SACHVERSTÄNDIGEN-BÜRO FÜR COMPUTERWESEN
PROF. DR. PAUSCH & PARTNER

Büro Darmstadt: 64289 Darmstadt, Heinheimer Strasse 38
Tel: 06151/9712640 Fax: 06151/9712641
Büro Kassel: 34277 Fuldabrück, An der Röthe 10
Tel: 0561/95339100 Fax: 0561/95339101
Büro Grünstadt: 67310 Hettenleidelheim, Im Park 9
Tel: 06351/1359000 Fax: 06351/1359001
Büro Pegnitz: 91257 Pegnitz, Reisach 16
Tel: 09241/7359000 Fax: 09241/7359001
Büro Wiesbaden: 65197 Wiesbaden, Geschwister-Scholl-Str. 26
Tel/Fax: 0611/2046273
Office Sydney (G+R IT-Experts): Level 6, Spring Street, 2000 Sydney NSW
Phone: 02 82960492 Fax: 02 82960411

Report 140222/02

Evaluation of the System *MaverikMonitor*

On behalf of

Maverickeye UG (haftungsbeschränkt)
Herrn Yanick Gabriel
Heilbronner Straße 150
70191 Stuttgart

Expert witness

Dr. Simone Richter

von der Industrie-und Handelskammer Darmstadt
öffentlich bestellte und vereidigte Sachverständige
für das Sachgebiet Systeme und Anwendungen
der Informationsverarbeitung

Data Privacy Declaration:
All information that were received for the purpose of preparing this report are stored on digital media. They will solely be used for the intended purpose described in this report.

1. Introduction

1.1 I was engaged as an expert by Maverickeye UG (haftungsbeschränkt) represented by Mr Yanik Gabriel, Heilbronner Straße 150, D-70191 Stuttgart, Germany.

1.2 I acknowledge that I have read, understood and complied with Practice Note CM7.

1.3 A copy of my CV may be found at appendix A of this report.

2. Instructions

2.1 My instructions were to provide answers to the questions set out below in relation to:

(a) the software program known as Maverik Monitor Version 1.47 (**Software**); and

(b) a system which monitors, detects and stores information in respect of data transfers of copyright material on the BitTorrent network of which the Software forms a part (**System**).

2.2 The questions which I was instructed to answer are:

(a) Is the Software capable of accurately detecting and recording copies of data being shared from a Subscriber's IP Address taking place on BitTorrent networks?

(b) How does the System function?

(i) Is there likelihood for erroneous information to be recorded by the Software?

(ii) Can the information recorded by the Software be manipulated by a user?

(iii) When the System engages in a BitTorrent transaction with a remote computer, is the System able to accurately record the IP address of the remote computer with the date and time?

(iv) What is the link between hash value and IP address, and does the System record this information accurately?

(v) What procedures are in place to ensure that the IP address recorded was allocated to the correct Subscriber at the time of the alleged copyright infringement?

(vi) What is the 'anti-leech mod' and how does it affect the System's reliability?

(vii) What are dynamic IP addresses and what effect (if any) do these have on the accuracy of the information recorded by the System?

(viii) What are hash collisions and how do they affect the accuracy of the information recorded by the System?

(c) Does the System download from a source computer the full file?

3. Source of Information and Scope of Work

3.1 In preparing this report, I have reviewed and relied upon:

(a) user documentation for the Software (**User Documentation**);

(b) the source code for the Software;

(c) my visual inspection at one of the data centers where the System's information is gathered and stored; and

(d) the test results from the physical test conducted between 16 and 20 January 2014, to test whether or not the System was capable of detecting distribution of a known data set being five different movie files using the BitTorrent protocol.

3.2 For the purpose of this report, I have accepted the information provided to me as accurate, unless otherwise stated.

3.3 I worked independently in my analysis of the Software and the System as laid out in the in deriving my opinion set out in this report.

3.4 My opinions are based wholly or substantially on the specialised knowledge set out in my CV at Annexure "A" to this report.

3.5 I do not express any opinion about the accuracy of the information provided to me upon which my findings are based. The conclusions in this report depend upon the accuracy of that information.

3.6 If any information I have relied upon is found to be inaccurate or incomplete, or further information is provided to me, I reserve the right to revisit my findings.

3.7 This report makes use of technical terminology, an explanation of which is set out at paragraph 6 of this report.

4. Assumed Facts and Limitations

4.1 The NTP servers used by the System to calibrate its internal clock for the purpose of logging the date and time of alleged infringements are correct. The NTP servers are provided by local authorities and, as such, are trusted to distribute accurate time signals.

4.2 The hash value creation protocol for the copyrighted work and its sub-pieces is correct. As demonstrated with the five test files distributed over the BitTorrent Network.

5. Summary of Findings

5.1 In my examination of the System, I found that:

(a) it is capable of monitoring traffic on the BitTorrent network;

(b) it identifies the IP addresses of users and records the time of the data transfer accurately. This information enables ISPs to identify the Subscriber whose internet connection was used to conduct the alleged copyright infringement;

(c) The System implements mechanisms to ensure that no erroneous IP addresses are collected by the System.

(d) The System establishes connections with remote computers and receives and stores pieces of data from those computers which make content available across the BitTorrent network.

(e) The received pieces of data are compared to a control copy using hash value comparison. If the hash value of the piece of data downloaded from the source is the same as the control copy, then the received pieces of data and the control copy are considered to be identical.

(f) Dynamic IP addresses have no influence on the accuracy of the System logging and recording IP addresses and associated time stamp of infringement.

(g) If the source computer's IP address changes during capture of a data transfer, no data is saved by the System, and therefore not used by Maverickeye UG.

6. Terminology

6.1 This report makes use of technical terminology. The technical terms used in this report are set out below.

BitTorrent

6.2 The BitTorrent network is a type of peer-to-peer network. Peer-to-peer networks are a conglomerate of computers that link together to share information, files or data with one another through the use of specialised software (**P2P Network**). The users or computers on a P2P Network can either receive or send information, files or data to other computers (or undertake both functions simultaneously).

6.3 On P2P Networks, a connection is established between the users of the network who are online at the time. Each participating computer can perform both the function of a "client" (i.e. the receiving or downloading computer) and that of a "server" (i.e. the sending or host computer). The computers can then both send data to, and receive it from, each other's computers. The data is exchanged directly between the participating computers and is never stored in a centralised place. The data distributed may have various origins, but data exchange takes place exclusively between two individual computer systems.

6.4 Users need to take active steps to set up and install software to enable participation in the P2P Network. These steps cannot be 'accidently' or inadvertently undertaken by a user. Any user wishing to participate in file sharing needs to install or actively start specialised software known as a BitTorrent client (**BitTorrent Client**). In order to do so, a user will need to download a

BitTorrent Client from a website that distributes BitTorrent clients. As part of the installation process, a user may configure the BitTorrent Client in accordance with his or her preferences, or adopt standard settings. A BitTorrent Client enables users to access a given P2P Network, such as eDonkey or BitTorrent. Some examples of BitTorrent Clients for the BitTorrent network include Azureus, BitComet and UT.

6.5 Once a user has installed the BitTorrent Client, the user may have an option to specify what is called the user's 'shared files' folder (**Shared Folder**) in which the user may place any files. Files placed in the Shared Folder are made available and may be distributed to other users requesting that file.

6.6 A user will then need to conduct a search of torrent files related to the data he or she wishes to acquire. Such data often includes copyrighted works such as films, television shows and music. Websites such as "The Pirate Bay" may be used to search for and obtain the relevant torrent files, as they offer torrent files and magnet links for download. The user then needs to download those torrent files and open them in their BitTorrent Client.

6.7 The BitTorrent network splits or separates a complete file, being for example, a movie or song (**Complete Data Set**) into pieces to enable efficient distribution to participants. Those pieces may be further broken down into sub pieces. When those pieces are reassembled, they constitute the Complete Data Set.

6.8 Once a user opens the torrent file in their BitTorrent Client, the BitTorrent Client queries the peers to which it is connected in order to ascertain which pieces of the Complete Data Set those peers have available to download. Some peers will have the Complete Data Set, and are known as "seeders" (**Seeders**). Other peers may have less that the whole file because they are still in the process of downloading it, but they will still be able to share the pieces that they have.

6.9 Over the time of downloading the Complete Data Set, pieces are requested and received by the BitTorrent Client from various other peers and are ultimately assembled together like a large jigsaw into the Complete Data Set. If, for example, the Complete Data Set is a film file, the film file will at this stage be in a state in which a user can view it.

IP address

6.10 An Internet Protocol (**IP**) address is an address which identifies a computer within an IP network. It is comparable to a postal address in the sense that it enables computers to exchange data with each other. The commonly used "IPv4" address consists of four numbers (values 0-255) separated by a dot. In computer terms it is a 32 bit large binary number.

6.11 There are more computers connected to the internet than available IPv4 addresses. To ensure that the information sent from another computer reaches the correct addressee, a dedicated technology is used called "Network Address Translation" (**NAT**).

6.12 Using NAT, an internet access point has at least one worldwide unique IP address. The internet access point is often a router or a a digital subscriber line (**DSL**) access point. This type of equipment is used to connect computers within an internal network. They are commonly found within households or businesses.

6.13 Each device connected to the internal network may have its own IP address assigned by a router. In circumstances where an internal network is present, the IP address logged by the Software will identify subscriber's internet access point only. That is, the IP address assigned to the router or DLS access point to which a number of computers can be identified.

6.14 It is not possible to determine if an internal network is present unless inquiries are made with the internet connection owner. There could very well be one computer connected to the router.

Dynamic IP address

6.15 An Internet Service Provider (**ISP**) has control of a large number of IP addresses. It assigns IP addresses to its account holders (**Subscribers**) in order to provide internet connections. An internet connection cannot function without an IP address.

6.16 The ISP may have more Subscribers than they control IP addresses. If this is the case, it may not assign fixed static IP addresses to its Subscribers.

6.17 An ISP may assign 'dynamic' IP addresses to its Subscribers. A dynamic IP address is only assigned for a limited time period. This period may vary from ISP to ISP. The time period could be up to six months or more and is dependent on each ISPs own internal policy.

6.18 Due to the nature of dynamic IP addresses, in order for ISPs to identify Subscribers from data logged by the System, two data points must be known: time of the data transfer and the associated IP address. Using this information, an ISP may cross-reference IP address, date and time with its Subscriber database.

Hash Values

6.19 An algorithm that correlates data of variable length to data of a fixed length is called a hash function. The value returned by the function is a hash value.

6.20 Two different sets of data may be compared using hash values. They are commonly referred to as digital fingerprints. They allow large quantities of data to be represented by a relatively small number of bytes. The determination of a hash value of a data set and the comparison of two hash values is more efficient than a byte-wise comparison of two files. The System uses hash value comparison to determine if two sets of data are identical.

6.21 For example, two different movies will have two different hash values. Furthermore, every piece of the movie will also have its own hash value.

6.22 There are various methods used to calculate hash values. The most commonly used hash value calculation method is called MD5. This method is known to have a mathematical "defect".

There is a possibility of creating the same hash value for two different sets of input data. If this occurs one calls this a hash collision. This method is not used by the System.

6.23 The second most applied method is called SHA-1. It is theoretically vulnerable to hash collisions, but so far there is no method known to create a hash collision. The method known to be collision free is called SHA-512 and the third most commonly used.

6.24 The P2P Network allocates a hash value to each file that is made available for sharing, so it can easily be identified by the P2P Network participants.

6.25 The System uses the SHA-1 and SHA-512 methods. The System uses SHA-1 as BitTorrent uses this to identify data. SHA-512 is used internally by the System to verify the downloaded sub-pieces against a control copy of data.

6.26 The Software uses hash value comparison to determine if two sets of data are identical. Accordingly, if two SHA-1 or SHA-512 hash values are the same, the data compared is said to be identical.

NTP Servers

6.27 A Network Time Protocol (**NTP**) server, or NTP Stratum-1 servers, is a networking protocol which synchronises all computers on an NTP server to within a few milliseconds by reference to Coordinated Universal Time (**UTC**), which is the primary time standard by which the world regulates clocks and times.

6.28 There are difference sources for UTC, such as the Global Positioning System and WWV, a radio station which continuously transmits official U.S. Government frequency and time signals. Both of these sources of UTC provide accurate time.

Module

6.29 A module is a part of a computer program which carries out a specific function and may be used along or in combination with other modules in the same program.

Data Structure

6.30 A data structure is the location where data is stored in a program. There are several different types of data structures which are capable of storing different types of data. If the data structure is not created for a particular type of data, then it will be unable to store that data.

Transport Control Protocol/Internet Protocol

6.31 The IP address attributes to a computer a unique reference number. By this number a computer is identified. The IP does not facilitate the transfer of data itself; therefore the Transport Control Protocol (**TCP**) was invented. It is a set of rules used along with the IP to send data in the form of message units between computers over the Internet.

7. Method of Work

7.1 I examined and reviewed the User Documentation of the Software and the complete source code of the Software, line by line, and the scripts therein. In addition I set up files within the Bit Torrent network to be shared an downloaded by various users to determine if:

(a) the System accurately monitors BitTorrent traffic;

(b) the System's BitTorrent Client initiates a TCP connection with the source computers (as explained in paragraph 8.2 below);

(c) the System accurately captures the IP address, together with time stamp and the port number used by the source computer when a data transfer between the Software and source computer is successful;

(d) the System records the hash value of the data received from the source computer; and

(e) the System accurately calculates the hash value of the sub-piece received from the source computer for the purposes of conducting a comparison with it and the reference file (a copy of the data set known to be a complete copy of the copyrighted work made available on BitTorrent networks).

7.2 I analyzed the source code of the Software line-by-line in order to ascertain:

(a) the way in which data identified by the Software is processed;

(b) the correctness of the various Data Structures containing the data identified; and

(c) whether it correctly stores the data in the Data Structures and extracts the correct information from the data structures to ensure proper identification and comparison of hash values.

7.3 Finally, I examined the source code of the Software to evaluate the consequences of potential errors, including:

(a) buffer overflows, which occur when data is being stored in the Software's module to capture amongst others the data set of IP-addresses, date, time and duration of the possible copyright infringements and the capacity to store the data is exceeded;

(b) wrong variable handling, which can occur, for example, when the Software is expecting a variable (i.e. a location where temporary data is stored) to contain an integer number but instead it contains a string of text and is therefore unable to process the variable; and

(c) logical errors resulting from bugs which cause the Software to operate incorrectly and produces an unintended or undesired output.

8. Live Tests performed in the BitTorrent Network to test the System

8.1 I was asked by Maverickeye UG to test the System for accuracy by distributing Complete Data Sets using the BitTorrent Network. The purpose of this test was to conclude whether or not the System was capable of detecting distribution of a known Complete Data Set using the BitTorrent protocol.

8.2 This test was designed to be a real life test which allowed different users or peers to share data with the System.

8.3 I was given permission by the copyright owners of the Complete Data Sets distributed in the live scenario to use the Complete Data Sets in the tests described below.

9. Test implementation

9.1 I informed Maverikeye UG of the time period that the tests would take place. I did so to ensure that Maverickeye's System was in operation during the time required to conduct the test. I also informed Maverickeye UG of the names of the files I was using as part of the test, so they would know which files to search for. I did not inform Maverickeye the precise point in time the downloading/uploading of the Complete Data Sets would occur. The test was performed between January 16th and 20th 2014 (**Test Period**).

9.2 In undertaking the test, I completed the following steps.

9.3 First, I uploaded four Complete Data Sets to my desktop computer. The Complete Data Sets used in the test were the following files:

(a) prepare-loopdevices.exe, being an executable file (i.e. software that causes a computer to perform tasks);

(b) AVI_0002.AVI, being a file containing both audio and video data;

(c) AVI_0004.AVI, being a file containing both audio and video data; and

(d) Vorlesungen.zip, being a compressed file containing multiple pdf files.

(together **Test Data Sets**)

9.4 Second, I created a torrent file in respect of each of the files in the Test Data Sets using the following BitTorrent clients:

(a) μTorrent version 3.3.2;

(b) KTorrent version 4.2.0 and 4.3.1; and

(c) Transmission version 2.82.

9.5 When creating the torrent files, I had to include a tracker address. The tracker address is what announces that the file is being made available and can be accessed by those using the BitTorrent network.

9.6 The trackers used for each of torrent files are set out below:

udp://tracker.openbittorrent.com:80/announce

udp://tracker.publicbt.com:80/announce

9.7 Third, during the Test Period, I used two mobile devices, named Samsung Mobile Device 1 and Samsung Mobile Device 2, to download the torrent files for each of the files in the Test Data Sets. Both of these devices are Samsung Tablets and each have internet connections provided by different ISPs and each have different IP addresses.

9.8 At the time of each download on each of the Samsung Tablets, I noted down the following information:

(a) the device being used, that is Samsung Mobile Device 1 or Samsung Mobile Device 2;

(b) the particular file being downloaded, e.g. prepare-loopdevices.exe;

(c) the IP Address of the device being used to download the file;

(d) the date on which the download commenced; and

(e) the time at which the download commenced.

A copy of the information that I have noted down in respect of the data transfers is set out in Schedule 2 to this Report.

9.9 I ascertained the date and time of the commencement of the download by monitoring my BitTorrent Client and noting the time at which the download appeared to commence on the BitTorrent Client.

9.10 I ascertained the IP address of each device by going into the settings of each device, where the IP address of the device is stated.

9.11 I was aware of each of the ISP providing the internet connection for each of the Samsung tablets, because I contracted with each ISP in relation to each device. In any event, using the IP address, it is possible to perform an internet search to obtain the ISP which owns the IP address.

9.12 After the termination of the test, Maverickeye UG provided me with an extract from their data-base, which had detected and stored information in respect of the data transfers which oc-curred during the Test Period.

9.13 I then compared the information I had noted down about the data transfers, with the information extracted from Maverickeye UG's databases.

9.14 An example extract of the System's database provided to me by Maverickeye UG is set out below. This extract is the information obtained by Maverickeye UG in relation one download by a user of the prepare-loopdevices.exe file. Each second line gives a short explanation of the meaning of the database entries.

ExportId	4771520570
	Unique number assigned automatically by the System
ExportTs	2014-01-16 16:44:22
	Timestamp of the export into the System's database
ClientinformationId	11313558141
	Unique number assigned to the user
Date	2014-01-16
	Date of the data transfer
Time	15:32:32
	Time of the data transfer
ClientProto	Bt
	Name of the protocol used in the data transfer. Bt means BitTorrent Protocol.
ClientIp	109.85.95.91
	IP Address of the user downloading the file
ClientPort	55980
	Port number used for the data transfer
ClientDhtPort	6881
	Special Port number used within the data transfer
ClientUserName	
	The username of the particular user with their ISP (unknown by Maverickeye UG)
ClientUserHash	2D4241333330302D9676CEC37D543F1A69838FD6
	Hash value
ClientVersion	-BA3300-
	Version of the Bit Torrent Client used
FileName	prepare-loopdevices.exe
	Name of the downloaded file
TrackName	prepare-loopdevices.exe
	Name of the part of the downloaded file
FileHash	8AAE50E22398E28AC4F20E5460A312B95E693F3A
	Hash value identifying the file
FileSize	16683
	Size of the file in bytes
BtBitField	0
	Value indicating the piece to download
GeoIpIsp	Vodafone D2 GmbH
	Name of the Internet Service Provider (ISP)

GeoIpOrg	
	Information about the ISP
GeoIpCity	Andernach
	City where the ISP is located
GeoIpZip	
	Post code of the city where the ISP is located
GeoIpCountry	DE
	Country where the ISP is located. DE means Germany.
GeoIpRegion	Rheinland-Pfalz
	Further information about the location of the ISP
GeoIpLon	7.4
GeoIpLat	Geographical coordinates of the ISP location 50.4333 Geographical coordinates of the ISP location
SessionStart	16/01/2014 3:31:58 PM
	Date and Time of the start of the detection of the download (UTC)
SessionEnd	16/01/2014 3:33:34 PM
	Date and Time of the end of the detection of the download (UTC)
TransferStart	16/01/2014 3:32:31 PM
	Date and Time of the end of the transfer (UTC)
TransferEnd	16/01/2014 3:32:32 PM
	Date and Time of the end of the transfer (UTC)
SessionDuration	96
	Duration of the session in seconds
LoggerId	clientng13
	Unique number of the logging process
LoggerIp	787014805
	IP number of the logging computer
LoggerCountry	DE
	Country where the logging computer is located
LoggerLon	8,4287
	Geographical coordinates of the logger's location
LoggerLat	49,0019
	Geographical coordinates of the logger's location
TotalPeers	0
	Number of additional peers
UniqTs	2014011615
	Another time stamp indicating the beginning of activity within the program. e.g. 16 Jan 2014 at 3pm (being 1500 hours).
RelatedTitleId	6467
	Internal unique number identifying the file
OwnerId	3668
	Internal unique number identifying the owner of the file

VerifyPartOk	YES
	Indicates whether the comparison of the data against a control copy was fine
ExporterVersion	2.5
	Version number of the Module forming part of the System used to export the data into the secure database
ExporterBinCRC32	c0dcf654
	The value assigned by a cyclic redundancy check[1] to the data transferred to the database. Upon retrieval of the data from the database, the cyclic redundancy check is repeated and if the two values match, then it can be assumed that the data is correct and has not been corrupted.

9.15 The System database extracts provided to me by Maverickeye UG are set out at Schedule 3 to this Report. These extracts set out all of the downloads of the Test Data Sets during the Test Period.

9.16 The data shown in the table above is extracted from row 5 of the file "6467 – Prepare-loopdevices.xls" of Schedule 3. This data correlates with the data I noted down in Schedule 2 at the tab titled "Prepare-loopdevices", that is, the ISP name, IP address, data and times of distribution correctly matched.

9.17 I can verify that the information that I noted down in Schedule 2 correlates with the data provided to me by Maverickeye UG.

9.18 It is my conclusion that the System is able to accurately detect data transfers on the BitTorrent network when deployed to do so.

10. Inspection of the data center

10.1 I visited the data center where the System was running.

10.2 The purpose of my visit was to check that data was stored securely and could not be tampered with. From my inspection, I am confident that the safety measures put in place provide sufficient security to ensure that the data collected by the System is secure.

10.3 I also confirm that write once read many (**WORM**) tape drives are used to store data collected by the System in a secure manner. WORM technology provides non-editable data storage such that any data stored by the System cannot be altered, overwritten or corrupted.

11. Answers to the Questions

11.1 The opinions set out below are based wholly or substantially on the specialised knowledge referred to in appendix A.

[1] Mathematical method developed by W. Wesley Peterson, 1961.

Is the Software capable of accurately detecting and recording copies of data being shared from a Subscriber's IP Address taking place on BitTorrent networks?

11.2 Based on my examination of the source code presented and the tests performed, the System correctly detects and records instances of copies of data being shared from a Subscriber's IP address on the BitTorrent networks.

11.3 It accurately records the IP addresses, port numbers, time stamps and the SHA-512 hash value of the sub-pieces received by the System in a secure database.

How does the System function?

11.4 The Software consists of various modules. Each of the modules has a distinctive task including (but not limited to) tracking IP traffic, calculating hash values and recording the correct time.

11.5 When the System is initiated, it checks the operating system, interfaces, clock and memory on the computer on which the Software is installed for potential errors. If any potential errors are identified, the Software and thus the System does not function.

11.6 The System ensures that an accurate time is used by calibrating its local clock with various NTP servers. The System uses amongst others, the time signals distributed by the Physikalisch-Technische Bundesanstalt (the German national metrology institute).

11.7 The System captures all data packets being transmitted between the local BitTorrent Client and a remote one. Once the local BitTorrent Client receives information about the availability and location of distributed data, it will initiate communications with those locations.

11.8 A TCP connection is then initiated by the local BitTorrent Client in order to inform source computers that the System is interested in acquiring data. Once transfer begins, all data traffic is accurately recorded with information such as time of the data transfer, IP addresses, port numbers, and if the parties in the TCP connection wish to send or receive data. If a data transfer between the System and remote computer is successful, the IP address is logged, together with time stamp and the port number used by the source computer.

11.9 Information is logged by the System only if the TCP connection is active two seconds before and after the data transfer. This ensures that users participating in transactions for a lesser period of time are not logged. This adequately addresses the issue of a dynamic IP address changing.

11.10 The hash value of the transmitted sub-piece is also recorded by the System.

11.11 Once the connection between the System's local BitTorrent Client and the source computer is closed the data captured is transferred to a secure database. The storage medium is a write-once read only. The database is backed up once per day. The backup carries a time stamp and is secured by a digital signature.

Is there a likelihood for erroneous information to be recorded by the System?

11.12 It is impossible for the System to record erroneous information. The System uses the TCP protocol to establish a communication channel with source computers. By design, the connections established are valid. If a TCP connection is not initiated between the System and source computer then no data is logged.

11.13 The System also checks the hash values of all sub-pieces received to ensure they are part of the full file.

11.14 If any technical problems or errors, including those described at paragraph 7.3 above, occur during the data capture and logging process, the relevant process is terminated and data captured during that process is not transferred to the secure database. This means that all data transferred to the secure database and stored by the System is error free.

Can the information recorded by the System be manipulated by a user?

11.15 This may be possible in limited circumstances where an individual has access to the database and the backup files with the required security clearance and the possession of the key to the digital signature.

11.16 The risk of database manipulation is mitigated as the database records all access to it. Those records may be used to trace unauthorised access to the user account responsible. As an added security measure, the backup files are secured by digital signatures.

11.17 The System saves all network traffic between the system and the remote computer. Manipulation of this without detection is difficult to accomplish given the security measures in place.

11.18 I consider the above stated limited circumstances are impossible to achieve.

When the System engages in a BitTorrent transaction with a remote computer, is the System able to accurately record the IP address of the remote computer and date and time?

11.19 The System is able to accurately record the IP address and the associated time stamp.

11.20 When a remote computer offers to distribute a piece of data to the System, both parties must establish a secure end-to-end connection using the TCP/IP protocol. When a transfer takes place, the IP address of the remote computer is logged by the System.

11.21 NTP servers ensure that the System receives accurate time information. This allows the System to accurately calibrate its clock. The System uses its clock to produce data transfer time stamps.

11.22 Using an IP address and associated time stamp of the data transfer, an ISP may correctly identify the Subscriber allocated that IP address at a specific time by searching databases in its possession. Most ISPs will operate a data retention policy. This means that the information

sought may be deleted after a period of time. It is therefore important for copyright owners who allege that the data transfer captured by the System contains copyrighted material to obtain disclosure of the names and addresses behind the IP addresses before such deletion occurs.

What is the link between hash value and IP address, and does the System record this information accurately?

11.23 When a remote computer responds to the local client's request for data, the remote computer will begin to transfer a sub-piece of the requested file. This data transfer is monitored and recorded by the System. Once the complete sub-piece is received successfully by the System, a record in the database is created documenting that transaction.

11.24 The database record contains all information relating to the transaction between the System and source. That information includes the following:

(a) IP address of the sender;

(b) IP address of the System;

(c) port numbers used for the transfer;

(d) time stamps for:

(i) the start of the conversation;

(ii) the end of the conversation;

(iii) the start and end of the download;

(e) the SHA-512 hash value of the data received from the source computer; and

(f) the position of the sub-piece in the complete file.

11.25 Once the above information has been logged, a hash value comparison is initiated by the System. The hash value of the sub-piece received from the source computer is calculated and compared to the hash value of the complete data set of the copyrighted work stored in a reference file. If the hash values and positions are the same, the System establishes that the downloaded sub-piece must have been part of the copyrighted work. In my opinion, this is the correct finding to make in the circumstances.

What procedures are in place to ensure that the IP address recorded was allocated to the correct Subscriber at the time of the alleged copyright infringement?

11.26 The Subscriber may be identified with the assistance of the ISP responsible for the allocation of that IP address. Only the ISP stores this information. Without the ISP's assistance, a copyright owner cannot identify the Subscriber.

11.27 In the case of dynamic IP addresses, various Subscribers may have the same IP address – but during different time periods. This is the reason it is important that the correct time stamp is recorded accurately by the System once an infringement of copyright is detected.

11.28 To ensure a correct time stamp the System synchronizes its time with NTP servers, which allows the System to produce accurate time stamps.

11.29 Logs with more than a 10 milliseconds time difference are not exported by the Software for use as evidence in legal proceedings, as these have the possibility of introducing potential inaccurate logs into the database. This scenario is avoided altogether by rejecting such logs. In my opinion, these safety protocols ensure that the Software produces accurate time stamps.

11.30 As the System establishes a connection to the source for at least 2 seconds after the successful data transfer, the accuracy of the time stamp recorded for the data transfer is valid.

What is the 'anti-leech mod' and how does it affect the System's reliability?

11.31 Within a P2P Network, users exchange data amongst themselves. A BitTorrent Network is a type of P2P network. It relies on the willingness of all participants to share their data. If a client is not sending data it will be marked as a 'leecher'. The System does not log users known as "leechers". A leecher is a user who does not act as a data source. The System only enters into transactions with those users who engage in the distribution of data. In other words, the users logged by the System have not only made available data, but also engaged in an actual transaction with the System and has been recorded distributing a piece of the data.

11.32 In the operation of the System, the local BitTorrent Client is created in a way which ensures that it never transfers pieces of copyrighted files to other users. Instead, the client in the System mimics a user willing to act as a source of data but no actual transfer takes place. It does not affect the accuracy or reliability of data logging by the System.

What are dynamic IP addresses and what effect (if any) do these have on the accuracy of the information recorded by the System?

11.33 An explanation of dynamic IP addresses may be found in section 6. In summary, an ISP may assign an IP address to different Subscribers at different periods of time. But as the time of the data transfer is recorded accurately by the System, an ISP is able to determine which Subscriber it assigned an IP address to at a particular moment in time using the IP address and time stamp.

What are hash collisions and how do they affect the accuracy of the information recorded by the System?

11.34 A hash collision occurs when two different data sets have the same hash value. The System uses the SHA-512 method to determine the hash value of the sub-piece transferred.

11.35 There are two potential sources of hash collisions:

(a) Trivial hash values, being:

(i) hash values calculated from a file containing either bit values of 0 or 1 or the same data in each block; and

(ii) bit patterns that describe the type of the file transferred.

The System checks the hash values against a blacklist and other trivial data sets. Data transfers involving these hash values are deleted by the System and are not logged.

(b) Hash collisions for the complete file and a sub-piece

The System identifies files by their name and SHA-1 hash value. It checks the value of the SHA-512 hash value of the downloaded sub-pieces against the calculated SHA-512 hash value of the reference file. To perform this calculation the position of the sub-piece must also be known. The probability of the same data of two different files being at the same place is extremely unlikely. This means that where the hash value and location match, it may be said that distribution of the relevant data set took place.

11.36 Thus, in my opinion, the System addresses the risk of trivial hash collisions and treats them correctly by not recording the transaction.

Does the System download from a source computer the full file?

11.37 The connection with a source is terminated after the System receives a sub-piece of the copyrighted work from the source. The System may connect with the same source again in relation to another sub-piece of the same work. However, the System does not acquire the whole file from a single source. This is not necessary or always possible.

11.38 However, each user does make the Complete Data Set available for download to other users. While it is possible for a user to make only one piece of the Complete Data Set available, this is extremely unlikely, as the user would need to possess highly specialised information technology skills to split a Complete Data Set into several pieces and make each piece available for download separately. This is also highly unlikely because users on the BitTorrent network are using BitTorrent network to download and make available Complete Data Sets, such as complete movies.

11.39 Users of BitTorrent networks seek to acquire various files. These may include copyrighted works such as movies, software and music. In relation to movies, the intention of a user is to acquire a full copy of a movie. It would be nonsensical to acquire only one piece, which may only be a few random seconds of the movie. When a user acquires software, it is unlikely that the software will function properly or even install if it is incomplete. A user therefore needs to acquire the Complete Data Set for the software in order to use it. Furthermore, in relation to music files, it is inconceivable that a user will only seek to acquire a random few seconds of a musical work.

11.40 BitTorrent users have little control over which pieces they acquire first. They acquire the pieces that are available and offered by other users. For these reasons, it is reasonable to say that

those wishing to acquire a file will participate in a BitTorrent network until the Complete Data Set is acquired.

11.41 A user who has acquired the Complete Data Set will continue to transmit pieces of that file to other users unless the client is instructed to stop doing so. It is not uncommon for users to continue distributing the files for long periods of time.

11.42 A user without the full data set will request the pieces it does not have from other sources. These transmissions may take place while that user is transmitting data to the System. Accordingly, the possession and distribution of a sub-piece is a very strong indication that the user is or will in the future be in the possession of the complete data set.

11.43 The reason why P2P Networks work so quickly and effectively is because different parts of a movie are obtained from different users until an entire film is downloaded. That is, when downloading a movie the user's computer may simultaneously download the movie from a number of different other peers' computers.

12. Declaration

I, Dr Simone Richter, declare that I have made all the inquiries that I believe are desirable and appropriate and that no matters of significance that I regard as relevant have, to my knowledge, been withheld from the Court.

Signature..... Date.....2. April 2014

Von der Industrie- und Handelskammer öffentlich bestellte und vereidigte Sachverständige
Dipl.-Phys. Dr.
Simone Georgia Richter
Sachverständige für
Systeme und Anwendungen
der Informationsverarbeitung

Dr. Simone Richter
permanently sworn in and officially installed expert witness for Systems and Application of Information Technologies.
Officially supervised by the Chamber of Commerce in Darmstadt

Appendix A

Curriculum Vitae

Name: Dr. rer. nat. Simone Richter

Born: December 13, 1969 in Berlin, Germany

Career Summary

Over the past 15 years of professional experience, I have gained extensive knowledge and expertise in the following areas:

- General and extended IT-knowledge
- Company Administration (Finance, Controlling, Human Resources, Technology Transfer, Legal, Site & Buildings, Risk Management, Quality Management, Procurement and Contracts)
- General Project Management
- Strategy Development and Deployment
- Change Management
- Technical Project Management
- Project Team Management
- ERP Project Management
- Project Evaluation
- Particle Accelerator Design and Operations

Due to my training and experience, I have specialised knowledge in all areas of I.T including computer hardware, software, data centres, networking and database management.

Since January 2003, I have provided independent expert witness testimony in Germany as an accredited I.T. expert. This accreditation is provided by the German government to individuals who demonstrate through rigorous testing, that they have the qualifications, training and experience to provide expert opinions and Court approved reports for the benefit of the Court in the field of IT.

Education / Training

2010 – 2011	Malik Master of Management, MZSG St. Gallen, Switzerland
2008 – 2010	Certificate, Helmholtz-Akademie für Führungskräfte – Helmholtz Management Academy, Germany

2002 – 2005	Certificate, Medical Physics and Technics, University of Kaiserslautern, Germany
2001	Dr. rer. nat. - PhD, Technical University Darmstadt, Germany. My phD
1996	Diploma –Master with Thesis in Physics, Technical University Darmstadt, Germany. My diploma involved writing approx. 8,000 lines of code
1988	Abitur – university-entrance diploma , Göttenbach-Gymnasium, Idar-Oberstein

Certifications and Accreditations

2009 –	Member of the Lenkungsausschuss Helmholtz-Akademie für Führungskräfte(Steering Committee of the Member of the Selection Committee for the Helmholtz-Gemeinschaft Deutscher Forschungszentren e.V. (Helmholtz Association - HGF) Mentoring Program
2009 – 2010	Elected Member of the GSI Wissenschaftlicher Ausschuss (internal scientific council)
2008 – 2010	Elected Deputy Chairwoman of the Wissenschaftlich-Technische Räte-Versammlung of Helmholtz-Gemeinschaft Deutscher Forschungszentren e.V. (HGF) (Committee of scientific-technical councils)
2006	Member of the Working Group "Strategy of the Helmholtz Gemeinschaft"
2005 – 2008	Chairwoman and Elected Member of the GSI Wissenschaftlicher Ausschuss
2004 – 2005	Member of the Internal Advisory Committee for the Definition of the Structure of the FAIR Project
2003 – 2004	Appointed Member of GSI Wissenschaftlicher Ausschuss
2003	Accreditation as expert witness (permanently sworn in as public officer) with formal knowledge examination according to German law §36 GewOo re-certified and re-newed oath
2005, 2010	My involvement at the GSI's internal scientific council and later on also within the committee of scientific-technical councils at the HGF gave me the opportunity to work on the steering level of the organizations. The boards I led as chair or deputy chair are asked to give strategic advice to the managing and governing boards of the respective organizations. I gained a deep insight in decision making pro-

cesses with high profile stakeholders such as ministries or other representatives of large research organizations.

Work History

2003 – Present

In January 2003 I was formally sworn in as a public officer as an expert witness in all matters related to Information Technology. In this position, I am regularly called to provide technical expertise for Courts, public prosecuting offices and private organizations as an independent expert for the Court.

Prior to my accreditation, I was required to demonstrate my IT knowledge, by working on 10 different cases with another IT expert. After this, I was required to undertake a two day test which involved a written component in respect of General IT knowledge including hardware, software and operating systems, and then a practical component in which I was interviewed and my technical knowledge was tested by a 3 person panel. Upon passing these tests, I was recommended by the Darmstadt Chamber of Commerce to be accredited as an independent expert. Constant training is required to be reaccredited every 5 years. The chamber of commerce in Darmstadt, Germany serves as my supervising agency on behalf of the German government. Currently only approx. 140 individuals in Germany hold this accreditation.

Key Achievements

- Giving technical expertise reports and testimonies in civil court cases
- Providing IT-related expert opinions and court approved reports in private cases

General Responsibilities

- Working with judges, lawyers, plaintiff and defendant representatives to solve court cases
- Providing Companies with technical background information about IT-project statuses
- Evaluation of IT-project statuses
- Acting as public office for (IT) technical questions in and before court cases

2011 – Present **GSI Helmholtzzentrum für Schwerionenforschung GmbH; www.gsi.de**

Interims Manager

In-house Consultant and Strategic Advisor to the Administrative Managing Director

GSI GmbH as member of the German Helmholtz-Association (www.helmholtz.de) provides a heavy-ion particle accelerator as a national large-scale research infrastructure open to national and international scientists. The Helmholtz-Association is Germany's largest research organization with more

than 33.000 employees and ad yearly budget exceeding 3.4 billion €. GSI GmbH with a work force of approx. 1100 employees and a yearly budget of over150 M€/year is structured in two business units: the research and scientific and the administrative branch. I advise the Administrative Managing Director and interim manage the Organization and Controlling Division.

Key Achievements

- structuring and shaping of the Organization/Controlling Division and the Finance/Procurement Division initiating and guiding the change process for the Administration to a customer and project oriented division
- member of the task-force for Merger of GSI GmbH and FAIR GmbH
- member of the working group on the supervision of the FAIR GmbH
- interims management for the Organization/Controlling Division, coaching of the Head of the Finance and Purchasing Division, coaching of the department leader of the Patent and Technology transfer department, focusing on developing the IP-Policy
- requirements analysis for Controlling and Business Intelligence System.

General Responsibilities

- ensuring the process set-up for company-wide processes like budgeting, quality management, project management
- giving expertise on all FAIR related issues like in-kind contributions, project steering and control, reporting, and merging of FAIR with GSI
- giving expert advice to the CFO and the CEO of GSI

2010 – 2011 Facility for Antiproton and Ion Research in Europe GmbH (FAIR GmbH), www.fair-center.org Administrative Managing Director, CFO.

- The FAIR project represented by the FAIR GmbH is currently the largest European Research Project under construction. Its estimated cost to completion is 1.6 billion€, the envisaged start of operation is 2018.
- FAIR GmbH was founded under the umbrella of multinational treaty – The FAIR Convention. So far there are eight signatory countries to the Convention (Finland, France, Germany, India, Poland, Russia, Slovenia, and Sweden) and Spain expected to sign soon. The signatory countries are represented in the shareholders assembly through their respective nominated shareholders. For Germany the GSI GmbH is the major shareholder of FAIR. Shareholders contribute to the facility either in cash or in-kind. The FAIR accelerator facility complex will provide to more than 3000 scientists high precision heavy ion and antiproton beams at various experimental stations.

Key Achievements

- bringing the company into operation

- shaping the companies organization, structure and major processes
- shaping and setting up the administration business unit including the site & buildings division
- major contracts – site & buildings, business management contract with GSI GmbH, in kind contracts

General responsibilities

- full personal liability for the company's assets
- reporting to the shareholders assembly
- liaisons to shareholders, ministries and other stakeholders of FAIR
- liaisons to local authorities and the public bringing up a fully-fledged research facility – construction budget 1.6 billion€, expected yearly budget for operations 150 million€ joint management with the scientific managing director of the company yearly cash budget ramping up from 2 million€ (shortened financial year Oct.2010 – Dec. 2012) up to 44 million€ in 2012
- Personnel ramping up from 2 employees in Oct 2010 up to 35 in 2011
- set-up of the company
- Finances, accounting, purchasing, legal services
- all issues related to site & buildings, building permits, etc.
- general services
- personnel and recruitment
- technical safety and security

2001 – 2010 **GSI Helmholtzzentrum für Schwerionenforschung GmbH; www.gsi.de**

2009 – 2010 **Designated Administrative Director for FAIR, FAIR Division, Prokura (proxy holder on behalf of GSI GmbH)**

Within the GSI GmbH as designated major shareholder, host lab and providing the injector accelerators to the FAIR facility, the FAIR division has been established providing the seed for the company to be founded. I took over the responsibility to coordinate and account for all necessary work and preparations to start-up the company FAIR GmbH after the signature of the Convention for FAIR.

Key Achievements

- providing advice to the international community (in this respect the representatives to the prospective signatory countries – mostly research ministries or research organizations) to formulate the legal documents
- formulating the contractual framework for in-kind contributions
- preparing the SAP-ERP system of GSI to accommodate FAIR's accounting area independently
- pre-defining structures and processes for the FAIR GmbH

General Responsibilities

- proxy holder for GSI GmbH for all FAIR issues
- recruiting in advance for the FAIR GmbH

2006 – 2009 Deputy Division leader FAIR Technical Division, Deputy Technical Project Leader Accelerator and Site & Buildings, Principal Investigator

Focusing on the technical challenges of the FAIR project (accelerator construction and issues concerning site & buildings), I was Deputy for the assigned Technical Director for FAIR. The GSI's FAIR Division was holding five departments and almost 100 employees. The main focus of the Division laid on elaborating all specifications for the accelerator complex and the construction and development of site & buildings.

Key achievements

- implementing project structures for the accelerator sub-project
- recruiting key personnel for the newly formed site & buildings department
- implementing the central documentation and document management system for FAIR
- structuring of the project and the project documentation

General Responsibilities

- assuring the completeness and integrity of the specifications
- managing the accelerator subproject within the EU-FP7 preparatory
- phase program for FAIR principal investigator within the Helmholtz
- Program "Structure of Matter" managing the division on behalf of the division leader
- budgeting for the division

2001 – 2006 Researcher Accelerator Division, Injectors Department

GSI runs a large scale accelerator on a 7/24 basis for approx. 6000 hours/year. Within the linear accelerator and injector group I was the responsible liaison person to the controls department

Key achievements

- redesign and supervision of the implementation of the experimental beam line for the Super Heavy Elements (SHE) Program@GSI (discovery of new elements) and bringing it into successful operation
- requirements specification of the FAIR control system and machine protection systems

General responsibilities

- responsible machine physicist on duty for the linear accelerator
- accelerator coordinator on duty

- beam-line responsible for the SHE-beamline and general beam diagnostics
- requirements analysis and specifications for operating software and beam diagnostics elements
  Work package leader EU FP6 Program GANMVL
- supervisor for summer students

1996 – 2001 Researcher, Institute for Nuclear Physics, Technical University Darmstadt

I worked on my PhD thesis on the "*Design and Implementation of a Local Control System for the S-DALINAC*". This involved the designing of hardware boards, and involved me writing approximately 15.000 lines of code. The work comprised the completely new set-up of the local control system, but still to work together with the remote controls. I was also involved in the supervision of students at the University.

Key achievements

- implementing a new local controls system for the accelerator
- implementing a new rf-controls system (supervision of a diploma thesis)
- implementing a controls system for the Helium refrigerator (supervision of a diploma thesis)
- implementing an archiving database for accelerator settings

General responsibilities

- ensuring the operationally of the accelerator
- deputy IT-Manager of the institute
- supervision of diploma/master thesis
- co-supervision of PhD thesis
- supervision of the practical training of graduate students (lab work)