Visualization Of A Data Leak

How Can Visualization Assist To Determine The Scope Of An Attack?

I.P. Swart
Council for Scientific and Industrial Research,
University of Rhodes

Prof M.M. Grobler
Council for Scientific and Industrial Research,
University of Johannesburg

Prof. B. Irwin
University of Rhodes

Abstract—**The potential impact that data leakage can have on a country, both on a national level as well as on an individual level, can be wide reaching and potentially catastrophic. In January 2013, several South African companies became the target of a hack attack, resulting in the breach of security measures and the leaking of a claimed 700000 records. The affected companies are spread across a number of domains, thus making the leak a very wide impact area. The aim of this paper is to analyze the data released from the South African breach and to visualize the extent of the loss by the companies affected. The value of this work lies in its connection to and interpretation of related South African legislation. The data extracted during the analysis is primarily personally identifiable information, such as defined by the Electronic Communications and Transactions Act of 2002 and the Protection of Personal Information Bill of 2009.**

Keywords-component; data leakage; Internet; privacy laws; security

I. INTRODUCTION

The recent breach of South African companies' security measures has released a trove of South African personal information to the Internet. Although small by international standards when compared with breaches such as the incident that saw Sony lose an estimated 77 million records [1], the South African incident saw as much as 700000 records exposed [2]. The number seems large but does little to describe the real value of the data lost: the dumped data ranges from government, banking, mining, petroleum, transport services, management, networking, construction, education and other enterprises, making this leak a very wide impact area [2].

The aim of this paper is to analyze the data released from the breach and visualize what has been lost by the companies affected. The value of this work lies in its connection to and interpretation of related South African legislation. The data extracted during the analysis is primarily personally identifiable information, such as defined by the Electronic Communications and Transactions (ECT) Act of 2002 and the Protection of Personal Information (POPI) Bill of 2009.

The paper will look at the impact areas of visualizing data and what can be considered as personally identifiable information according to South African regulations. It further provides an overview and brief analysis of the publicly available data obtained as a result of the data leak. The paper concludes with a discussion on why it is deemed important to protect private information, and how this data leak serves as warning to South African organizations that better security controls are required.

II. THE IMPACT OF VISUALIZING DATA

Assessing the impact of visualizing data leakage is fairly difficult to determine. Based on the literature reviewed by the authors, three possible impact areas can be identified. These are listed below and discussed thereafter.

- Human emotions.
- Complex datasets.
- Processing of unstructured data sources.

The effect of human emotions is demonstrated in a recent case in America. Reporters from the *Journal News* newspaper took publicly available text-based information of all gun-owners in the New York area and plotted it to an interactive Google map to visually display the data. Although the data has been available for years, the act of visualizing the data sparked intense uproar with the newspaper receiving almost 2000 letters of protest and experiencing retaliation against their reporters [3].

Researchers are well aware that humans have difficulty grasping complex datasets without prior exposure. As such, there are several studies already underway to improve the visualization field [4]. This trend is completely understandable when the work of Grady [5] is taken into account. Her research concluded that the human brain dedicates almost 30% of its capacity to sight with the closest other sense, the sense of touch, only utilizing 8% of the brain's capacity. In her findings, Grady explains that humans are exceptionally gifted at evaluating visual input and making decisions based on what they see. Work conducted by Bertini and Lalanne [6] supports this theory with their results when they investigated the complementary role of data analysis and visualization techniques and found that human operators in data rich areas achieved a higher degree of understanding regarding their environment with the help of computer visualization techniques.

Unfortunately as Marty [7] highlights, the biggest problem with visualization is not the actual data display since there are many gifted artists available, but the processing of the unstructured data sources. To this extent researchers such as Loukissas and Mindell [8] are performing research on the viability of a visual language with syntax and semantics.

As such, the impact of the data leak in South Africa, and the visual representation thereof, is in particular difficult to measure. At the time of writing, the extent of the data leak was not yet publicly visualized, thus it is not clear how human emotions may impact people's understanding thereof. Therefore, the first impact area, *human emotions,* is not applicable in this specific instance. The second impact area, *complex datasets*, is satisfied, given that the data leak affected at least 14 different companies in various sectors of business, all with their own structure and semantic nuances; it can thus be said that it created a data rich environment. Unfortunately since the data is from various companies, the structure of the data is completely different and presents a number of analysis problems. This satisfied the third impact area, *processing of unstructured data sources.* The remainder of this paper looks into the type of information that was leaked and the potential impact that this leaked data have on various entities.

III. Personally Identifiable Information

According to South African legislation, the definition of personal information is extensive and includes the following:

- information relating to an identifiable, living natural person and, where applicable, an identifiable, existing juristic person;
- information relating to the race, gender, sex, pregnancy, marital status, national, ethnic or social origin, color, sexual orientation, religion, age, physical or mental health, well-being, disability, conscience, belief, culture, language and birth of the person;
- information relating to the education, financial, medical, criminal or employment history of the person;
- any identifying number, symbol, email address, physical address, telephone number or other particular assignment to the person;
- the blood type or any other biometric information of the person;
- the personal opinions, view or preferences of the person;
- correspondence sent by the person that is implicitly or explicitly of a private or confidential nature; and
- the view or opinions of another individual about the person and the name of the person if it appears with other personal information relating to the person or if the disclosure of the name itself would reveal information about the person [9][10].

In essence, any piece of information that can be either directly or indirectly linked to the identity of a specific person can be regarded as personally identifiable information. Due to the sensitive nature of many of these pieces of information, the first identified impact area of visualizing data, *human emotions* (refer to Section II), plays an important part. The next section presents an overview of the personally identifiable information that was leaked in this particular instance.

IV. Analysis of the Data Obtained

Team GhostShell announced on 28 January 2013 that they have obtained data leaked from South African companies. The proclaimed reason for the data leak attack was to retaliate against corruption in the South African government and to send a message that corruption will be uncovered [11]. Although the text was worded in both English and Afrikaans, it is clearly visible that Google Translate was used to translate the English text to Afrikaans. This initially gives the attack a local perception but could simply be due to a smart attacker that did his research regarding the country's languages.

The data was posted on various paste sites such as http://textsnip.com and http://everfall.com with each dataset posted on more than one paste site to create redundancy in case one paste site removed the data. As at 14 April 2013, all data sources were still available for all companies affected by the breach. Some paste sites contained a modification to the primary trunk of the data so that a less technical operator might mistakenly think the data has been removed. It should be noted that no verification of the data available was given by any of the companies involved and as such should not be trusted as 100% accurate. The focus of the paper is only to analyze the content of the purported breach, highlighting the difficulties in assessing the data leaked.

According to the analysis performed by the authors, the attack itself seems to have been conducted via Structured Query Language (SQL) injection techniques. The authors found artifacts of the software Acunetix (used to perform SQL injections) in the data and the data format of more than one of the datasets posted on the various paste sites. Acunetix has a freeware version of the software available to the public, and since it is widely used, it cannot indiscriminately lead to the identity of the attackers. Table I presents a summary analysis of the leaked data.

TABLE I. Summary Analysis of Data Available (Own Compilation)

Company Attackers Claim Information was Obtained From	Username/ Password Available	Password Hashed	Encryption Type	Data Removed	Type of Data Lost
African Reinsurance Corporation	Yes	Yes	MD5/Custom	No	Email, Username, Password
Woolworths Holdings Limited	Yes	No	N/A	No	Email, Username, Password, ID, Address, Marriage, Phone, Employment history
South African Diamond Corporation	No	N/A	N/A	No	Email
African Mining	Yes	Yes	N/A	No	Email, Title, Position, Username, Password
BEE Network	Yes	Yes	MD5/Custom	No	Email, Title, Username, Password, Address, Financial

Company Attackers Claim Information was Obtained From	Username/ Password Available	Password Hashed	Encryption Type	Data Removed	Type of Data Lost
Allied Technology International	Yes	Yes	MySQL/MD5	No	Email, Username, Password
I Llovo Boulevard (I llovo Corporation)	Yes	Yes	MySQL/MD5	No	Email, Username, Password
Genesis Insurance Brokers	Yes	Partial	MD5/MD4	No	Email, Username, Password, Phone, Correspondence
Omni ID Company Part	Yes	No	N/A	No	Email, Username, Password, Phone, Correspondence, Address
Ornico Marketing	Yes	No	N/A	No	Email, Username, Password, Phone, Address
Moolmans Africa Mining Corporation	Yes	Partial	Custom	No	Email, Username, Password
Angola's National Diamond Corporation	No	N/A	N/A	No	Email
Angola's Oil field industries	No	N/A	N/A	No	Database Structure
South African Express Petroleum	Yes	Yes	SHA-1/ MySQL5	No	Email, Username, Password
State University Part	Yes	Partial	MD5/MD4/My SQL 160 bit	No	Email, Username, Password, Phone, Address, Correspondence
Westcol College	Yes	Yes	MD5/MD4	No	Email, Username
The Inc Company	Yes	Yes	SHA-256/ Haval-256	No	Email, Username, Password, Phones
Sasol Corporation	Yes	Yes	MD5/MD4	No	Email, Username, Password
Kenyan Business Directory	No	N/A	N/A	No	N/A
Algerian Government website	Yes	Yes	MD5/MD4	No	Username, Password
PressOffice linked to BidOrBuy, South Africa's largest online store	No	N/A	N/A	No	Names, Phones
FreightForwarders	Yes	No	N/A	No	Email, Username, Password, Phone, Correspondence, Address
PostNet Internet Services	Yes	Partial (NedCare)	MD5	No	Email, Username, Password, Phone, Correspondence, Address

From the analysis performed and presented in Table I the following personally identifiable information was extracted by the authors:

- 5107 identifiable physical addresses.
- 18004 telephone numbers.
- 11703 unique username and password combinations.
- 19849 South African ID numbers.
- 1641 tertiary and secondary school education details.
- 45721 email addresses.
- 1632 marriage statuses.

One of the surprise findings of the data analysis was that although the breach was touted as a South African breach, it affected quite a number of companies and people outside the South African borders, as depicted in Figure 1. This has an impact on several areas. For example, although some of the leaked data might not at first glance be considered as personally identifiable information since it relates to publicly known business addresses, it should be noted that the place where the data was obtained can make a significant contribution to the value of the data. Since the leaked business addresses were located in a database and linked with a South African business address, the implied fact is that the one business would be acting as the supplier and the other business acting as the consumer. With this knowledge a competitor can already determine the potential supplier list of a competitor's business and potentially gain an unfair advantage.

When focusing more on the information related directly to South Africa, the leaked data shows that the affected companies have definite space for improvement in terms of their information security measures. For example, passwords in databases were predominantly stored in clear text and when encryption was used, it was mostly plain MD5. Most of the MD5 encrypted passwords could be decrypted with a simple Google query and translated to either "admin" or "1234". No other fields except session cookies were encrypted at all, making the theft of personal information all the more viable to an attacker. Figure 2 depicts the address information that is available in the dataset within the South African borders. The authors made use of the Google Streetview Application Programming Interface (API) to plot the data for both Figure 1 and Figure 2. This provides the ability to zoom in to street level to look at the residence of a target, further visualizing the extent of the data breach and perceived violation of personally identifiable information.

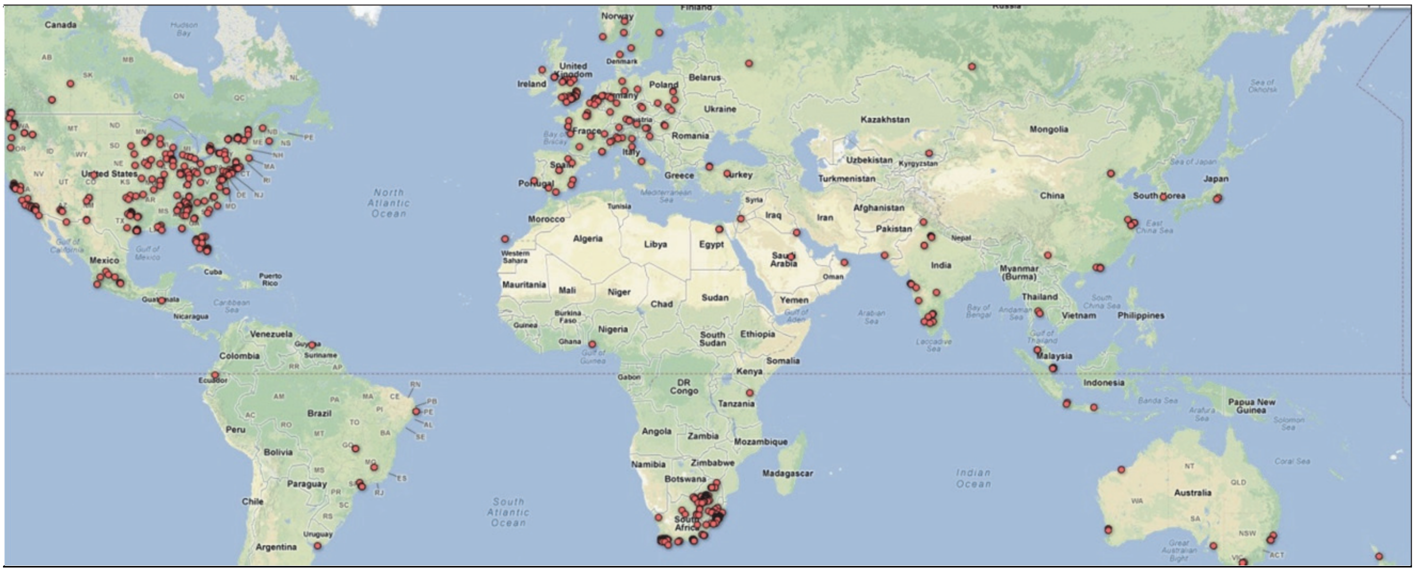

Figure 1. Worldwide Addresses Disclosed (Own compilation)

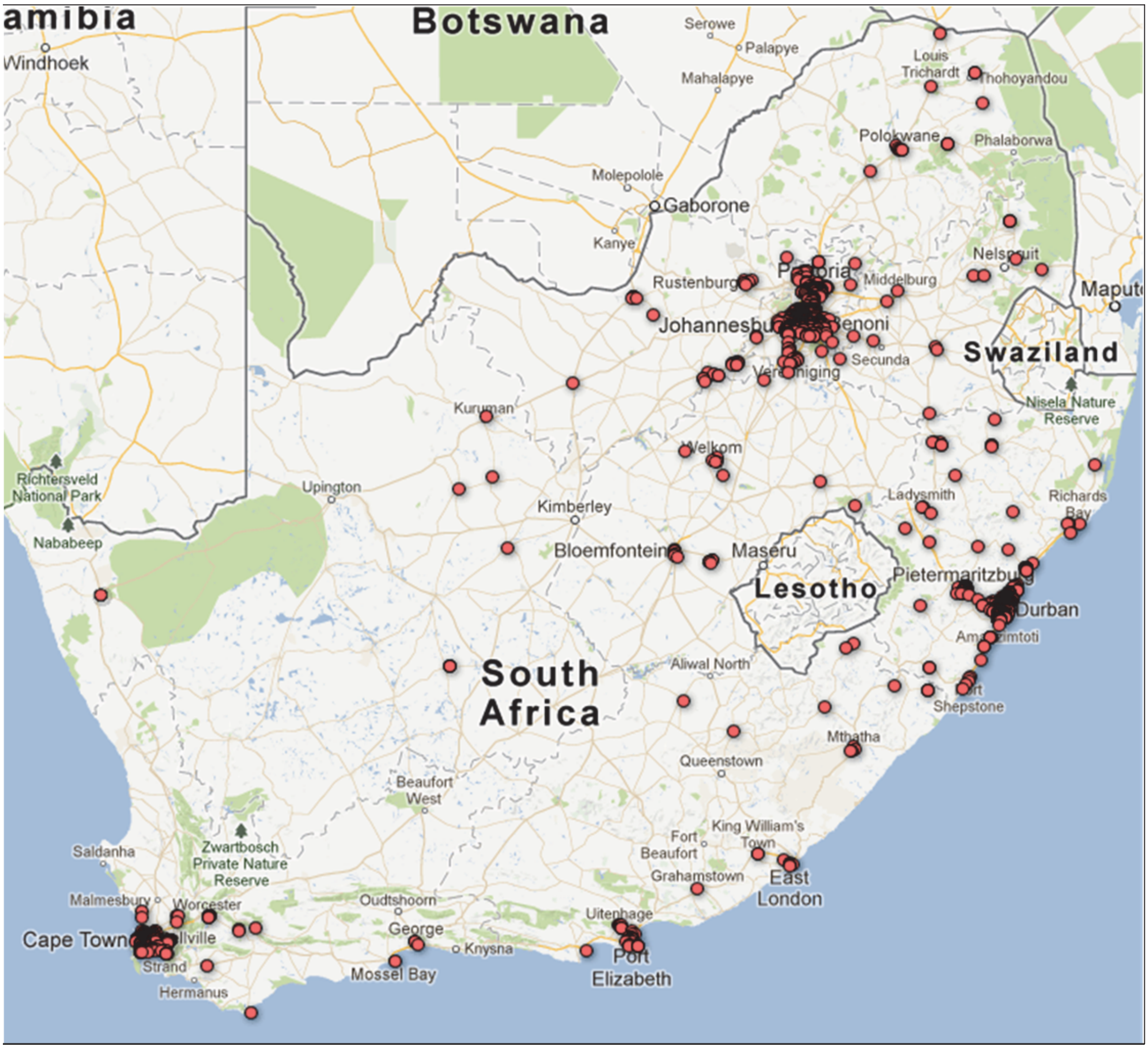

Figure 2. South African Addresses Disclosed (Own compilation)

V. Information Available from the Leaked Data

As an example of the type of information that can be deduced from a data leak of this magnitude, Table II compares the most commonly used passwords from the leaked South African data with AVG Australia's published list of most commonly used passwords [12]. Although the South African list shows skewed results (for example, the 213 instances of the password *omni-id)*, Table II provides valuable insight into South African information security practices. (This high occurrence relates directly to the dataset used - the password *omni-id* was preset as a user's default password upon account creation within one of the affected companies. Due to the volume of records from that company, it obtained first place as most popular password.)

TABLE II. Comparison of Commonly Used Passwords – AVG Australia and South African Operation Sunrise data leak (Own Compilation)

Ranking	**Commonly Used Passwords - Australia**	**Amount Found**	**Commonly Used Passwords – South Africa**	**Amount Found**
1.	password	53	omni-id	213
2.	123456	79	123456	79
3.	12345678	13	password	53
4.	abc123	4	1234	43
5.	qwerty	8	12345	23
6.	monkey	3	postnet	22
7.	letmein	6	12345678	13
8.	dragon	4	admin	11
9..	111111	6	favour	11
10.	baseball	1	marketin	10
11.	iloveyou	3	john	10
12.	trustno1	1	michelle	10
13.	1234567	4	richard	10
14.	sunshine	5	P@ssw0rd	9
15.	master	2	love	9
16.	123123	0	jesus	9
17.	welcome	2	amanda	8
18.	shadow	3	andrew	8
19.	ashley	1	qwerty	8
20.	football	0	louise	8
21.	jesus	9	martin	8
22.	michael	6	vanessa	8
23.	ninja	0	123456789	7
24.	mustang	1	1111	7
25.	password1	2	Mickey	6

The disconcerting fact is that the 25 most commonly used passwords for both Australia and South Africa are either too short or blatantly obvious. International standards recommend at least eight characters with various casing and special characters, mixed in a not-pattern like fashion. From Table II it can be seen that this is not applied in practice.

Table III shows the top 25 domain names found in the leaked dataset. Although not conclusive, the amount of unique occurrences of the respective domains raises questions regarding the accountability of the involved entities. Should the affected domains analyze the data and notify other organizations that could be impacted by the breach? Are the compromised domains responsible to inform their clients that they are victims of a data leak? Although very relevant questions about personally identifiable information, the in depth investigation of these questions are beyond the scope of this paper.

TABLE III. Most Leaked Domain Names in the South African Operation Sunrise data leak (Own Compilation)

Ranking	**Domain**	**Unique Occurrences**
1.	gmail.com	1994
2.	mail.uovs.ac.za	1680
3.	yahoo.com	1525
4.	postnet.com	1149
5.	hotmail.com	720
6.	mweb.co.za	500
7.	webmail.co.za	491
8.	telkomsa.net	450
9..	absa.co.za	332
10.	norcrossa.com	280
11.	yahoo.co.uk	212
12.	aforbes.co.za	191
13.	oldmutual.com	177
14.	ukzn.ac.za	176
15.	standardbank.co.za	165
16.	iafrica.com	164
17.	tileafrica.co.za	154
18.	tiberty.co.za	149
19.	vodamail.co.za	143
20.	absamail.co.za	143
21.	sanlan4u.co.za	114
22.	liblink.co.za	109
23.	fnb.co.za	108
24.	qwa.uovs.ac.za	102
25.	sanlam4u.co.za	98

VI. Why Protecting Personally Identifiable Information is Critical

Besides the obvious concerns regarding privacy in the modern era, several key factors mandate that the accessibility of personally identifiable information be strictly controlled. Consider that something as mundane as spam has a historic open rate of approximately 3%. In contrast, targeted spear phishing attacks have an estimated 70% open rate. In addition, research shows that spear phishing victims will actually follow links included in the phishing message approximately 50% of the time [13]. With the amount of personal information available on the Internet after a data leak of this magnitude, attackers have a lot of information that can be used to target specific people in spear phishing attacks. Thus, a data leak can actually serve as a catalyst for further cyber attacks.

Perhaps one of the most notable public instances where spear phishing has proven its effectiveness was in the breach that occurred at the RSA security corporation in 2011. According to an investigative report by Matrix Global Partners [14], the initial foothold into the company was obtained by sending spear phishing emails. The emails were crafted to target selected employees and were tailored to align the content as much as possible with the individual. With leaked data available on the Internet, this becomes relatively easy to do.

Identity theft is another thriving criminal activity and it is not surprising that in a country that has a staggeringly high unemployment rate, South Africa has also seen a huge increase in that area. Recent statistics place the estimated loss due to identity theft in the region of R1 Billion per year [15]. Once again the obvious benefits for attackers from a data leak are that the imposter can assume the identity of the victim and with as much information as possible available on the Internet after a data leak, the odds significantly reduce the chance of detection. Taking into account the information available in the leaked dataset, the imposter can impersonate any of the individuals in possession of a degree and with reasonable certainty obtain a position or credit at a less than vigilant organization. This is an illustration of just one scenario, but the scope for identity theft is significantly increased by personally identifiable information that is made available on the Internet.

One of the main problems with guarding personal information is that the information is so intrinsically linked; one piece of information can lead to the disclosure of various other pieces of information. Many online companies can provide credit, education and employment history checks for a nominal fee via either a cell phone or ID number (for example, visit http://demo.traceps.co.za). This ability to link various pieces of information is due to the South African Regulation of Interception of Communications and Provision of Communication-Related Information Act (RICA) requiring that cell phones and ID numbers be linked for identification purposes. Similarly it is possible for law enforcement to determine all required information such as financial records due to legislation such as the Financial Intelligence Centre Act (FICA). Although this aids law enforcement, it also contributes to the need to protect all personal information. Figure 3 displays the intrinsically linked information that is available from a South African ID number, using just the Internet. Although the online services do prohibit the use to gather information regarding a third party, it is doubtful that determined ID thieves would be deterred.

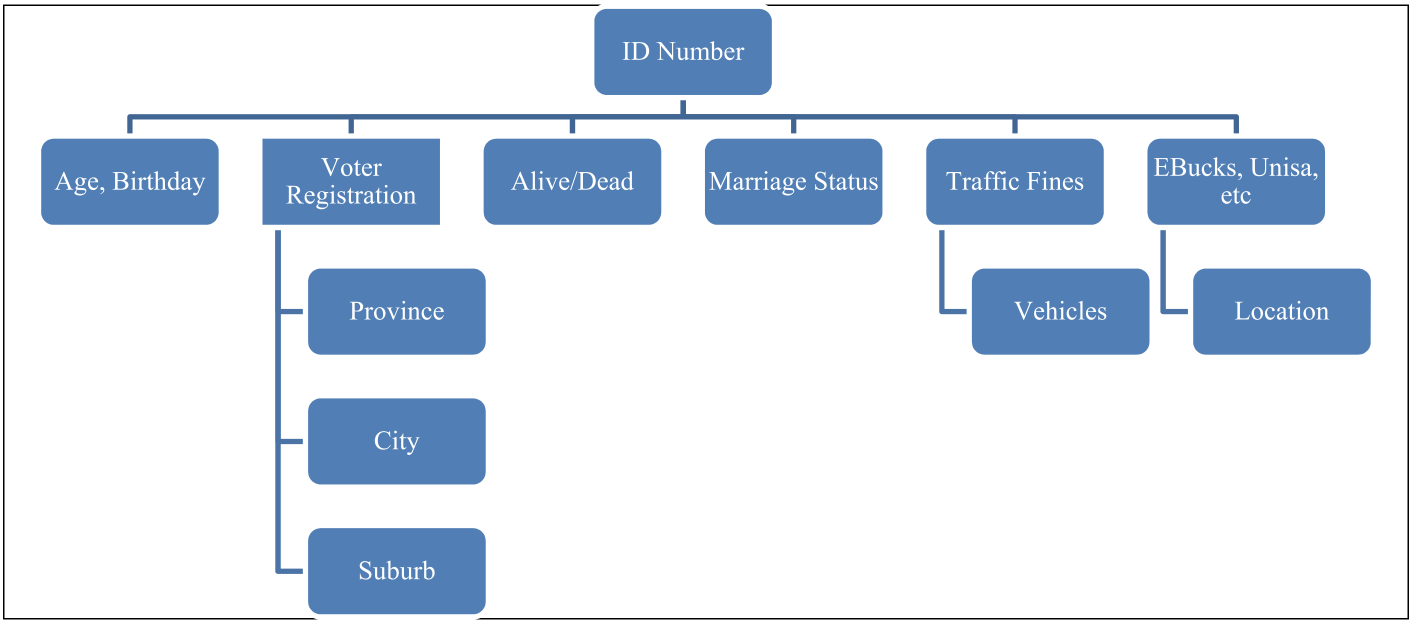

Figure 3. Basic relationship of personal information to South African ID number

Consider research by Florencio and Herley [16], proving that the average user shares the same password across 3.9 different accounts. Add to the equation an attacker that has the ability to, with the aid of personally identifiable information, determine what services an individual has subscribed to and the scenario for estimated loss changes significantly (both in monetary value and potential identity compromise). Even if just one of a user's accounts combined with personally identifiable information has been compromised, the impact could potentially be felt across a number of related accounts, compromising additional sensitive information (this is taking into consideration Florencio and Herley's observation that users share the same password across multiple accounts; thus, if one account is compromised by means of a password breach, there are potentially another three accounts for the same user with the same password).

With the release of the POPI Bill of 2009, it is no longer a case of common sense to protect personally identifiable information but it will also be a legal requirement. The Bill clearly specifies what is considered personally identifiable information (refer to Section III), the penalties involved for transgressing the law and security requirements for a company storing personal information [9]. In this sense, losing/leaking data is expensive both from a holistic view (refer to Figure 3) and a monetary perspective. According to research done the cost of data breaches where personally identifiable information is involved has increased significantly to $3.94 per record at the lowest estimate and to $194 at the highest with an average breach cost of $3.7 million, up from $2.4 million in 2001 [17]. In addition, the estimated revenue lost due to reduced customer loyalty after a data leak varies by country but ranges from $289060 in India to $3 million in the USA [18]. At the time of writing, no official figures for South Africa could be obtained to compare the costs of data breaches.

VII. Lessons to be Learned

There are a number of lessons to be learned from the South African data leak of January 2013. These lessons can largely be mapped against the three areas identified that can potentially be impacted by the leak (refer to Section II).

The first identified impact area relates to human emotions. As stated in Section II, at the time of writing the extent of the data leak was not yet publicly visualized, thus it is not clear how human emotions may impact people's understanding thereof. However, as part of the exploratory process, the authors mapped the timeline (refer to Figure 4) from the date of the data leak to the date that the data was removed from the Internet (refer to Table 1, column *Data Removed*). This is done for the available datasets and is not a reflection on the companies involved in the breach. Due to the nature of personally identifiable information leaked, the authors assume that the extended period of time during which the information remained available on the Internet may act as a catalyst for human emotions. It should be understood that even though the data is removed from the Internet, the attackers that placed it there could still have access to the data. The possible attack footprint is reduced by removing the public data but it is not removed completely.

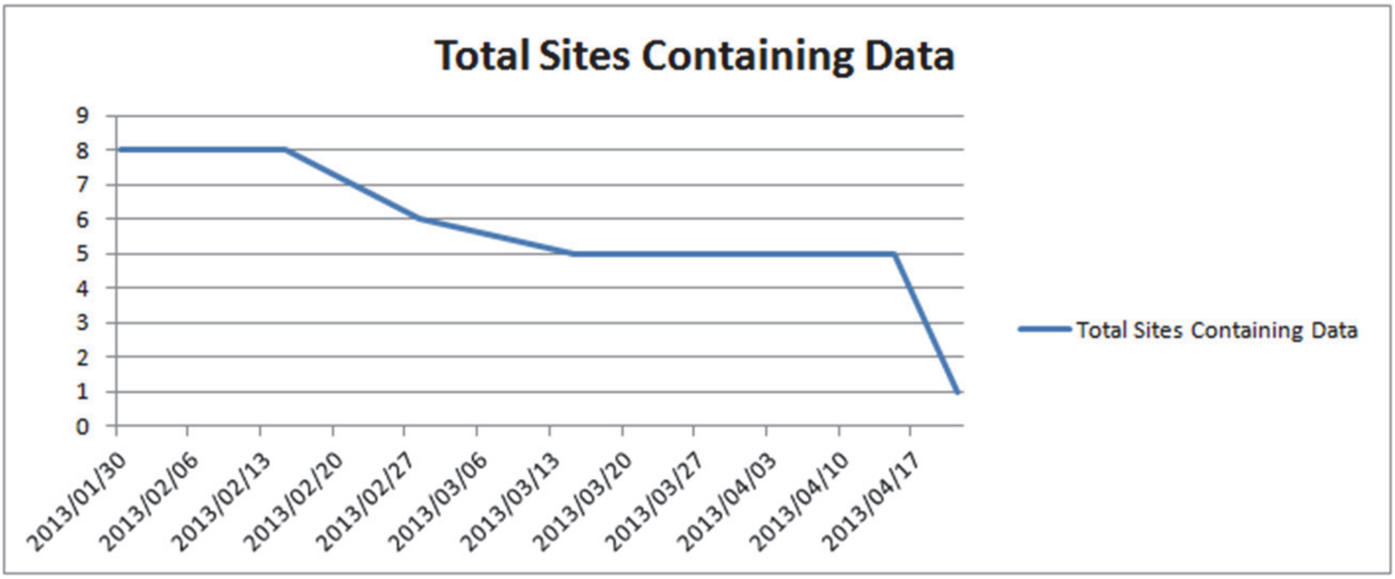

Figure 4. Timeline of datasets' availability on the Internet

The second identified impact area relates to complex datasets. Not only does this paper show that all domains are at risk (the leaked data affected government, petroleum, banking, construction, management, networking, mining, transport services, education and other enterprises), but it raises a number of questions regarding the personally identifiable information and the responsibility of those entities tasked with handling personally identifiable information. The intrinsically linked nature of personally identifiable information, as presented in Figure 3, touches on the complex nature of these specific datasets.

The third identified impact area relates to the processing of unstructured data sources. This area emphasizes the important role that legislation plays when working with personally identifiable information – special care should be taken to ensure that South African legislation provides unambiguous guidance on personally identifiable information, associated responsibilities (of compromised companies, affected domains and victims) and measures of protection.

Cachin and Schunter [19], amongst other researchers, have shown that the rate of attack on companies might have decreased a bit, but the volume of data leaked with each attack has grown tremendously. Research institutions such as Deloitte [20] and Cibecs [21] respectively present surveys stating that 53% of companies are not ready for POPI and 26% of companies have not even started to prepare. Attackers are thus getting better at extracting information from leaked datasets while the victims are still struggling to understand the legal requirements presented to them in the form of South African legislation. With the data from the relatively small South African attack presented in this paper, it remains an open question if both business and administrators are ready for the new laws being signed into effect.

VIII. Conclusion

The data leak affecting several South African companies has had a significant impact on the way that the security of personally identifiable information can be perceived. Compared to international security breaches, the South African data leak is relatively small and contained. However, it provides a number of lessons to be learned from.

The information and analysis presented in this paper is the result of research and interpretation performed by the authors. The findings and assumptions are based on information that was retrieved from publicly accessible websites. The content of the datasets were not verified with the affected companies, and as such none of the findings or assumptions made in this paper are binding or place an obligation or blame on the companies involved. The findings are the interpretations of the authors, based on publicly available information of the data leak. The aim of this paper was to loosely analyze the data released from the breach and to visually present the extent of the compromise to aid in furthering the importance of implementing and maintaining adequate information security measures, as well as awareness towards the responsibility associated with handling personally identifiable information.

Although the nature of this research is exploratory, the paper provides insight into the value of leaked personally identifiable information on the Internet. It further touches on the contribution that South African legislation has on the intrinsically linked nature of personally identifiable information, and what the resultant impact can be if this type of information is made available on the Internet.

References

[1] Hanlon, T. 2011. *PlayStation Network hacked, personal information of 77 million accounts accessed.* Available from: http://www.gizmag.com/playstation-network-hacked/18501/ (Accessed 15 April 2013).

[2] Selvan, S. 2013. *#ProjectSunRise: Team GhostShell leaked 700000 accounts from South Africa.* Available from: http://www.ehackingnews.com/2013/01/projectsunrise-team-ghostshell-leaked.html (Accessed 15 April 2013).

[3] Carbone, N. 2012. *Outrage after New York paper posts map of gun owners' names and addresses.* Available from: http://newsfeed.time.com/2012/12/26/outrage-after-new-york-paper-posts-map-of-gun-owners-names-and-addresses/(Accessed 14 February 2013).

[4] Chen, C. 2005. Top 10 unsolved infromation visualization problems. *IEEE Computer Graphics and Applications.* 25(4):12-16.

[5] Grady, D. 1993. *The vision thing: Mainly in the brain.* Discover magazine. Available from: http://discovermagazine.com/1993/jun/thevisionthingma227#.UWvKcLX-FCg (Accessed 10 April 2013).

[6] Bertini, E. & Lalanne, D. 2009. *Surveying the complementary role of automatic data analysis and visualization in knowledge discovery.* Proceedings of the ACM SIGKDD workshop on visual analytics and knowledge discovery. Pp 12 – 20. Paris, France. ACM.

[7] Marty, R. 2009. Applied security visualization. Addison-Wesley, 2009.

[8] Loukissas, Y. & Mindell, D. 2012. *A visual display of sociotechnical data.* CHI'12 Extended abstracts on human factors in computing systes. Austin: ACM.

[9] Protection of Personal Information Bill. 2009. Available from: http://www.justice.gov.za/legislation/bills/B9-2009_ProtectionOfPersonalInformation.pdf (Accessed on 2 February 2013).

[10] Electronic Communications and Transactions Act. 2002. Available from: http://www.info.gov.za/view/DownloadFileAction?id=68060 (Accessed on 2 February 2013).

[11] Cyber War News. 2013. *Break down and insight into project sun rise African data leak. Available from:* http://www.cyberwarnews.info/reports/break-down-and-insight-into-project-sun-rise-african-data-leak/ (Accessed 14 February 2013).

[12] AVG. 2012. *World's 25 worst passwords revealed – Is yours one of them?*. Available from: http://wwwresources.avg.com.au/security_risks/worlds-25-worst-passwords/ (Accessed 16 February 2013).

[13] FireEye. 2012. *Spear phishing attacks – Why they are successful and how to stop them*. Available from: http://www.computerworld.com.au/whitepaper/370771/spear-phishing-attacks-why-they-are-successful-and-how-to-stop-them/download/ (Accessed 20 March 2013).

[14] Matrix Global Partners. 2011. *RSA breach: Analysis and protection recommendations*. Available from: http://www.matrixgp.com/Files/StormShield/RSA_Whitepaper_Aug_2011.pdf (Accessed 10 April 2013).

[15] Hoffman, S.K. & McGinley, T.G. 2011.*Identity Theft.* ABC-CLIO. 2011.

[16] Florencio, D & Herley, C. 2007. *A large-scale study of web password habits*. Proceedings of the 16th international conference on World Wide Web. Pp 657- 666. Banff, Canada. ACM.

[17] NetDiligence. 2012. *Cyber liability & data breach insurance claims, A study of actual payouts for covered data breaches*. Available from: http://www.netdiligence.com/files/CyberClaimsStudy-2012sh.pdf. (Accessed 6 April 2013).

[18] Ponemon. 2011. *Global cost of data breach study.* Available from: http://www.symantec.com/content/en/us/about/media/pdfs/b-ponemon-2011-cost-of-data-breach-global.en-us.pdf (Accessed 10 April 2013).

[19] Cachin, C. & Schunter, M. 2011. A cloud you can trust. *IEEE.* 48(12):28-51.

[20] Deloitte. 2012. *Why SA companies should take heed of the Protection of Personal Information Bill.* Available from: http://deloitteblog.co.za/tag/protection-of-personal-information-bill/ (Access 3 March 2013).

[21] Cibecs. 2012. *Survey results: Only 26% of South African companies preparing for POPI Bill.* Available from: http://www.itweb.co.za/office/cibecs/PressRelease.php?StoryID=234936 (Accessed 3 March 2013).