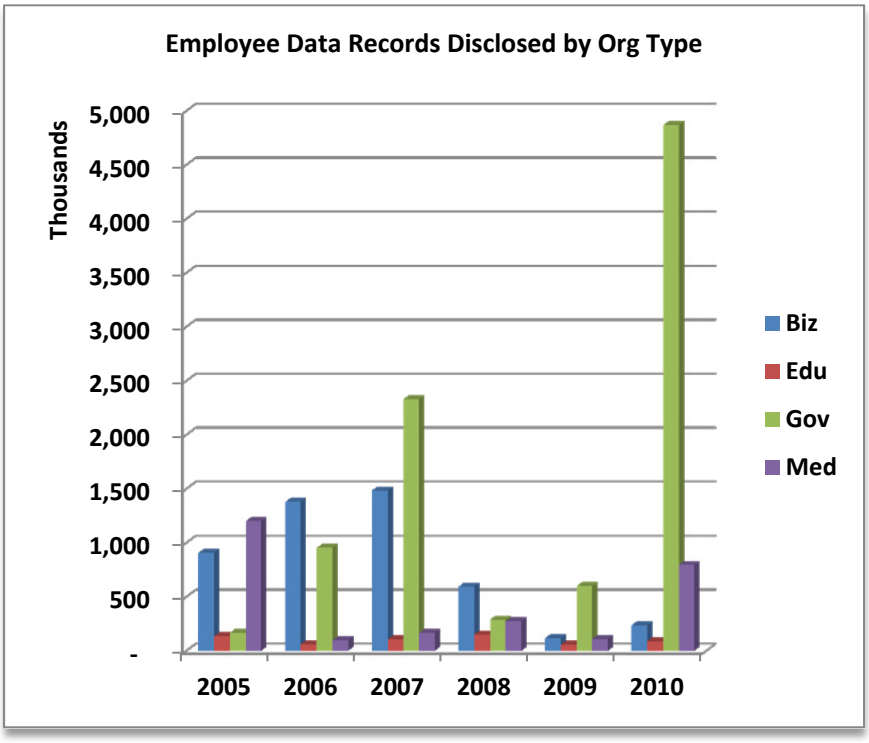


Figure 58 shows the organizational break down per year for Employee records. While the Business sector held the lead position for the first couple of years, you can see that the Government sector has increased significantly in the past year. The overall number of records disclosed is fairly low by comparison with the Customer records; the scale of this chart is in thousands rather than millions.

The large spike in Government records in 2010 is primarily due to two breaches. The first was a missing laptop containing 207 thousand records; and the second was a missing hard drive, containing over 250 thousand records. Between these two incidents, they account for 48% of the records disclosed in 2010 for the Government sector.

The top incident vector for Employee data remains the Laptop vector, with an increase from 225 to 251. The Web vector increased from 68 to 87, and the Documents vector displaced the Drive/Media vector this year.

Figure 59: Customer Incident Breach Vectors (2005 - 2010)

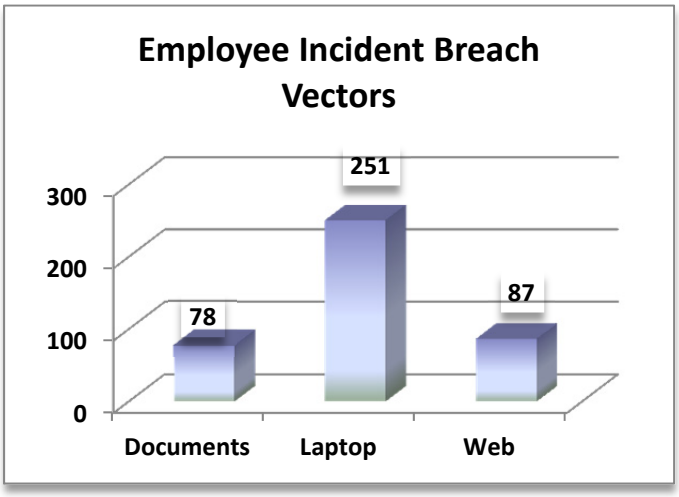
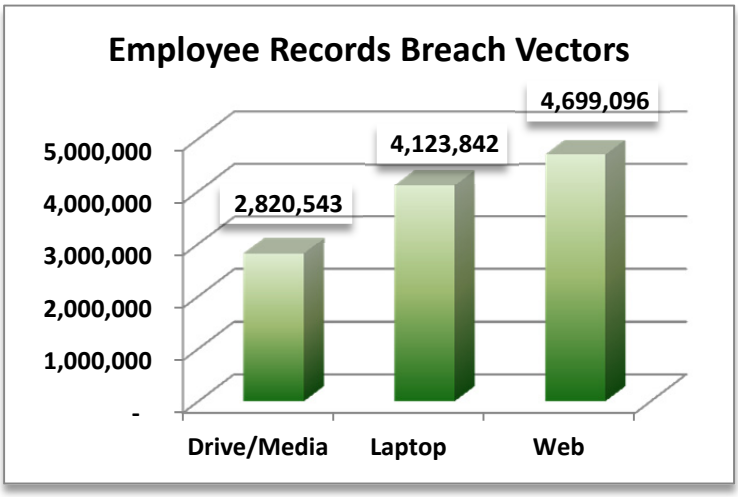


Figure 60: Employee Records Breach Vectors (2005 - 2010)

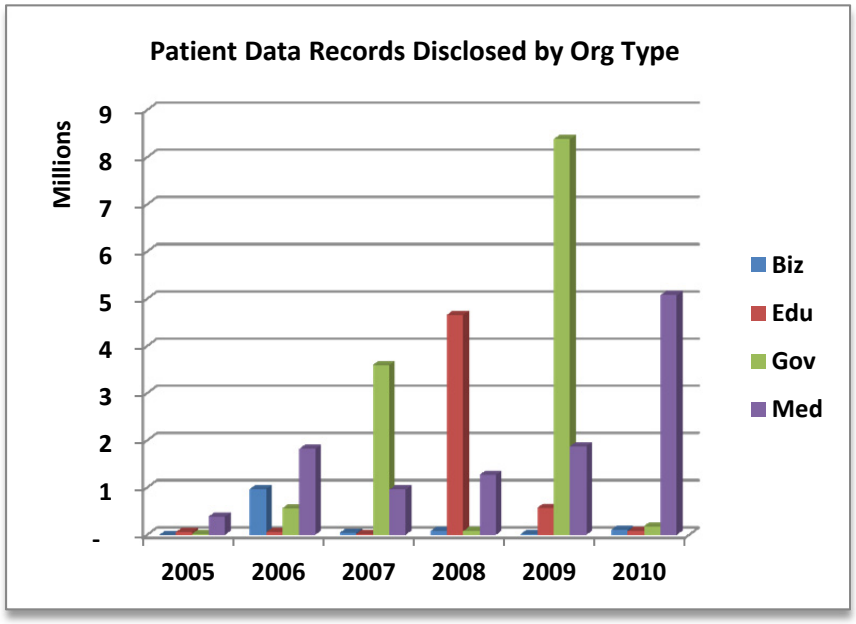


For records, the top vector was Laptop, but is now Web, as you can see. The Web vector was not in the list last year at all for top records vectors. The Laptop vector increased by over 234 thousand records. The Drive/Media dropped to third place, although it almost doubled in size between TLV and TLV2011.

The median records per breach for Employee data was 2,000. There were 818 incidents accounting for over 17 million Employee records disclosed throughout the study.

Patients

Figure 61: Patient Data Records Disclosed by Organizational Type (2005 - 2010)



The Government sector has the lead in disclosing Patient data as well. As seen in 2009, there was a sharp spike, followed by a significant drop. The 2009 figure was dominated by the National Archives and Records Administration breach, which was responsible for 76 million records. The Medical sector has not held the dominant position in this vector, which is counter intuitive—they would be the largest repository of medical data after all. The leadership in 2010 is a new factor, and likely influenced by the breach reporting requirements coming into effect for medical incidents. There were three large breaches in the Medical sector in 2010 that drove this total particularly high in relation to other years. They included a Laptop breach that affected 1.2 million records; a Fraud-SE breach that disclosed 1.5 million records; and a Tape breach that accounted for 1.7 million records. Between those three, they accounted for 92% of the records in the Medical sector.

An increasing trend in the data has been in the area of medical fraud, particularly in the 2010 reports. When the data on patients is compromised, sometimes the goal is not to steal identities in the traditional sense; it is to work the medical insurance system to file false claims of durable goods and services. These incidents are becoming more prevalent in the news reports; potentially because of the reporting requirements, but also because the increased public debate on medical care has made this more newsworthy.

Figure 62: Patient Incident Breach Vectors (2005 - 2010)

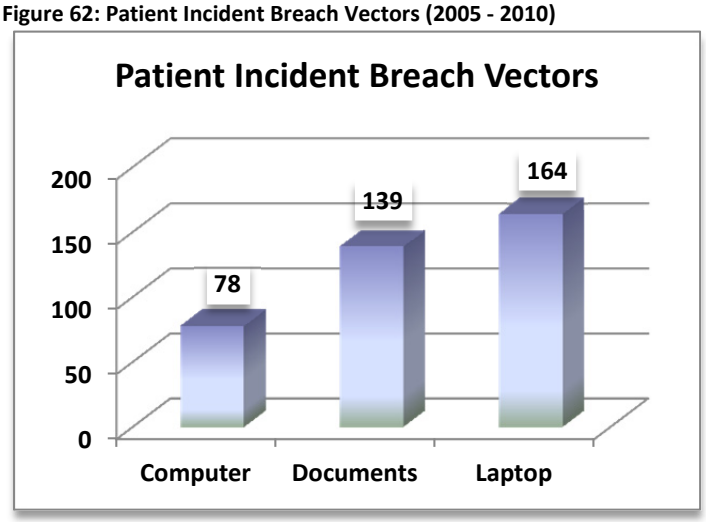


Figure 63: Patient Records Breach Vectors (2005 - 2010)

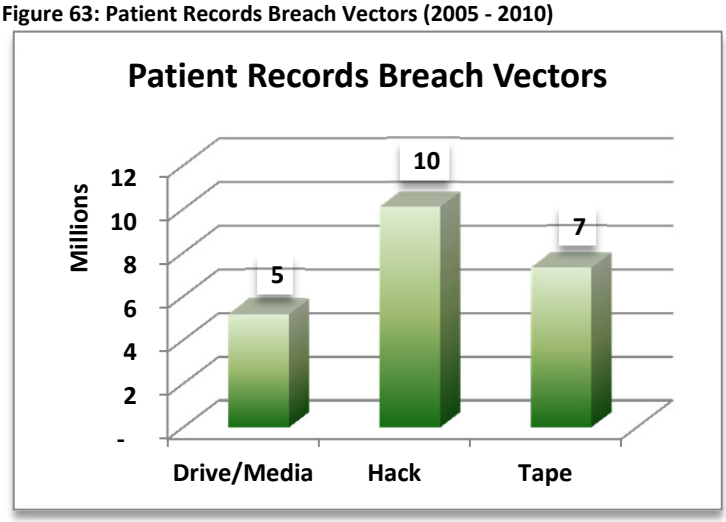
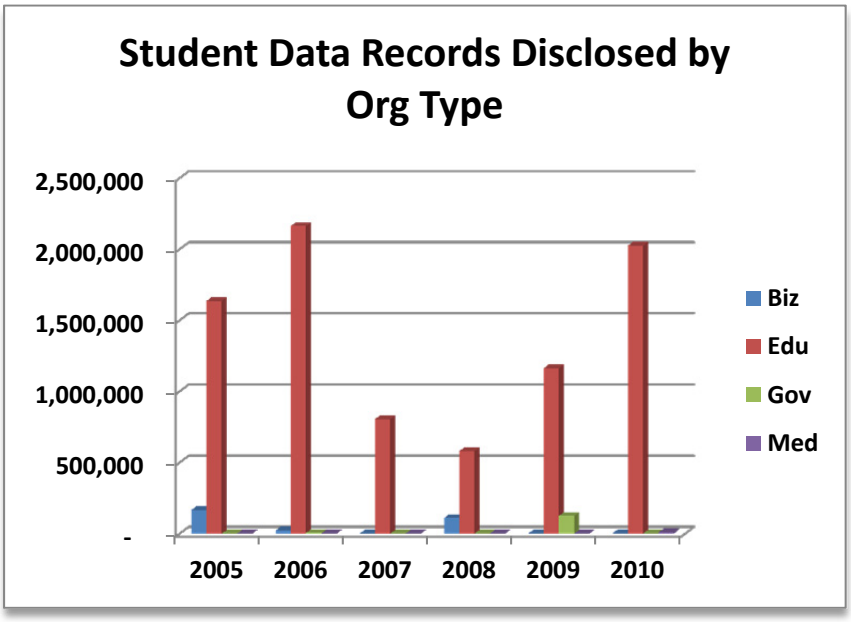


Figure 62 shows the top breach vectors for incidents. The members of this chart and their order remain consistent with the prior report. The Laptop vector gained 67 incidents; the Documents vector increased by 77; and the Computer vector increased by 37. The records vectors also remained the same in order and membership. The Hack vector was 9.3 million; the Tape vector was 5.2 million; and the Drive/Media vector was 3.6 million in the previous report.

The median records disclosed for Patient data is 2,000. There were 657 incidents and over 30 million patient records disclosed.

Students

Figure 64: Student Data Records Disclosed by Organizational Type (2005 - 2010)



For Student record disclosures, as expected the Educational sector is dominant—so much that the other sectors barely register in this chart. It is interesting to see the 2010 data keeps pace with other high years, despite the sharp drop in records disclosed.

When the data relationship is Student, the data element disclosed almost always includes the data subject's Social Security Number.

For incident breach vectors, the Hack, Web and Laptop vectors are still the same three top vectors, and their order has not changed. The Hack vector increased from 106 to 123; the Web vector went from 94 to 112; while the Laptop vector changed from 44 to 51.

The records vectors also remain the same in their membership and ranked order. The Hack vector was 3.6 million; the Web vector was 1.1 million; and the Drive/Media vector was only 484,000.

Figure 65: Student Incident Breach Vectors (2005 - 2010)

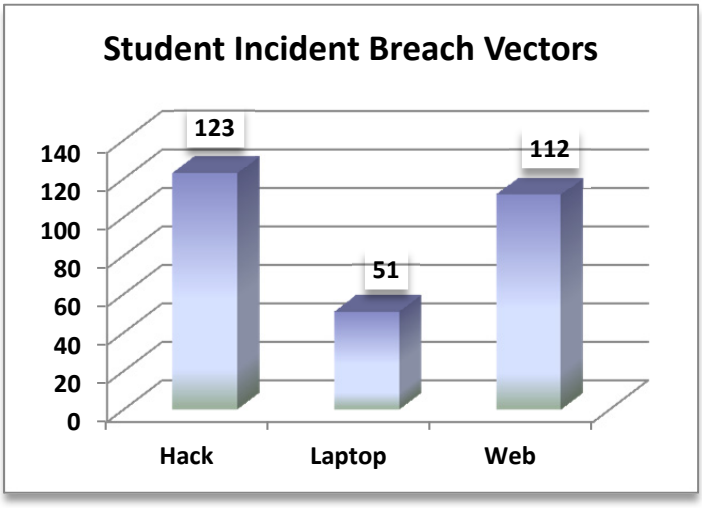
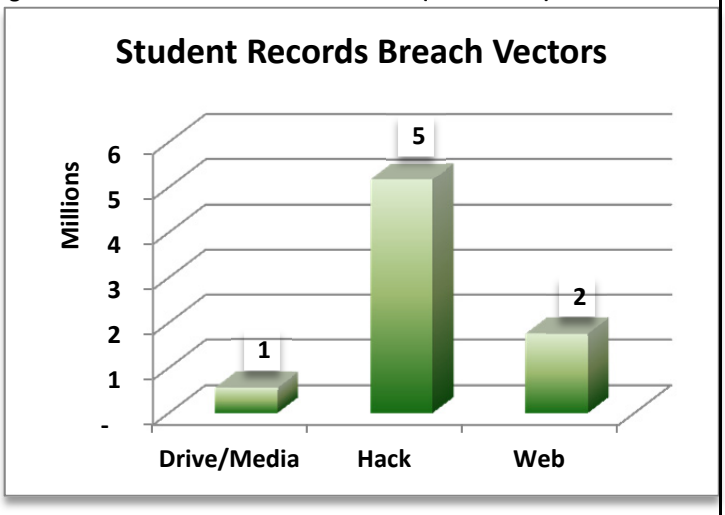


Figure 66: Student Records Breach Vectors (2005 - 2010)



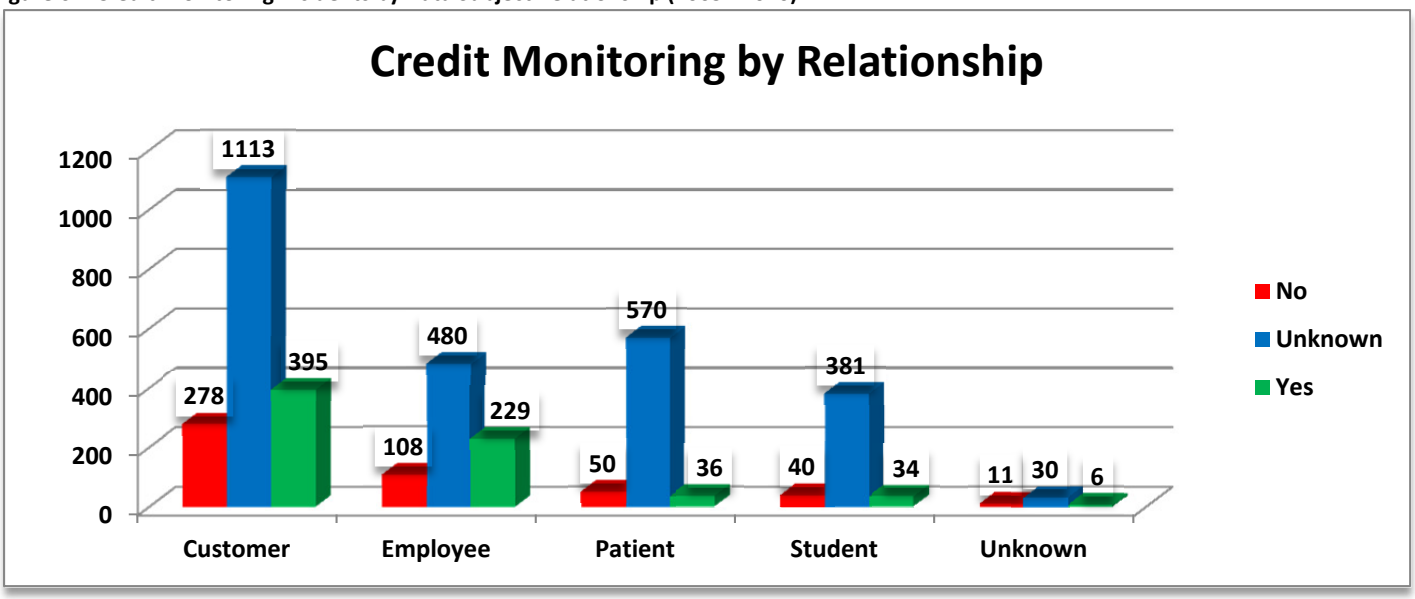
The median records disclosed for student data is 2,400. There were a total of 455 incidents and almost 9 million Student records disclosed in the course of the study.

Credit Monitoring

Earlier, we looked at whether credit monitoring services were offered by organizational sector. This time, we're looking at this service offering based on the relationship to the data subject.

In most cases where the answer to credit monitoring services being offered is known, companies are choosing to offer this service to the data subject victims. For Customer data, the ratio of yes:no is 1.42 for TLV it was 1.47. For Employee data, the current ratio is 2.12; in TLV it was 2.01. For Patient data the ratio is 0.72; while in TLV it was 0.75. For Student data it is 0.85; in TLV it was .068.

Figure 67: Credit Monitoring Incidents by Data Subject Relationship (2005 - 2010)



The higher the value in this ratio; the more cases where organizations are offering credit monitoring services. So when this figure decreases, such as with the Customer data, it means that there were more organizations offering this service in TLV than there are now. The number of "No" values increased faster than the number of "Yes" values, which decreased the ratio. The unknown records are not considered in these calculations.

Cost

In the previous report, the estimated total cost of breach incidents in the dataset was over $139 billion. The same calculations have been applied with the recent data, including the 2010 information.

Table 14: Estimated Cost of Data Breaches/Year

Year	Records Disclosed	Cost Per Record	Total Breach Cost
2005	68,555,563	$138.00	$9,460,667,694.00
2006	80,377,865	$182.00	$14,628,771,430.00
2007	164,813,878	$197.00	$32,468,333,966.00
2008	182,707,769	$202.00	$36,906,969,338.00
2009	261,759,494	$204.00	$53,398,936,776.00
2010	48,080,863	$204.00*	$9,808,496,052.00
Total	**806,295,432**		**$156,672,175,256.00**

***Cost figure from 2009.**

As with TLV, the latest year's per record figure is not yet published by the Ponemon Institute. The 2009 figure is being used in place for the purposes of this estimate. The 2009 figure was not published at the time of the release of TLV, and has since become available. It has been incorporated into these calculations [12].

The prior report estimated the total cost of breaches in the 2005 – 2009 timeframe to be over $139 billion. With the additional records for that timeframe since initial publication, the total has been increased to over $146.8 billion. As shown above, add in the records for 2010, and the total grows to over $156 billion.

The above table gives us the estimate in terms of those incidents with finite numbers of records disclosed associated. Since we were able to estimate the additional records that may be involved based on the median figures in Table 6, we can now use this data to get an estimate of the change in the cost figure. These are shown below.

Table 15: Estimated Cost of Data Breaches/Year with Median Estimated Records

Year	Records Disclosed	Cost Per Record	Total Breach Cost
2005	69,440,995	$138.00	$9,582,857,310.00
2006	81,318,959	$182.00	$14,800,050,538.00
2007	166,205,264	$197.00	$32,742,437,008.00
2008	183,409,547	$202.00	$37,048,728,494.00
2009	263,043,348	$204.00	$53,660,842,992.00
2010	48,666,013	$204.00*	$9,927,866,652.00
Total	**812,084,125**		**$157,762,782,994.00**

***Cost figure from 2009 [12].**

As before, these estimates only include the costs incurred by the breached organization. They do not include those suffered by the data subjects or the other companies that may have consequences due to a breach at another organization.

CONCLUSIONS AND RECOMMENDATIONS

Information security is a broad topic, and no single document can hope to provide recommendations that will suit every situation. However, a number of recommendations were made throughout this document, and they are summarized here for convenience.

1. Know where your data is from inception to disposal. If you do not know where it comes into the organization, where it is transformed, stored, shared with outside parties, archived, and finally how it is disposed of—you cannot hope to keep it secure.
2. Trace each sensitive data type from when it is created, to when it is disposed of, and all the places it is used in between. Without making these types of data flow maps, organizations are operating on an incomplete risk picture.
3. When a laptop is issued to an individual, it should be accompanied by a set of rules for the custodian of the device to follow. This should include direction for maintaining physical control offsite and onsite, as well as fallback controls for when these rules either are insufficient to keep the asset safe.
    a. Information Security professionals who can influence the contents of their organization's awareness training should be lobbying to have something included about laptops.
    b. In no case should the laptop be left overnight in the vehicle—particularly at the employee's residence.
    c. Do not neglect physical controls to protect electronic data. The number of laptops stolen from offices illustrates the need for locking mechanisms for the laptops when unattended at work.
4. Organizations should either put controls in place that notify when a device is tampered with, or have regular inspections of point of sale devices, gas pumps and ATMs to mitigate this risk.
5. Attention should also be given to the use of production data in test and development environments, since those environments typically have less stringent security controls in place.
6. Organizations must "bake" security controls into contracts with third party partners. This means the Information Security personnel should be involved early in the selection and vetting of potential business partners where sensitive data is concerned.
7. While prevention of malicious activity is ideal, detection is critical to minimize the damage of an incident, regardless of actor.
8. If organizations are still using SSNs as their unique identifier, they should be taking steps to eliminate them wherever possible. Reducing the locations where this highly sought after data element is stored will only help to reduce the risk of their disclosure. Data masking and encryption should be considered in cases where they must be stored and used.

This report is the second in The Leaking Vault series, and over time the data should become more precise and reliable as the number of data points in the study increases. The breach vectors should help show where the highest risk resides, and thus help organizations determine where best to spend their limited security budgets.

Until we have a federal breach disclosure law, the data sources will continue to be from disparate sources and have variable metrics available. Organizations dealing with cross border privacy laws have an even greater challenge, and that complexity will likely continue to increase. In the meantime, organizations will continue to struggle to comply with this maze of requirements for reporting data breaches as best they can.

REFERENCES

[1] Attrition.org. (2008). DLDOS (Data Loss Database - Open Source). http://attrition.org/dataloss/dldos.html

[2] Baker, W., Hutton, A., Hylender, C.D., et. al. (2011). 2011 Data Breach Investigations Report. Verizon RISK Team.

[3] Britec09. (2009). Kon-Boot Lets You Bypass Logon for Windows and Linux. http://www.youtube.com/watch?v=2lr7SYER8x4&feature=fvst (About 5,130 results for "bypass windows password" on youtube.

[4] Chickowski, E. (2011). Shortened Breach Disclosure Periods Could Hurt Consumers. Dark Reading. http://www.darkreading.com/database-security/167901020/security/news/231002965/shortened-breach-disclosure-periods-could-hurt-consumers.html

[5] Edwards, C., Kharif, O., & Riley, M. (2011). Human Errors Fuel Hacking as Test Shows Nothing Stops Idiocy. Bloomberg. http://www.bloomberg.com/news/2011-06-27/human-errors-fuel-hacking-as-test-shows-nothing-prevents-idiocy.html

[6] Florencio, D. & Herley, C. (2011). Sex, Lies and Cyber-crime Surveys. The Tenth Workshop on Economics of Information Security (WEIS 2011). George Mason University.

[7] Identity Theft Resource Center. (2011). Identity Theft Resource Center 2010 Breach Report. http://www.idtheftcenter.org/ITRC%20Breach%20Report%202010.pdf

[8] Lynn, W. (2010). Defending a New Domain, The Pentagon's Cyberstrategy. Foreign Affairs Magazine. http://www.foreignaffairs.com/articles/66552/william-j-lynn-iii/defending-a-new-domain

[9] Maryland Office of Attorney General. (2011). Maryland Information Security Breach Notices – 2010. http://www.oag.state.md.us/idtheft/breachNotices2010.htm

[10] Moker, Kevin. (2008). Sound Assurance Incident Report Database. http://www.soundassurance.com/docs/incident_report_executive_summary_details.pdf

[11] Open Security Foundation (2011). DataLossDB. http://datalossdb.org/.

[12] Ponemon Institute, LLC. (2011). Sixth Annual U.S. Cost of a Data Breach Study. Symantec Corporation.

[13] Ponemon Institute, LLC. (2009). How Global Organizations Approach the Challenge of Protecting Personal Data. Accenture.

[14] Ponemon Institute, LLC. (2010). Privacy & Data Protection Practices, Benchmark Study of the Financial Services Industry. Compuware.

[15] Ponemon Institute, LLC. (2011). The State of USB Drive Security. Kingston. http://www.kingston.com/secure/PDF_files/MKP_272_Ponemon_WP.pdf

[16] Privacy Rights Clearinghouse. (2011). A chronology of data breaches reported since the choicepoint incident (list). http://www.privacyrights.org/ar/ChronDataBreaches.htm.

REFERENCES (Cont.)

[17] State of California Department of Consumer Affairs/Office of Privacy Protection. (2011). State Security Breach Notification Laws. National Conference of State Legislatures. http://www.ncsl.org/programs/lis/cip/priv/breachlaws.htm

[18] United States Department of Health & Human Services. (2011). Breaches Affecting 500 or More Individuals. http://www.hhs.gov/ocr/privacy/hipaa/administrative/breachnotificationrule/breachtool.html

[19] Walters, Chris. (2009). Here's What A Card Skimmer Looks Like On An ATM. The Consumerist. http://consumerist.com/2009/04/heres-what-a-card-skimmer-looks-like-on-an-atm.html

APPENDIX A: DATA BREACH VECTOR DEFINITIONS

Computer: The Computer vector involves a non-laptop computer—frequently a desktop, but potentially a server or larger computer.

Disclosure: The Disclosure vector involves data that is disclosed by the data holder. An example is the insider snooping cases where hospital personnel viewed the medical records of celebrities without a valid reason. The data was not copied, stolen or lost, but the data subject's privacy was violated. This category was added to more accurately classify this type of event.

Documents: The Documents vector involves the loss, theft or inappropriate disposal of printed material.

Drive/Media: The Drive/Media vector involves portable hard drives, memory sticks, USB sticks, CD-ROMs and any other portable storage device.

Email: The Email vector involves data that is disclosed via email—whether to the wrong person, or other concerns.

Fax: The Fax vector involves the use of a fax machine to disclose information inappropriately.

Fraud - SE/Fraud-Social Engineering: The Fraud-SE vector involves malicious activities specifically designed to gain the attacker access to data via social engineering or other misrepresentation/pretexts.

Hack: The Hack vector involves attacking an organization's systems by exploiting vulnerabilities in the system's software, hardware or networking.

Laptop: The Laptop vector involves the loss, theft or disposal of portable computers.

Snail Mail: The Snail Mail vector involves the disclosure of information via the Postal Service or other courier.

Tape: The Tape vector involves the loss, theft or disposal of data stored on tape.

Unknown: The vector was not specified in the incident reports.

Virus: The Virus vector involves the disclosure of data as a result of a computer virus infection.

Web: The Web vector involves the disclosure of information by posting it on the web—either intentionally or accidentally.