Perception of Data Privacy in Digital Forensic Investigation

Robin Verma and Gaurav Gupta

robinv@iiitd.ac.in, gauravg@iiitd.ac.in,
Indraprastha Institute of Information Technology Delhi, India

Abstract

The use of digital technology in conventional as well as new age crimes is increasing throughout the world, and researchers are working to develop digital forensics techniques to investigate such crimes. Digital forensic investigation process requires the digital forensic investigator to manually examine the forensic image of the seized storage media. The investigator gets full access of all the data contained in the forensic image including the accused or the victim's private or sensitive data that may be entirely unrelated to the given case. The unrestricted access on the seized media or its forensic image becomes a significant threat to the accused or the victim's data privacy as the investigator can view or copy the data at will. There is no legal or technical infrastructure in place to stop such abuse. The paper presents a study containing three different surveys for the three stakeholders in the digital forensic investigation and aim to capture their perception of the accused or the victim's data privacy. The surveys were circulated in India among the digital forensic investigators, cyber lawyers and the general public respectively. These surveys included questions related to the privacy of digital data present on the storage media during the course of digital forensic investigation and subsequent trial in court of law. The surveys collected 15, 10 and 1889 responses from participants of respective target audience classes. The responses show lack of professional ethics among investigators, lack of legal support for lawyers to protect data privacy during digital forensic investigation and confusion among the general public regarding their data privacy rights. The findings would help in justifying the need for privacy preserving digital forensic investigation framework that protect privacy during the digital forensic investigation process without compromising on efficiency and performance.

Keywords: Data Privacy, Digital Forensic Investigation, Survey

1. Introduction

Privacy is a very complex concept to define, as one can have various definitions of privacy depending on the context where it has to be used. Privacy can be considered as a tool that enable an individual to control access to his personal space [1]. In the digital world, an individual's personal space is much larger than his personal space in the physical world. The personal objects of an individual in the digital world comprises of the digital information in the form of files on his digital devices and other places. Digital forensics investigation aims to find pieces of evidence that link a malicious activity carried out using a computer or any other digital device, to the person responsible. Digital forensic investigators usually search the entire storage media while investigating a particular case to find all possible pieces of evidence that may help in solving it. The point when investigators stop the investigation is usually when sufficient number of case relevant evidence get collected. While investigating a case, a digital forensic investigator has access to all data and files present on the seized storage media. Apart from the potential evidence files, the storage media also contain its owner's private data like personal or family photographs, videos, business plans, emails, medical documents, financial details to name a few. Investigator's access over case irrelevant files including owner's private file is a significant threat to owner's data privacy. The authors have addressed the same problem of data privacy breach of the owner during a digital forensic investigation and subsequent trial of the case in court of law. The authors wanted to collect ground truth about this data privacy problem in the field of digital forensics by circulating survey questionnaire among the three parties involved in a digital forensic investigation, namely the investigator, the cyber lawyer and the general public. The third party general public represents the accused, and the victims who's storage media is seized for investigation. The surveys are focused on the Indian context of digital forensic investigation, as all survey participants are from India. Researchers have made attempts to provide privacy in digital forensic investigation. Law et.al [2] proposed a cryptographic model where the data owner was allowed to encrypt the storage media and run indexing over it. The investigators could perform a keyword search over this with the help of encryption keys the owner used before. Croft et.al [3] proposed a privacy protecting layering model restricting investigators from accidental infringement of owner's privacy. The layers act as a guiding channel for the investigator to cross-check the information he is accessing. Other solutions that came later also used cryptography to protect privacy of owner's data in different use case settings [4][5][6]. The authors have carried out three surveys to understand the notion of privacy in the digital forensic investigation from the perspective of the digital forensics inves-

☆

tigators, cyber lawyers and general public. The three surveys focus on finding out how the privacy of data on storage media belonging to a victim/ accused is catered during the digital forensic investigation process as well as the trial in court. In the first survey authors have taken the views of digital forensic investigators about the privacy of data contained in forensic images of storage media belonging to the victim/accused involved in a given case. The survey came out with surprising findings that investigators do not pay respect to the privacy of accused/victim's data during the investigation process. The respondents have accepted that case irrelevant private data is often viewed and sometimes even copied from a case image. Similarly the second survey that aims to capture cyber lawyers' views on how privacy of data on digital forensic images belonging to accused/ victim are treated during the proceeding of the respective case(s) in court of law also reveals some interesting results. Two of the respondents have said that they know instances where the accused/victim has reported misuse of information gathered during investigation to threaten him. The third survey however captures responses from the general public about what their perception of privacy of the data they store on their digital devices. The responses show that people are either confused or unaware about protecting their data privacy in case their storage media is seized by law enforcement agencies. All the detailed findings related to the three surveys are stated in section 4, 5 and 6 respectively. To the best of authors' knowledge, this is the first study to collect investigator's, lawyer's and general public's perception of data privacy during the digital forensic investigation. The responses have been collected from all three stakeholders to present a comprehensive insight into the problem of privacy breach in digital forensic investigation process. The survey has collected responses from India, however, the findings show valid concerns confronting the global digital forensic community. The findings would help in justifying the need for a new privacy preserving digital forensic investigation framework that protect privacy during the digital forensic investigation process without compromising on efficiency and performance. In the rest of the paper, the word 'investigator' in place of digital forensic investigators, the word 'lawyer' in place of cyber lawyers and the word 'investigation' in place of a digital forensic investigation have been used. Similarly, the word 'investigator survey' in place of the first survey with digital forensic investigators as target audience, the word 'lawyer survey' in place of the second survey with the cyber lawyers as target audience and the word 'public survey' in place of the third survey with general public as target audience have been used. The words private and personal are used interchangeably throughout the rest of the paper.

2. Methodology

The three surveys were hosted by the authors to target respective audience within India. Surveys research is a well established field in computer science. It has been used by researchers in digital forensics too to understand target audiences' thought process and perceptions[10][11]. The understanding gained by analyzing the survey results can help researchers to explore new directions and solutions for a particular problem or field.

Authors' first step of survey design included personal interviews with one candidate each for the investigator survey and the lawyer survey. Afterwards five potential candidates were also interviewed for the public survey. The interviews helped authors to frame focused questions for the respective surveys.

The second step included converting the insights gathered during the interviews into set of objective questions for the three surveys. The respective surveys were then hosted online on the survey hosting website surveymonkey.com. The initial set of questionnaire was again shown to the interview candidates, and their the feedback on question formulation was collected. The feedback helped the authors to improve the relevance, readability and comprehensiveness of the respective surveys. The investigator survey's questionnaire flow is divided into three subsections based on grouping of similar questions; i) Following the forensic procedure; ii) Suitable time to stop the investigation and evidence gathering; and iii) Accessing accused/victim's private files. The lawyer survey questionnaire is divided into four sub-sections; i) Minimum number of evidence requirement; ii) Investigation of one case lead to prosecution of the other; iii) Accused/victim asked for privacy of their data; and iv) Misuse and threatening. On the same lines, the flow of questions in the public survey is divided into two sub-sections; i) Gathering general attitude towards privacy of data and personally identifiable information; and ii) Checking awareness about the digital forensics and the investigation process.

The third and final step included sending the online link of the respective surveys to their target audiences. The investigator and the lawyer survey went online in the month of August 2013. The last entry in the investigator survey was received in the month of January 2014, and that for the lawyer survey was received in the month of February 2014. The public survey went online on early September 2013, and the last entry received was in August 2014. The paper includes analysis of the entries received till August 2014 for the three respective surveys. The researcher believes that these three surveys cover a holistic picture of privacy of data during the process of digital forensic investigation and following trial in court of law. The surveys include all of the three parties involved in a computer forensic case, namely the accused/ victim from the general public, an investigator and a lawyer.

3. Demographics

The investigator and the lawyer survey includes participants who are experts from the respective fields. The investigator survey respondents include digital forensic investigators working on real life cases. They have experience of working in criminal cases as well as corporate investigations. In the investigator survey, a total of 15 digital forensic investigators filled in their responses. There is a bigger participation from the private sector, 11 out of 15 respondents are from privately owned digital forensic labs or companies. Rest 4 investigators are working for government forensic labs. 10 out of 15 have a degree in computer science, and the rest 5 are from different background. 7

out of 15 have less than 2 years of working experience in digital forensics field; another 4 are working as investigators for the last 2 to 5 years. Rest 4 have an experience of 5 to 10 years in the field. The table 1 shows the number of cases the participants have solved during the course of their career as an investigator.

The lawyer survey respondents are working as cyber lawyers in reputed courts in India including the Supreme court of India. 5 out of 10 participants work with a privately owned firm including one participant who owns a law firm. Another 3 participants work as independent law consultants and the remaining 2 work with a government agencies. The experience of participants as cyber lawyers, in a number of years, is presented in table 2. The number of computer fraud and cyber crime cases solved or handled by the participants while the course of their respective career is shown in table 3.

Table 1: Number of digital forensic cases solved as an investigator

Number of cases Solved	Percent of responses	Actual responses (out of 15)
Less than 10 cases	40.00%	6
Between 10 to 30 cases	13.33%	2
Between 30 to 50 cases	20%	3
Between 50 to 70 cases	13.33%	2
Between 70 to 100 cases	6.67%	1
More than 100 cases	6.67%	1

Table 2: Experience as Cyber lawyer

Experience (in years)	Percent responses	Actual responses (out of 10)
0 to 2	40%	4
3 to 5	20%	2
6 to 8	20%	2
8 to 10	0%	0
10 and above	20%	2

Table 3: Number of digital forensic cases solved as cyber lawyer

Number of cases Solved	Percent responses	Actual responses (out of 10)
Up to 10 cases	30%	3
Up to 25 cases	30%	3
Up to 50 cases	10%	1
Up to 75 cases	0%	0
More than 75 cases	30%	3

In the public survey 1235 participants filled the complete demographics; 654 people quit before reaching the demographic questions. There are 4% of the respondents who are under 18 years old, 61.1% of the respondents are between 19 and 24 years old; 17.8% of them are between 25 and 34 years old, 6.8% of them are between 35 and 44 years old and 10.3% of them are above 45. The number of male respondents are 66.6% and rest 33.4% are female.

17.2% of the respondents have 0-4 years of computing devices usage experience, 21.5% have 4-6 years of experience, and 61.4% of the respondents have more than 6 years of experience. The demographic of the survey show that the participants are well educated and have sufficient experience of using computing devices. The authors framed a hypothesis that participant's level of awareness about various privacy issues related to digital documents would be high. This was proved otherwise when all the response were compiled later. Due to a shortage of space the authors have removed most of the tables and figures from the survey, and used text to explain findings in as concise manner as possible.

4. Privacy from Investigator‘s Perspective

The aim of this survey is to assess how digital forensics investigators cater privacy of the accused/victim's data on the seized storage device. The survey's questionnaire was designed after taking inputs from a pilot interview with one investigator. The categorization of the questionnaire is discussed in the following subsections.

4.1. Following Procedure: Chain of Custody(CoC)

Chain of custody is a legal document that is used to track the evidence from the time of its seizure from the incident scene till it is presented in court of law or it is released back to the owner. CoC contains information about the artifact being seized, like the brand name, color, type and other physical and technical parameters, combined with name and position of people who handled it with the time of custody. It is maintained in order to fix accountability and bring fairness in the process of digital forensic investigation and subsequent trial in court.

Author's intention to put the first two questions on CoC is to check if the investigators are well versed to the basics of their trade. The questions also aim to evaluate investigator's level of seriousness towards their job by asking whether they follow CoC regularly or not. When asked about do they follow CoC in cases 14 out of 15 respondents said they fill the CoC form and only one of the respondent said he do not know about it. 11 out of these 14 respondents follow CoC in all of the cases at all times, however the other 2 said they follow it most of the times not always and the last candidate said he follows CoC only sometimes.

When asking the 14 who answered with a yes in previous question about what encourages them to follow CoC, 2 out of the 14 said they follow CoC only if the case is going to court of law. Another 3 out of the 14 said they follow CoC only if they think the case is important enough, whereas the rest 9 said they follow CoC in all cases irrespective of the case going to court of law or it is an internal corporate investigation.

4.2. Suitable time to stop the investigation

The next multiple choice question asks investigators whether they stop after finding case related potential evidence or they have a tendency of further exploring the forensic image increasing the chances of encountering owners personal file which are not relevant to the case. 8 out of 15 responded that they consider to stop after gathering all possible evidence present on the storage media that include those related to the case as well as those that are not related to the case, but can be used to prosecute the owner of the media in a new case. the next 6 out of 15 responded that they would stop their investigation after they have gathered all the evidence present in the storage media that are related to the given case. And lastly only 1 out of the 15 said he would stop investigation after he has gathered a minimum amount of evidence needed to prove or disprove the given case.

In another subsequent multiple choice question it was asked whether the investigators ever experienced a situation where they get hold of some proofs/ potential evidence for other activities not related to the case which can be used in forming another separate case against the accused/ victim. Surprisingly, 7 out of 15 (nearly 50%) responded with a 'yes, most of the times', another 4 out of 15 (26%) also responded with a 'yes, only sometimes'. Only the last 4 out 15 said that they do not get evidence for new fresh cases while investigating for a particular case. Combining the responses of both of these above stated questions, we can infer that gathering potential evidence that are not even related to the case under investigation is a regular practice among digital forensic investigators. The habit of exploring more evidence than required opens ground for privacy breach of data belonging to the accused/ victim.

4.3. Accessing accused/victim's private files

The Next question of the second survey, asked digital forensic investigators what do they do when they encounter owners private files (like personal photographs, videos, business plans, or other intellectual property) during the investigation process. 6 out of 15 said they would view all such files, copy files which are related to the case under investigation as well as some other files which are not related to the case but are illegal in nature. Other 4 out of 15 said they would view and copy all private files because these files have more probability of becoming evidence files in potential cases, including the case in hand and all other possible cases. The rest 5 out of 15 said that they would view all private files, but copy only those which are related to the case under investigation. The results show that all of the participant investigators accepted that they access users private files which may or may not be related to the case in hand. Moreover, 10 out of 15 say they copy users private file if they find it related to the respective case or illegal in some way.

Another interesting question which asks the digital forensic investigator whether they have seen any forensic investigator, in their laboratory or elsewhere, who while investigating a given case copies files like wallpapers, songs, movies, software games or commercial software from the case image under investigation. 3 out of 15 replied that they have seen investigators doing the same in their own laboratory only. Another 4 replied that they have seen investigators doing it but in some other forensic laboratory. Only one responder replied that he has not seen anyone copying, however he does not see any problem if one does so. 7 out of 15 on the other hand replied by saying that they have not seen it anywhere, and feel that it is not a good thing to do. Surprisingly half of participant investigators have seen other fellow investigators in their laboratory or elsewhere copying user content from the forensic images seized for investigation purposes.

The last question before demographics asks the investigators whether they have heard of any incident where the accused/victim has reported misuse of information, potential pieces of evidence gathered during the case investigation, to threaten the accused/victim. One responder out of 15 said he knows about such a reported case. 9 out of 15 respondents replied with a no, saying they haven't heard of any misuse of information like this during their carriers. Rest 5 respondents are not sure if such a thing can even happen.

5. Privacy from Cyber Lawyer‘s Perspective

The aim of the lawyer survey is to get insights into legal aspect of how privacy is catered during a digital forensic investigation and the trial of respective case in court of law. The authors conducted pilot interview of one cyber lawyer who is currently working in the supreme court of India which helped in framing a comprehensive questionnaire for the survey. The grouping of the questions is represented by the following sub sections.

5.1. Completion of case

The first question which the survey asks cyber lawyers is that in a case of a Cyber Crime and Computer Fraud, at what time does their preparation for a case is complete. 7 out of 10 answered that their work for a case is completed after they have gathered all possible pieces of evidence present on the storage media of digital devices including evidence related to the case as well as those that are not related to the case, but can be used to prosecute the owner of the media in a new case. 2 out of 10 said they would stop after they have gathered all pieces of evidence related to the case. Only one of the respondent said he would stop after he has collected minimum number of evidence.

In the question that followed the researchers asked what is the minimum number of evidence which are sufficient to prove or disprove a Cyber Crime and Computer Fraud case in the court of law. 4 out of 10 responded that '1 or 2' number of evidence are sufficient. 3 respondents said '3 to 5' evidence are enough whereas the remaining 3 said '6 to 10' evidence per case are required to get a conclusion. The responses to this question are important because they set an upper limit on number of evidence for majority of digital forensic cases. At maximum 10 evidence are sufficient for a digital forensic case, which actually starts with seizure of digital devices containing hundreds and thousands of files. Rest of the files in the seized digital device are irrelevant to the case and may be labeled as private to the accused/victim.

5.2. Investigation of one case lead to prosecution of the other

The next question asked in survey three wants to verify the results from survey two which asks the investigator that while investigating a particular case does he get hold of some proofs/ potential evidence for other activities not related to the case which can be used to file a new separate case in favour or against the accused or victim. While asking the same question from cyber lawyers make sense as all the evidence collected by digital forensic investigators are compiled and used by cyber lawyers before the case goes to court of law. 1 out of 10 respondent said he always get such situation where the evidence collected can be used to start a new fresh case against the accused/ victim. 5 respondents say they also get the same results most of the times, followed by 3 respondents who say that they too get similar multipurpose evidence sometimes. Only one respondent replied in a no.

5.3. Accused/victim asked for privacy of their data

After this the next three questions ask specifically about three privacy supportive laws written in the Constitution of India or the Information Technology Act 2000 & 2008amendment. The first one asks the respondents about how many cases in their career as a cyber lawyer have they handled in which a PC/ laptops or Smartphone/ tablet/ PDA etc. were seized and the accused or victim have applied for right of privacy referring to either the freedom of speech and expression under Article 19(1) (a) or right to life and personal liberty under Article 21 of the Constitution of India, or both. 5 out of 10 said they have experienced 10 such cases, 3 of the respondents said they have seen 10 to 30 of such cases. One of the respondent each have observed 30 to 50 and 50 such cases. The second question asks for the same scenario where investigative agencies are accused for privacy breach under section 72A of the (Indian) Information Technology Act, 2000. [i.e. the accused/victim accuses the agencies for accessing and disclosing their private information, which is irrelevant to the case being investigated. Example, the access and/ or disclosure of personal/ family photographs and videos, when the person is being investigated for a financial fraud.]. 6 out of 10 answered in a yes, and the rest 4 answered in a no. The positive responses range from 2 to 5 instances of such cases.

The third question asks for the same scenario where the investigative agencies are accused for privacy breach under section 43A of the (Indian) Information Technology Act, 2000. [i.e. the accused/victim accuses the agencies for improper or negligent handling of their sensitive personal data or information during investigation of the case]. 6 out of 10 answered in a yes, and the rest 4 answered in a no. The positive responses range from 1 to 5 instances of such cases.

Another question following the same lines asks cyber lawyers about how many cases they have solved and others they have knowledge about where the accused or victims have asked/ requested the court to preserve their private data or files on their seized digital devices. 3 out of 10 responded that they have seen up to 10 such cases, 2 of them said they have seen 10 to 20 cases, 1 of them said he has seen more than 90 cases. Rest 4 out of 10 said that they have not seen a case till date, where the accused or victim asked to preserve his/her private data or files on the seized digital device(s).

5.4. Misuse of information for threatening

The subsequent question which asks the cyber lawyers whether they have heard of any incident where the accused/victim has reported misuse of information (the gathered pieces of evidence after the completion of case investigation) to threaten the accused/victim, also got interesting responses. The question is exactly similar to what researchers asked to forensic investigators, in section 5.4. Two responder out of 10 said they know about such a reported case (one of them has seen 2 such cases, the other has seen 1). 3 out of 10 respondents replied with a no, saying they haven't heard of anything during their carrier about such a cases. Rest 3 (as 2 of the remaining 5 skipped the question) respondents are not sure if such a thing can happen. The results are again encouraging as there are two responses which accept misuse of information gathered during digital forensic investigation, which otherwise is not reported in general.

6. Privacy from General Public‘s Perspective

After successful acquisition of the exhibit disk, the digital forensic investigator has full access over the data on disk. The accused or the victim has no way to be assured that the private data on his/her disk that is not related to the case is not accessed by the investigators. For example, if a person is suspected of a financial fraud, then his family holiday trip photographs and videos, that are not related to the case should not be accessed during the investigation. This is a privacy breach and should not happen during a digital forensic investigation. The results of the investigator survey show participants accepting that they have seen investigators coping non-relevant data including private data from accused/victim's acquired images. The results of the lawyer survey show participants accepting that they have seen cases where the accused or victim has been threatened by the investigative officer using the case non-related data gathered during investigation. Keeping all these findings in mind the authors designed the public survey to assess the people's sensitivity towards privacy of their own data. The author also created a hypothetical scenario where participants had to assume that their data storage device has been seized by law enforcement agencies for investigation, and framed questions in the survey to check if people's sensitivity towards the privacy of their data gets effected. The public survey questionnaire can be broadly divided into two sub sections that are explained below:

6.1. General attitude towards privacy of data and Personally Identifiable Information (PII)

The questions in this sub-section aim to understand how people handle their private data. What all files do people consider private and where so they store them. Another parameter on which level of privacy can be measured is access control, so some of the questions are focused on passwords management. The protection of Personally Identifiable Information (PII) is another dimension of privacy in digital world. Remaining questions in the subsection collect responses on PII protection.

6.1.1. Storage of personal information on digital devices/places

The first question asks the participants how frequently they store private data, that is defined in the question as data one would not like to make public, on digital devices they own or use. 70.3% people store their private data on mobile phones, 75.1% store on laptops and 54.9% store on desktops. Talking of external storage devices, 45.4% people store their private data on portable hard disks and 58.1% store on pen drives. The percentages stated above are obtained by adding up values from the sometimes, usually and always responses.

Table 4: Digital Devices lost in the past five years

Device	Percentage
Mobile	33%
Tablet	0.7%
Laptop	3.1%
Portable Hard Disk	3.3%
Pen Drive	39.9%
None of above	41.5%

Considering the private data stored on above stated devices, losing one could be a serious privacy threat to the owner. The next question asked whether people have lost any of their digital devices in the past five years. The responses are shown in table 4. The statistics show that 59.5% people have lost at least one of the digital devices in the past five years. This high number shows either the owners are not so careful in keeping their devices secure or their device got stolen at some point of time. Valuable items like smartphones and laptops could be on target of thieves, but the loss of low cost devices like pen-drive can only be attributed to casual behavior of the owner. People take backups of sensitive data on portable storage devices like pen-drive and the survey responses show that around 40% people have lost one in the past five years.

6.1.2. Common passwords for different accounts

One in every three, 32.6% people, use a common password or pass phrase across multiple online accounts. 22.5% people prefer not to say anything in response, whereas 45% people do not use common passwords for their multiple accounts. The results show casual behavior towards security of private data stored online.

6.1.3. Storage of passwords on digital devices

The responses for this question say that 24.6% people store their passwords on either their mobiles or tablets. 25.6% people store passwords on laptop or desktop. Although two in every three, 63.9% people, do not store their passwords on their devices, the digital devices of the remaining one in every three would have their passwords stored in them. If seized for investigation one in every three device would contain stored password and out of that one in every three owner would have used common password for a variety of online accounts.

6.1.4. Personal files stored on digital devices

The aim of framing this question was to enlist a comprehensive list of private files to show what all is at risk if the devices are seized for investigation. The question asks people where do they store these private files. The output would provide a relative ranking of the devices/storage-places where people prefer to store a particular class of private files. The private files include media content like photographs, audio, video and text. Another source of private data could be digitized physical documents that people keep on their devices. The documents comes onto a digital device either when the document gets digitized using some scanning device, or a person simply clicks a digital photograph of the document.

A total of 1474 respondents answered the question and due to space limitation in the paper, only the relevant findings are enlisted in this sub-sub section. 84.27% people store their personal photographs on their laptop/desktop and 30 to 35% people store personal photographs on pen drives, portable hard disks, email accounts and smartphones each. The size of individual digital photographs is small enough to be accommodated on variable storage capacity devices that makes it the most prevalent personal file across all digital devices and online storage services.

Other prominent files stored on laptop/desktop include 69.13% video, 62.89% audio, 43.55% bank statement, 50.34% air/railway/bus booking, 67.30% mark-sheet/degree/admit card, 71.91% resume/ CV/ biodata, 58.28% job offer letter, 49.25% passport scan, 52.51% Permanent Account Number card (a identity card provided by income tax department of India), 44.17% Aadhar card (biometric identity card issued by Government of India), 45.39% birth certificate and 46.13% voters identity card (issued by Elections Commission of India). For every type of private file specified in the question, a person's laptop or desktop stores the highest percentage of them as compared to every other device or online storage service. This finding endorses our hypothesis that a person's laptop and desktop tend to contain a lot of private data whose privacy is at stake if it gets seized for a digital forensic investigation.

6.1.5. Rating personal files

The next two question were framed to obtain a relative ranking of the personal files and Personally Identifiable Information (PII) respectively. The participants were asked to assign a rank to the entries on the scale of 1 to 5, where 1 is for the least important and 5 is for the most important. The reason for collecting the ranking of a one's personal files and PII is to identify and collect the top entries. The authors ask survey participants in later questions to assume the hypothetical scenario where their storage devices get seized by law enforcement agencies and rethink about their rankings. The authors are interested to know whether the hypothetical scenario encourages people to change their previous rating about the personal files and the PII.

The first question that is discussed in this sub-sub section talks is about ranking of the personal files; and it got responses from 1474 people. After adding the values of rating 5 and rating 4 for every entry, 63% of the respondents rated personal photographs as important. Other notable example are 61.2% bank statement, 72% mark-sheet/admit card, 61.9% resume/CV, 62.8% job offer/ appointment letter, 68.6% passport, 73.10% PAN card, 65.60% Aadhar card, 76.9% credit/debit card, 68.3% license, 71.4% voter ID and 62.7% birth certificate.

6.1.6. Rating Personally Identifiable Information (PII)

This question aims to get a relative ranking among various Personally Identifiable Information (PII). 1287 people finished the survey and assigned a rank to given PII on the scale of 1 to 5, where 1 is for the least important and 5 is for the most important. 70.39% people find their full name to be important, similarly 67.66% people rated fathers name as an important PII. 61.53% people consider Mothers maiden name an important PII.

Other notable entries include 68.6% date of birth, 68.5% residential/ office address, 74.6% phone numbers, 72.2% email address, 68.3% passwords, 69.9% bank details, 72.3% Permanent Account Number, 65.6% passport number, 63.6% license number, 64% Aadhar card number, 62.2% ATM PIN number, 69.1% biometrics.

6.2. Awareness about digital forensics investigation

The questions in this sub-section aim to understand how people's perception about their private data, that they disclosed in the sub section before, changes in a hypothetical situation where their digital devices get seized by law enforcement agencies. One can expect a radical shift in a person's privacy ratings for his private data when it is secure with him versus when the security agencies seize his devices to investigate some case. The shift in privacy perception would be indirectly proportional to the trust the people have in law enforcement agencies. The possible change in perception would also depend on the individual's awareness about digital forensic investigation process and the fact that most of the digital forensic tools can find hidden and deleted data.

6.2.1. Belief in law enforcement agencies

People have firm belief that law enforcement agencies would not misuse their data, in case they seize it for a hypothetical investigation. The values present in table 5 show that 56.21% and 53.45% people would have no effect on the privacy ratings or they would be rather less concerned about the privacy of their data in case of personal documents and PII if the law enforcement agencies seize their devices.

6.2.2. Awareness about digital forensics

When asked whether law enforcement agencies have tools to recover deleted data, 32.21% people said they are not sure if it is possible and other 20.25% people said they don't believe that deleted data can be recovered. Only 47.4% people said they know that deleted data can be recovered by the law enforcement agencies. store their data temporarily on their office digital devices and deleted them after use. Even after nearly half of the people know that the deleted data is recoverable, in next question that followed 40.95% people said that they temporarily store their personal information on their office devices before deleting them.

Table 5: Rating when law enforcement agency acquire device

Data Type	No Effect	May Increase	May Decrease
Personal Document (1304 responses)	47.3%	43.8%	8.8%
Personally Identifiable Information (1304 responses)	46.7%	46.6%	6.8%
Online Social Networks (1286 responses)	52.9%	36.5%	10.6%

7. Legal status on privacy in Digital Forensic Investigation

The authors have also tried to find the current legal status of the term 'privacy during digital forensic investigation' in light of the legal systems of the Republic of India and the European Union. The Information Technology (IT) Act 2000 and later the IT (amendment) Act 2008 are the primary laws passed by the government of India to deal with cybercrime and electronic frauds at large. The initial version of the law, the IT Act 2000, defined all possible cybercrimes in legal terms and also proposed punishments of each of the individual offences. The IT (amendment) Act 2008 broadened the scope of the law by enhancing some of the definitions and provisions initially introduced in the IT Act 2000 [7]. The edits included usage of the word 'electronic signature' in place of digital signature to widen its scope, introduction of the term 'corporate responsibility' for data protection, defining the 'intermediary' to include service providers under the law, just to name a few [7]. However, neither the IT Act 2000 nor its evolved version IT (amendment) Act 2008 talks about the possibility of an investigator misusing his power by accessing private data that is totally unrelated to the investigation or about the rights of an accused/victim to ask for privacy of his/her data.

Europe has a reputation of being the most privacy-sensitive region in the world. In case of cyber laws the European countries follow the Council of Europe's Cybercrime Convention (Council of Europe Treaty Series (CETS) 185) [8] supplemented with the cybercrime convention on racism through computer systems (CETS 189) and the Lanzarote convention on the protection of children against sexual abuse (CETS 201). European Countries follow the treaties however implement the law in accordance to their respective legal structure. All of the European countries respective legal implementations define the types of cybercrime and related offences [9]. However, the issue of privacy of accused/victim's data under a digital forensic investigation has not been mentioned anywhere. Moreover, there is no clause regarding the misuse of power by investigator like their Indian counterpart. It appears that the investigator's work ethics is highly trusted both in India and the European Union.

8. Conclusion

The paper presents the results and analysis of three surveys aimed at gathering perception on data privacy during digital forensic investigation from the three stakeholders involved namely, the investigator, the lawyer and the general public. The analysis of results shows a lack of professional ethics among the investigators, lack of legal structure to check privacy abuse during investigations and lack of awareness in general public regarding their privacy rights.

The investigator survey results show that, in order to ensure completeness of the investigation and gather as many potential pieces of evidence as possible, more than half of the investigators almost invariably access the private files of the accused or victim on the seized image. Moreover, the survey findings also show that some investigators copy the accused or victim's private data from the case image for their personal use. These practices by the investigators are privacy abuse to the accused or victim's data, and show lowering work ethics among them. To prevent privacy abuse of accused or victim's data during digital forensic investigation process, the survey findings justify for the requirement of a new framework that protects privacy of data without compromising on completeness and efficiency.

The lawyer survey results indicate towards the requirement of privacy protection laws fixing accountability of the investigators in case of an abuse. The current cyber laws in India and across the world should include data privacy violations during the digital forensic investigation. And lastly, the general public results show that there is a lack of awareness among people regarding their personal data privacy. The results also show that people the

To the best of author's knowledge, this is the first study to collect investigator's, lawyer's and general public's perception of data privacy during the digital forensic investigation. The responses have been collected from all three stakeholders to present a comprehensive insight into the problem of privacy breach in digital forensic investigation process. Although the survey has collected responses from India, the findings show valid concerns confronting the global digital forensic community.

9. Limitations and Future Work

It should be also noted that in the public survey nearly 654 out of 1889 (34.6%) of the respondents did not finish the survey. Some of the participants who quit the survey in between pointed out in their comments that the survey was too exhaustive and too long. Although the number of participants in the investigator survey (digital investigation experts) and the lawyer survey (cyber law experts) are limited, yet the collected responses are valuable enough due to the expertise level of these participants in their respective fields. As future work, the authors wish to explore new ways of incorporating privacy of data into the to the standard digital forensic investigation model. The new privacy preserving mechanisms should fit in the existing models, and should not compromise on completeness of the investigation and the efficiency and performance of the tools.

10. References

1. Moore, Adam. ”Defining privacy.” Journal of Social Philosophy 39.3 (2008): 411-428.
2. Law, Frank YW, et al. ”Protecting digital data privacy in Computer Forensic Examination.” Systematic Approaches to Digital Forensic Engineering (SADFE), 2011 IEEE Sixth International Workshop on. IEEE, 2011.
3. Croft, Neil J., and Martin S. Olivier. ”Sequenced release of privacy-accurate information in a forensic investigation.” Digital Investigation 7.1 (2010): 95-101.
4. Hou, Shuhui, et al. ”Privacy preserving confidential forensic investigation for shared or remote servers.” Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2011 Seventh International Conference on. IEEE, 2011.
5. Hou, Shuhui, et al. ”Verifying data authenticity and integrity in server-aided confidential forensic investigation.” Information and Communication Technology. Springer Berlin Heidelberg, 2013. 312-317.
6. Hou, Shuhui, et al. ”Application Of Secret Sharing Techniques On Confidential Forensic Investigation.” The Second International Conference on Cyber Security, Cyber Peacefare and Digital Forensic (CyberSec2013). The Society of Digital Information and Wireless Communication, 2013.
7. Seth, Karnika. ” IT Act 2000 vs 2008- Implementation, Challenges, and the Role of Adjudicating Officers” National Seminar on Enforcement of Cyberlaw , New Delhi 2010
8. De Hert, Paul, Gloria Gonzlez Fuster, and Bert-Jaap Koops. ”Fighting cybercrime in the two Europes.” Revue internationale de droit pnal 77.3 (2007): 503-524.
9. Casey, Eoghan. Digital evidence and computer crime: forensic science, computers and the internet. Academic press, pp - 123-129 2011.
10. Ruan, Keyun, et al. ”Survey on cloud forensics and critical criteria for cloud forensic capability: A preliminary analysis.” Proceedings of the 2011 ADFSL Conference on Digital Forensics, Security and Law. 2011.
11. Ruan, Keyun, et al. ”Cloud forensics definitions and critical criteria for cloud forensic capability: An overview of survey results.” Digital Investigation 10.1 (2013): 34-43.