RETENTION OF COMMUNICATIONS DATA
UNDER
Part 11: ANTI-TERRORISM, CRIME & SECURITY ACT 2001

VOLUNTARY CODE OF PRACTICE

FOREWORD

The Anti-Terrorism, Crime & Security Act was passed in December of 2001 (the Act) Part 11 of the Act aims to allow for the retention of communications data to ensure that the UK security, intelligence and law enforcement agencies have sufficient information available to them to assist them in protecting the UK's national security and to investigate terrorism.

Communications data are retained by the communications service providers to enable them to carry out their business effectively. Such information could be divided into three broad categories these being subscriber information (identifies user); traffic data (identifies whom was called etc); and use made of service (identifies what services are used). The Act recognises that communications data are an essential tool for the security, intelligence and law enforcement agencies in carrying out their work to safeguard United Kingdom national security. These agencies, which are authorised to acquire communications data under statutory provisions, would be greatly assisted if they could rely on the communications data being available when they required it.

Part 11 of the Act provides only for the retention of data that communication service providers already retain for business purposes. Its object is not to enlarge the fields of data which a communication service provider may (or must) retain, but to encourage communication service providers to retain that data for longer than they would otherwise need to do so for their own commercial purposes. The Act identifies that the purpose of the retention period is the safeguarding of national security or for the prevention or detection of crime or the prosecution of offences which relate directly or indirectly to national security.

This Code of Practice relates specifically to the need for communications service providers to retain data for extended periods of time in order to assist the security, intelligence and law enforcement agencies in carrying out their work of safeguarding national security or in the prevention or detection of crime or the prosecution of offences which relate directly or indirectly to national security.

This Code of Practice does not address issues relating to disclosure of data, it simply addresses the issues of what types of data can be retained and for how long it will be retained beyond a particular company's existing business practice. The Code explains why communications service providers have the ability to retain data beyond their normal business purposes for the reasons outlined in the Act.

Communications data may be obtained by security, intelligence and law enforcement agencies under the Regulation of Investigatory Powers Act 2000 and other statutory powers. This Code does not deal with these provisions.

The Data Protection Act 1998 requires that personal data are processed lawfully. In retaining communications data for longer than needed for their own business purposes and for the purposes identified in the Act communication service providers will process personal data. The Information Commissioner's Office (ICO) has accepted that such processing will not, on human rights grounds, contravene this requirement of the Act.

However, individual communication service providers must satisfy themselves that the processing is "necessary" for one of a range of functions. In doing so they are entitled to rely heavily on the Secretary of State's assurance that the retention of communications data for the periods as specified in this Code is necessary for the government's function of safeguarding national security, and on the fact that the Code has been approved by Parliament.

The ICO has though expressed concern about such retained data being acquired for purposes that do not relate to national security. Acquisition of communications data is not addressed in the Act and therefore is not within the proper ambit of this Code.

CONTENTS

Purpose of the Code

1. In section 102 of the Act, Parliament has given the Secretary of State the power to issue a Code of Practice relating to the retention of communications data by communication service providers. This Code of Practice is intended to outline how communication service providers can assist in the fight against terrorism by meeting agreed time periods for retention of communications data that may be extended beyond those periods for which their individual company currently retains data for business purposes.

2. After consultation with the security, intelligence and law enforcement agencies, the Secretary of State has determined that retention of communications data by communication service providers in line with the Appendix to this Code of Practice is necessary for the purposes set out in section 102(3) of the Act, namely;
   (a) the purposes of safeguarding national security
   (b) the purposes of prevention or detection of crime or the prosecution of offenders which may relate directly or indirectly to national security.

3. The Code of Practice is intended to ensure that communication service providers may retain data for the two purposes identified at 2 a & b, after the need for retention for business purposes has elapsed and there is otherwise an obligation to erase or anonymise retained data. It does not provide guidance on the manner in which data retained for these purposes should be processed; nor does the Secretary of State consider it necessary to impose new standards on the conditions in which the data are stored, e.g. technical media, security, ease of access, indexing or other.

4. The Code does not relate to the powers of public authorities to obtain communications data retained in accordance with the Appendix to the Code. Acquisition of communications data is provided for by Chapter II of Part I of the Regulation of Investigatory Powers Act 2000, as well as other relevant statutory powers. See paragraphs 25 to 28.

Human rights and data protection considerations

5. This Code has been drawn up in accordance with existing legislation, including the Human Rights Act 1998, and the Data Protection Act 1998, and the Telecommunications (Data Protection and Privacy) Regulations 1999, together with their parent directives.

6. Data retained under the Code are subject to the data protection principles found in the Data Protection Act 1998. Under the first data protection principle personal data may only be processed if at least one of the conditions in Schedule 2 to the 1998 Act is met. The processing of data retained under this Code falls within paragraph 5 of Schedule 2 of the Data Protection Act 1998 in that it is necessary for the communication service provider to retain data to enable the Secretary of State to fulfil his function for the protection of national security. Some communications data may in certain circumstances constitute sensitive personal data. Processing of such data is permitted by virtue of Schedule 3, paragraph 7 of the 1998 Act.

7. Data retained under the Code will, at least for a certain period, be data that are needed by the communication service provider for business purposes. Its processing will therefore initially be undertaken for a dual purpose: (a) business purposes, (b) national security purposes, where “national security purposes” includes both the purposes set out in section 102(3) of the Act. Since both purposes of retention will apply to all data simultaneously during the ‘business purpose time period, there is no need for separate storage systems for “business data” and “national security data” under this dual-purpose scheme.

   However, once an individual company has exceeded the business purpose time period then data will be retained specifically for the purposes described in Section 102(3) of the Act. The system deployed by individual companies will need to identify that the data has exceeded the business purpose time period. Individual communication service providers will need to ensure that they do not access those data for their own purposes. At the end of the retention period necessary for ‘business purposes’ the only data that a communication service provider should retain are that data identified in the ‘Technical Specification’ attached as Appendix A to this Code.

8. The fifth data protection principle provides that personal data processed for any purpose or purposes shall not be kept for any longer than is necessary for that purpose or those purposes. The periods for which it appears necessary to the Secretary of State for communication service providers to retain communications data for national security purposes are those set out in Appendix A. The periods for which it is necessary for communication service providers to retain communications data for business purposes is a matter for each communication service provider, and they might be longer or shorter than the retention periods the Secretary of State has set out are necessary for national security. Compliance with the fifth data protection principle under the dual-purpose scheme requires that after the expiry of the shorter of these two periods, communications data may only be retained further for the period required by the remaining purpose. When the retention periods for both purposes have expired, the data must be either anonymised or erased.

9. As indicated the Secretary of State considers the retention of data in accordance with Appendix A to be necessary for the purpose of national security and accordingly retention for those periods should comply with the fifth data protection principle. However, because the purpose of retention is to safeguard national security were it to be suggested that retention in accordance with this Code did not comply with the fifth principle, the national security exemption in s 28 of the Data Protection Act 1998 could be relied on to exempt such data from the fifth principle so enabling it to be retained in accordance with the Code. If necessary the Secretary of State would issue a certificate under s 28.2 confirming the same.

10. The data subject access provisions set out in the Data Protection Act 1998 continue to apply to communications data retained under this Code, that is to say that data subjects may request access to their personal data whether it is held for national security purposes or for the communication service provider's business purposes. In addition, subscribers should be notified where their personal data will be retained for the purpose of the Act, as well as for the communication service providers business purposes, and that it may be disclosed to relevant public authorities, as set out in paragraph 27 of this Code. Every effort should be made to ensure that this is brought to the attention of the subscriber for example this could be added to billing information or sent by way of text message or e-mail.

NB. Communication service providers will need to ensure that their entry in the register of data controllers maintained by the Information Commissioner describes the processing of personal data involved in retention of communications data for the national security purposes. The Information Commissioner's advice is that they should notify that they are processing for the following purpose
"*NATIONAL SECURITY:~ Retention of communications data for the purpose of safeguarding national security or for the purposes of prevention or detection of crime or the prosecution of offenders which may relate directly or indirectly to national security*"

This is not one of the standard purpose descriptions that the Information Commissioner provides so communication service providers will need to complete it in full, together with details of the associated data subjects, classes and recipients, when they apply to add a new purpose to their existing notification.

11. The retention specification set out in Appendix A to this Code has been drafted taking into account a number of factors, including the right to respect for private life under Article 8 of the European Convention of Human Rights. The Secretary of State considers the retention periods set out in Appendix A to be both necessary and proportionate in light of the individual's right to respect for private life and the national security purposes for which the retention of data is required.

Jurisdiction and types of operators covered by the Code of Practice

12. The Code of Practice applies to all communication service providers who, provide a public telecommunications service in the United Kingdom as defined in section 2 of the Regulation of Investigatory Powers Act 2000, and who retain communications data in line with the provisions of the Act. The Secretary of State considers it necessary for the national security purposes outlined in the Act, for communications data held by communication service providers, which relates to subscribers resident in the UK or subscribing to or using a UK-based service, to be retained in accordance with the provisions of the Code, whether the data are generated or processed in the UK or abroad. However, if data relating to a service provided in the UK are stored in a foreign jurisdiction it may be subject to conflicting legal requirements prohibiting the retention of data in accordance with this Code. In such cases, it is accepted that it may not be possible to adhere to the terms of this Code in respect of that communications data.

13. The data categories and retention periods in the Appendix to this Code have been determined with regard to considerations of necessity and proportionality. The data categories and retention periods relate to communications data generated and retained by communication service providers who provide a service to the general public in the United Kingdom. This Code is not intended to apply to individuals or organisations who do not provide such a public service (e.g. private networks).

14. In some cases, two or more legal entities may be involved in the provision of a public telecommunications service, e.g. backbone/virtual service provider model. In such cases, the provisions of this Code apply to data retained by each legal entity for their own business purposes.

Types of data and retention periods

15. Communications data can be divided into three broad categories, corresponding to the definitions in section 21(4) of the Regulation of Investigatory Powers Act 2000, which can be summarised as follows:

a) **traffic data** – including telephone numbers called, email addresses, and location data etc.
b) **use made of service** – including services subscribed to, etc.
c) **other information relating to the subscriber** – including installation address, etc.

"communications data" as defined by RIPA means any of the following-
(i) any traffic data comprised in or attached to a communication (whether by the sender or otherwise) for the purposes of any postal service or telecommunication system by means of which it is being or may be transmitted;
(ii) any information which includes none of the contents of a communication [apart from any information falling within paragraph (i)] and is about the use made by any person-

(1) of any telecommunications service; or
(2) in connection with the provision to or use by any person of any telecommunications service, of any part of a telecommunication system;

(iii) any information not falling within paragraph (i) or (ii) that is held or obtained, in relation to persons to whom he provides the service, by a person providing a telecommunications service.

"traffic data", as defined by the Regulation of Investigatory Powers Act 2000 in relation to any communication, means-
(i) any data identifying, or purporting to identify, any person, apparatus or location to or from which the communication is or may be transmitted,
(ii) any data identifying or selecting, or purporting to identify or select, apparatus through which, or by means of which, the communication is or may be transmitted,
(iii) any data comprising signals for the actuation of apparatus used for the purposes of a telecommunication system for effecting (in whole or in part) the transmission of any communication, and
(iv) any data identifying the data or other data as data comprised in or attached to a particular communication, but that expression includes data identifying a computer file or computer program access to which is obtained, or which is run, by means of the communication to the extent only that the file or program is identified by reference to the apparatus in which it is stored.

References, in relation to traffic data comprising signals for the actuation of apparatus, to a telecommunication system by means of which a communication is being or may be transmitted include references to any telecommunication system in which that apparatus is comprised; and references to traffic data being attached to a communication include references to the data and the communication being logically associated with each other.

16. The maximum retention period for data held under the provisions of this Code is 12 months, without prejudice to any longer retention period which may be justified by the business practices of the communication service provider.

17. For data categories 15(a) and 15 (b) above the period of retention begins at the point when the call ends, for subscriber-related data category 15 (c) the period of retention begins when the data are changed or subscriber leaves the service.

18. The retention periods given in Appendix A recognise that types of communications data, as personal data, vary with respect both to their usefulness to the agencies, and to their sensitivity. It is recognised that the usefulness of different types of communications data for the purpose of safeguarding national security will vary and this is reflected in the different retention periods.

19. The data categories listed in Appendix A will not all be relevant to every communication service provider. Whether or not a data type will be relevant to a communication service provider and therefore retained will depend on the services which it provides, for example, an internet service provider will not retain IMEI data. Communication service providers will not be expected to retain additional categories of data to those which they routinely retain for business purposes. In other words if a data type is not already captured for the business purposes of an individual company then there will be no expectation that this data type is retained for the purposes of the Act.

Agreements

20. The Secretary of State may enter into agreements with individual communication service providers who receive requests for communications data stored under these provisions. The purpose of these agreements is to communicate the retention practices of those communication service provider to public authorities listed in Chapter II of Part I of the Regulation of Investigatory Powers Act 2000. They will play the role of Service Level Agreements (SLAs) and will include any arrangements for payments to cover retention costs. These SLAs will be based on an open document outlining the agreement between the Secretary of State and the company concerned. Each of these will differ with respect to the appendices which will outline the services that a particular provider is able to deliver. Those parts of these agreements that do not contain commercially sensitive material will be publicly available. The appendices will remain commercially sensitive.

21. The agreements will be drafted within the framework provided by this Code. An agreement may not set a retention period for any type of data which is greater than the period set out in Appendix A to this Code.

22. Any agreement will be made between the Secretary of State and the communication service provider and must be entered into voluntarily by both sides. It may be terminated by either side subject to a period of notice set out in the agreement.

Costs arrangements

23. Where the period of retention of data for national security purposes is not substantially larger than the period of retention for business purposes, the retention costs will continue to be borne by the communication service provider.

24. Where data retention periods are significantly longer for national security purposes than for business purposes, the Secretary of State will contribute a reasonable proportion of the marginal cost as appropriate. Marginal costs may include, for example, the design and production of additional storage and searching facilities. This may be in the form of capital investment into retention and retrieval equipment or may include running costs.

Acquisition of data retained under the terms of this Code of Practice

25. It is outside of the scope of this Code of Practice to address the issue of acquisition of data after it has been retained. It can only address the issue of retention of data for the purposes of the Act. The Act establishes the framework for communication service providers to retain data for the purposes of safeguarding national security and for the prevention or detection of crime and prosecution of offenders which may relate directly or indirectly to national security.

26. The Code sets out a retention specification which is designed to meet the two aims set out above, both relating to national security. That is to say that any particular piece of data is retained because it belongs to a certain data type, and it is necessary to retain all data of that type for the purpose of safeguarding national security or for the purpose of the prevention or detection of crime or the prosecution of offenders which may relate directly or indirectly to national security.

27. The retention of such data is necessary so that it is available to be acquired by relevant public authorities under Chapter II of Part I of the Regulation of Investigatory Powers Act 2000, or otherwise, to assist them in safeguarding national security. However, whilst restrictions exist elsewhere, this Code cannot itself place restrictions on the ability of these bodies or other persons to acquire data retained under the Code for other purposes through the exercise of any statutory power. In particular, this Code cannot place any restrictions on the ability of the public authorities listed in Chapter II of Part I of the Regulation of Investigatory Powers Act 2000 to acquire data retained under this Code for any of the purposes set out in section 22 of that Act which do not relate to national security.

28. In addition data access requests can also be received from data subjects under the Data Protection Act 1998 and from civil litigants.

Oversight mechanism

29. The retention of communications data is a form of personal data processing. As such, it is subject to the Data Protection Act 1998. Oversight of the 1998 Act is by the Information Commissioner.

Transitional arrangements

30. All data collected after the communication service provider adopts the Code should be processed in accordance with both the national security purposes and the business purposes from the point that it is generated. Data already held by the communication service providers at the time of adopting the Code will be processed only in accordance with the purpose for which it was originally collected.

31. Subscribers should be notified of the new purpose for which data is being retained. This may be done by sending out a general notification to all customers. The national security purpose must be made clear to any new subscribers at the time they subscribe.

32. During the period of time that a communications service provider is building the technical capacity to extend retention of specified data beyond their normal business time periods, the company's standard retention practice takes precedence. Once the individual communication service provider has the technical capacity to retain data for the extended time periods set out in this voluntary Code of Practice, then the communication service provider shall inform existing and new customers that the purpose for retention and the periods of retention have been varied to meet with the needs of the Act. Only after this information has been passed on to existing customers and new customers can the communication service provider then retain the data for the extended time periods for the purposes of national security. There may be a period after the communication service provider has adopted the Code when he cannot retain data for the full period set out in Appendix A owing to the need to introduce technical adaptations. The agreement with the communication service provider will set out how long it will take to reach full compliance.

Criteria for assessing the effectiveness of the Code of Practice

33. The Code will be reviewed three months from the date when it first receives parliamentary approval, in accordance with the following criteria:

    (a) Has it improved investigative work?

    (b) How many request for data have been made?

    (c) Is the voluntary system working?

    (d) What percentage of the market is covered by communication service providers who have adopted the Code of Practice?

    (e) Are sectors of the industry which have not adopted the Code enjoying an unfair commercial advantage?

    The SLAs introduced under this Code will require communication service providers to keep records of all enquiries made for data retained under the Act from the date an individual service provider enters into a voluntary agreement with the Secretary of State, in order to enable a comprehensive survey to be undertaken.

APPENDIX A

Data retention: expansion of data categories

SUBSCRIBER INFORMATION **12 months**
(From end of subscription/last change)

Subscriber details relating to the person

e.g. Name, date of birth, installation and billing address, payment methods, account/credit card details

Contact information (information held about the subscriber but not verified by the CSP)

e.g. Telephone number, email address

Identity of services subscribed to (information determined by the communication service provider)

e.g. Customer reference/account number, list of services subscribed to

Telephony:	telephone number(s), IMEI, IMSI(s)
Email:	email address(es), IP at registration
Instant messaging:	Internet Message Handle, IP at registration
ISP - dial-in:	Log-in, CLI at registration (if kept)
ISP - always-on:	Unique identifiers, MAC address (if kept), ADSL end points, IP tunnel address

TELEPHONY DATA **12 months**

e.g. All numbers (or other identifiers e.g. name@bt) associated with call (e.g. physical/presentational/network assigned CLI, DNI, IMSI, IMEI, exchange/divert numbers)
Date and time of start of call
Duration of call/date and time of end of call
Type of call (if available)
Location data at start and/or end of call, in form of lat/long reference.
Cell site data from time cell ceases to be used.
IMSI/MSISDN/IMEI mappings.
For GPRS & 3G, date and time of connection, IMSI, IP address assigned.
Mobile data exchanged with foreign operators; IMSI & MSISDN, sets of GSM triples, sets of 3G quintuples, global titles of equipment communicating with or about the subscriber.

SMS, EMS and MMS DATA **6 months**

e.g. Calling number, IMEI
Called number, IMEI
Date and time of sending
Delivery receipt - if available
Location data when messages sent and received, in form of lat/long reference.

EMAIL DATA **6 months**

e.g. Log-on (authentication user name, date and time of log-in/log-off, IP address logged-in from)
Sent email (authentication user name, from/to/cc email addresses, date and time sent)
Received email (authentication user name, from/to email addresses, date and time received)

ISP DATA **6 months**

e.g. Log-on (authentication user name, date and time of log-in/log-off, IP address assigned)
Dial-up: CLI and number dialled
Always-on: ADSL end point/MAC address (If available)

WEB ACTIVITY LOGS **4 days**

e.g. Proxy server logs (date/time, IP address used, URL's visited, services)
The data types here will be restricted **solely to Communications Data and exclude content of communication**. This will mean that storage under this code can only take place to the level of www.homeoffice.gov.uk/…….

OTHER SERVICES **Retention relative to service provided**

e.g. Instant Message Type Services (log-on/off time) If available.

COLLATERAL DATA **Retention relative to data to which it is related**

e.g. Data needed to interpret other communications data.for example -the mapping between cellmast identifiers and their location –translation of dialling (as supported by IN networks)

Notes:

All times should include an indication of which time zone is being used (Universal Co-ordinated Time is preferred).

An indication should also be given of the accuracy of the timing.

To assist in the interpretation of Internet terminology the Home Office have, with the permission of the Internet Crime Forum, reproduced at Appendix C the document written by the Data Retention Project Group of the Internet Crime Forum.

The Home Office recognises the effort that has gone into producing this document and would thank all those responsible for its production.

APPENDIX B

Agreements

To be written as single document outlining voluntary agreement and requirements of Appendix A. To include separate appendices relative to individual company's additional storage.

APPENDIX C

Principal Current Data Types

Howard Lamb
Chair
ICF Data Retention Project Group

1 Introduction

1.1 In December 2001, the Internet Crime Forum (ICF) established a project group the primary aim of which was to identify current data types in use by subscribers who have access to the Internet.

1.2 The group was not tasked with debating the legal issues in relation to the data types identified. There are many legal issues relating to data retention and these will undoubtedly be discussed in other documents.

1.3 The group was established with a view to producing a document that would provide a better understanding of the technology used and the information that law enforcement is seeking from its investigations. It is not intended to be a standard or a best practice document. The document is intended to be a reference to what data may be available and which of those data types are likely to be useful to law enforcement when conducting an investigation.

2. Group Members

2.1 The group is restricted to technical and investigation experts, as explained in 1.2, this group does not hold a view on the value or legality of access to this data.

2.2 The ICF Data Retention Project Group called upon experts from the Internet industry who gave advice on the numerous data types that are created when a subscriber connects to and communicates via the Internet. This connection could be through an Internet Service Provider (ISP), a Virtual Internet Service Provider (VISP) or by other connection to the Internet.

2.3 The group also engaged the services of Computer Forensic experts whose work regularly involves liasing with various Law Enforcement Agencies and assisting with their investigation involving the Internet. Representatives of various Trade Associations were involved in the process together with several members of various Law Enforcement Agencies.

3. Acknowledgements

3.1 I would like to acknowledge the support given to this project by Chief Superintendent Len Hynds of the National High Tech Crime Unit, members of the Internet Crime Forum and to those experts from the Internet and Forensics Industry who have assisted in the process. All participants gave freely of their time as they agreed it was vital that this type of work be carried out.

4. Current Data Types

4.1 This document seeks to identify the principal known Data Types that a subscriber to an Internet Service may create whilst they are actively subscribing and utilising their Internet account.

4.2 It is accepted that this document could not be a definitive document of all data types due to the rapid development of technology.

5. Service Providers

It must be appreciated by the reader of this document that not all Internet Service Providers retain the data types that are mentioned within this document.

Each service provider is aware of their current data retention practices and may be able to advise on the detail. Communication should in the first instance be routed through a Law Enforcement Single Point of Contact for Law Enforcement personnel. Requests for data retention policies made from outside the SPOC regime may be liable to conditions determined by individual ISPs.

There are service providers, known as Virtual Internet Service Providers (VISPs), who utilise most, if not all, the infrastructure of a large service provider. They may utilise various elements of a service, such as mail, sign up servers, radius servers, web cache and news and badge them as their own. In these cases the data that a subscriber generates may be spread across several companies.

Even amongst traditional ISPs some parts of their service may be provided by third parties. In this case as well, information may be held by many different companies and may or may not be accessible to the primary ISP.

Furthermore, some data types for example, web server log information, may be owned by and under the control of the customer rather than the ISP.

6. Glossary

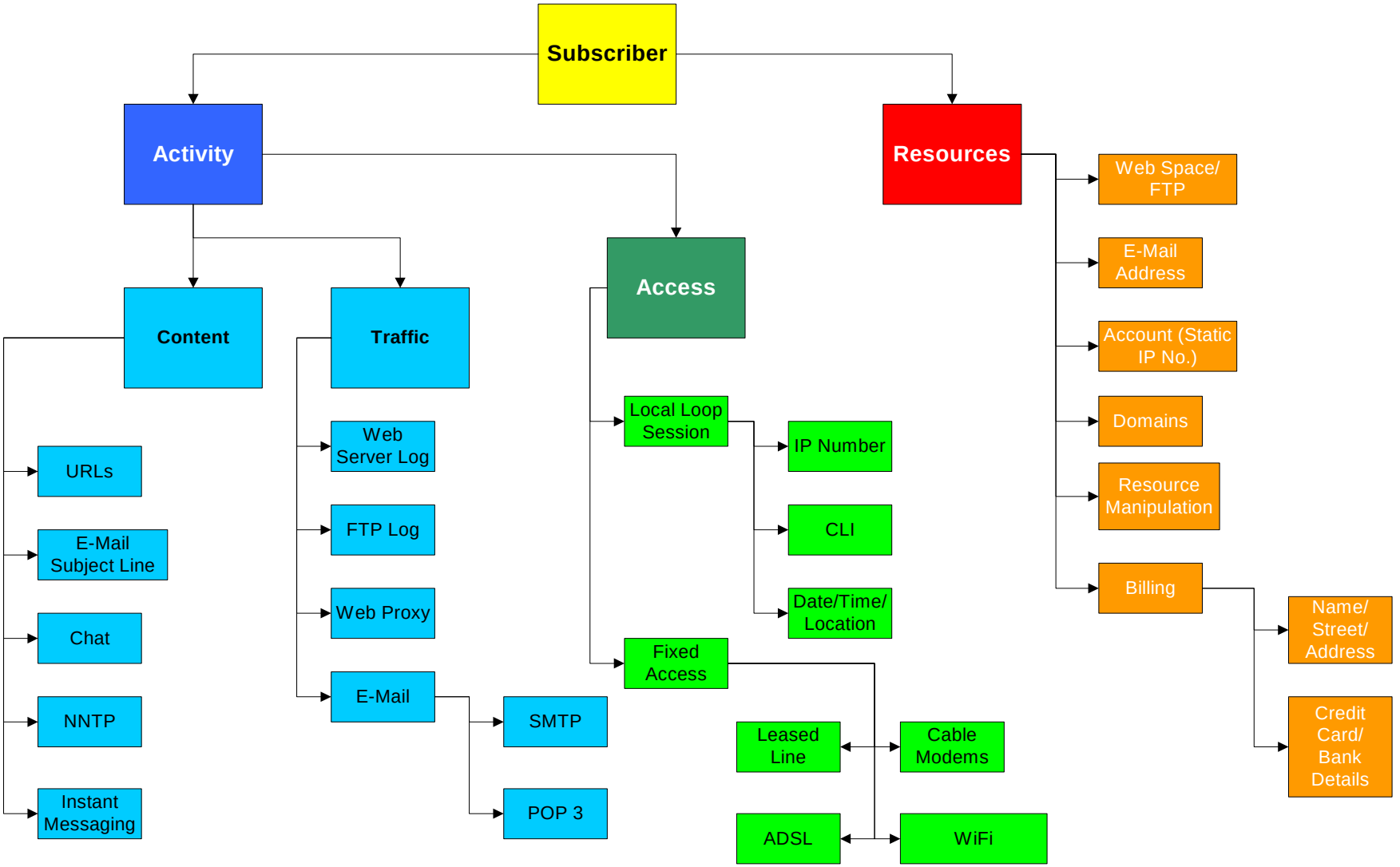
There is a glossary attached to this document that informs the reader of what the various data types are and it is advisable that this is read in conjunction with the rest of the document.

7 Subscriber Data Types

7.1 This next section of the document identifies the principal data types that may be created when a subscriber accesses the Internet.

7.2 The data types have been broken down into two main areas. The first being the activity of the subscriber and the second the resources that a subscriber could utilise.

7.3 When matching events on the Internet with details recorded in ISP logs it is absolutely essential to ensure that time and date information is correctly recorded. It is Best Practice for ISPs to synchronize their systems with global time standards using protocols such as NTP, however consideration should always be given to this not being the case for particular logs. Equally it is essential that enquiries about logging information provide accurate timing information. A frequently encountered pitfall is incorrect handling of timezone offset information and careful attention should be paid to this.


Subscriber
Activity
Resources
Content
Traffic
Access
URLs
E-Mail Subject Line
Chat
NNTP
Instant Messaging
Web Server Log
FTP Log
Web Proxy
E-Mail
SMTP
POP 3
Local Loop Session
IP Number
CLI
Date/Time/ Location
Fixed Access
Leased Line
Cable Modems
ADSL
WiFi
Web Space/ FTP
E-Mail Address
Account (Static IP No.)
Domains
Resource Manipulation
Billing
Name/ Street/ Address
Credit Card/ Bank Details

8. Potential Value

8.1 This section of the document identifies and details the potential value of the various data types to investigations. This is not a definitive list of data types and it must be appreciated that advances in technology may well mean that some of the data types that are currently of little value may at some stage in the future generate logs that could be useful for the purposes of investigations.

8.2 The table below identifies each of the data types and the data that could be generated by the subscriber.

8.3 Data can only be obtained in accordance with UK Law and international treaties. This document does not address this issue any further.

8.4 Internet Service Providers retain data for business purposes. The procedures surrounding this data retention may affect the way in which data could be used for evidential purposes.

Activity	Data Type	Comment
Content		
	URLs	A URL (Uniform/Unique Resource Locator) is the address of a file (resource) accessible on the Internet. The type of resource depends on the Internet application protocol. Using the World Wide Web's protocol, the Hypertext Transfer Protocol (HTTP), the resource can be an HTML page, an image file, a program such as a common gateway interface (CGI) application or Java applet, or any other file supported by HTTP. The URL contains the name of the protocol required to access the resource, a host name that identifies a specific computer on the Internet, and a hierarchical description of a file location on the computer. The host name can be used to determine the physical location of the computer and its logical ownership.
	E-mail	ISPs may hold e-mail on behalf of subscribers. Much of the e-mail is content and a number of different legal regimes apply to the divulgence of this. Some of the header is communications data. In addition details of what e-mail has been sent and received may be recorded in logs. Some of the information in these logs may be content.
	Chat	Depending upon the technology the service provider will not normally retain the content of an individual Chat Room session but individual participants will be able to make their own record. It may be possible to trace and identify participants in a chat session providing the IP address or for some ISPs the screen name is obtained together with an accurate time stamp.
	NNTP/ Usenet	In order to trace the author of a Usenet article, the article headers will need to be inspected. These will usually contain the posting IP address and time stamp. The system through which it was originally posted should then be able to identify the account responsible for creating the posting. Provision of Usenet services is increasingly performed by third parties so it may be necessary to make further enquiries with a connectivity ISP to determine where the account was used from. The content of an NNTP (Usenet) session will not be retained by a service provider. Therefore the readership of an article is unlikely to be available. Usenet postings are commonly exchanged between ISPs. This means that an article may well have been hosted on a different service provider from the one on which it is read.

	Instant Messaging	Instant messaging (sometimes called IM or Iming) is the ability to easily see whether a chosen friend or co-worker is connected to the Internet and, if they are, to exchange messages with them. Instant messaging differs from ordinary e-mail in the immediacy of the message exchange and also makes a continued exchange simpler than sending e-mail back and forth. Most exchanges are text-only. However, some services allow attachments. In order for IMing to work, both users (who must subscribe to the service) must be online at the same time, and the intended recipient must be willing to accept instant messages. (It is possible to set your software to reject messages.) An attempt to send an IM to someone who is not online, or who is not willing to accept Ims, will result in notification that the transmission cannot be completed. The ISPs will not in general have any records of the messages which have been exchanged because they flow directly between the participants (Peer to Peer). If a 'rendez-vous' server is involved in the initial connection between the participants then some logging information about their identities may be retained. The rendez-vous server may be totally independent of any connectivity ISP.
Traffic		
	Web Server Logs	These typically contain the source IP address, requested content, submitted data e.g. username, password and previous site visited. Some of the data may be content rather than traffic data. Some of the data may be anonymised in near real time. Some of the IP addresses may be proxy caches rather than the actual requestor.
	FTP Logs	These contain source IP, account details and details of the file names uploaded into or downloaded from. Although most sites appear to have a username/password login, anonymous guest accounts are also common and although an e-mail address is traditionally provided as identification there is seldom any validation of this whatsoever. Some of the data may be content rather than traffic data. It is quite common for customers to upload the content of their web pages using FTP.

	Web proxy	A proxy server is a server that acts as an intermediary between a user and the Internet. A proxy server receives a request for an Internet service (such as a Web page request) from a user. If it passes filtering requirements, the proxy server will access the remote site and pass the information to the user. A Web cache maintains a store of previously downloaded items from the Web such as an HTML page. If it is asked for a page that is already in its store, then it returns it to the user without needing to forward the request to the Internet, though it may need to check if its cached copy remains up-to-date. If the page is not in the cache then the cache server acting as a client, on behalf of the user, uses one of its own IP addresses to request the page from the server out on the Internet. When the page is returned, the proxy server relates it to the original request and forwards it on to the user. The user's general impression of using both proxy servers and caches will be of a direct connection to the remote site. In the ISP context it is usual to combine these two functions, and the result may be, rather confusingly, called a web cache, a web proxy or indeed a proxy-cache. Some ISPs use a "transparent" scheme that intercepts, for example, all HTTP (port 80) traffic and sends it via a proxy-cache. In other cases the use of a proxy-cache is entirely under the users' control, though the ISP may encourage usage by means of the default configurations shipped to its customers. These servers can produce logs of the data handled, giving the local customer IP address, details of the requested content and details of any connections made to remote sites. Complete logging may only be enabled for troubleshooting, but even incomplete logging can create very substantial volumes of data and these logs are not kept for long periods of time. The presence of a proxy may mean that the user never accesses the target web site. The access that is made will show the IP address of the server. The server may be configured to pass information to the target web site giving some details of the user, but some servers are configured specifically to obscure the true identity of the user. Some web pages are designed so that they cannot be cached and so this traffic will flow directly and hence proxy-cache logs will be incomplete. Similar effects will be caused by the use of protocols such as HTTPS which often avoid the use of a proxy-caches altogether.

E-mail		SMTP (Simple Mail Transfer Protocol) is a TCP protocol used in sending and receiving e-mail between permanently connected machines. However, because mail is "pushed" rather than "pulled" it works very poorly with intermittently connected machines. Therefore is usual to provide mail delivery to dialup customers via POP3 or IMAP. These provide a store and forward system, so that users can periodically "pull" any new e-mail. SMTP is the standard method for ISP customers to send their e-mail. Although delivery can be made directly to remote systems it is common to relay e-mail traffic via an SMTP server at the ISP called a "smart host". Some ISPs will intercept outgoing SMTP (port 25) traffic and force it to use the smart host. POP3 is a relatively simple protocol for e-mail reception. It is usual to configure clients to delete the e-mail once it has been fetched. If it is not deleted then the ISP will delete it after a preset period. Long term storage is done on the client machine. IMAP is somewhat more complex and provides a client/server implementation of a fully featured e-mail interface – with all the e-mail held on the server machine, possibly for very long periods. IMAP clients will hold very little state from one session to another.
	SMTP	Mail will be held on the server until it can be passed to a destination but most service providers do not routinely keep content thereafter. A service provider may retain summary logging details of e-mail that has been received from or sent to their customers. This would include a unique message identifier, who the mail was alleged to be from, who the mail was addressed to, the IP address of the immediately previous hop and the time and date the mail was sent. Further information such as size and content such as the subject line may sometimes be recorded. When the intended destination of e-mail is unavailable it may be routed via intermediate machines (using lower priority MX records). This will reduce the usefulness of IP address logging details. It is also essential to view the "from" details with caution since they are trivial to forge. Finally, many ISPs have outsourced virus scanning and "spam" deletion services. Initial delivery is made to a third party who will only forward genuine e-mail to the ISP's systems. In all cases, the e-mail itself should contain full details of all the machines it has passed through, but these machines will almost invariably only record one part of its journey.

	IMAP & POP3	IMAP & POP3 logs typically contain just brief summary details of connections. These may extend as far as recording the connecting IP address and how many e-mails were read or deleted. It would be very unusual indeed to record anything which references the content, sender or path associated with transmission of the e-mail. IMAP & POP3 servers can usually be accessed from anywhere on the Internet so any IP address recorded may well require further tracing to be useful.
	Webmail	Acesss to e-mail via a web interface ("webmail") may be provided as a front end to POP3 services or as a service in its own right. Logging will typically record the IP address that accesses the mail box and may record which items of mail were looked at. Webmail services are almost invariably designed to be used from any Internet address.
Access		
Circuit Switched		
	IP Address	An IP address can either be static (allocated on permanent basis) or dynamic (a different IP address allocated each time authentication is made). When mapping a dynamic address to an account it is therfore essential to provide accurate timing information (date, time and timezone).
	Account Usage	Logging of account usage will record the date and time that the connection was established, and the date and time that it. Further details such as the number of packets transferred may also be available from some ISPs. Account authorisation if often done using a system called RADIUS so these logs are often referred to as RADIUS logs. Further logs, holding much the same information, from the Network Access Servers (NASs) may also be available at some ISPs.. A number of different versions of RADIUS are in use, so that the actual format of the logs (and indeed the format of the time and date information) may vary from ISP to ISP.
	CLI	If CLI is captured by the service provider it is most likely to be recorded within the RADIUS logs. Not all types of accounts will require CLI to be presented to the ISP. Some types of account will require access only from authorised CLI. At present, ISP equipment will not usually record the CLI if it is marked to be withheld.

Fixed access		Unlike circuit switched (dial-up) services which may be provided "free", fixed access systems are invariably charged for. This means that a valid billing address will be present in the ISP's accounting systems. In addition an installation address will be recorded, though in some cases the service may be moved to another address without the ISP becoming aware of it or bothering to record the change.
	Leased Line	A leased line is a link via the local exchange that has been provided for private use. In some contexts, it's called a *dedicated* line. A leased line is usually contrasted with a *switched line* or *dial-up line*.
	Cable Modem	The connection may be a dedicated modem or integrated into a television set top box.
	ADSL	A method of providing broadband access over a standard local loop. Within the UK this is mainly provided by British Telecom who route the traffic over an ATM cloud and provide an IP data connection to the ISP.
	Satellite	A system designed for rural areas to provide high bandwidth Internet access. The return path from the user to the ISP may be direct or via the satellite.
	WiFi	A fixed base station is connected by one of the previously described methods to the Internet. Clients may access by wireless within a limited distance. Currently these systems are insecure even where encryption has been used to protect the connection. Logs may only show that someone (necessarily physically close to the base station) has been using the system but local logging may provide further traceability. Some WiFi installations are deliberately made open for public access and some commercial operations provide access for payment, which may be made in cash.
Resources		
	Shell Sessions	Details pertaining to telnet and other 'shell' login sessions may be held in several files (telnet connections are typically logged in 'last' and 'messages' files on UNIX based systems). Shell sessions may log a variety of data including start/stop and source IP.
	Web/ FTP Space	Web and FTP Space may be provided separately or as part of a service package. All of the remarks relating to billing (to identify the owner) to server logs (to identify readers) and to FTP Logs (to identify up loaders) apply to this section.

	E-mail Addresses	There is a mapping between the e-mail address and the account. This may vary between ISPs. An account may have one or more e-mail addresses associated with it. Users may have the ability to change e-mail addresses at will. Some ISPs may not hold data on previous e-mail addresses.
	Domains	ISPs provide Domain Name Service (DNS) to allow mapping between domain names and IP addresses as well as information such as where to deliver e-mail. Details of who actually owns the domain name will be held by the appropriate registrar. The ISP will have some records as to which of their customers is controlling it. There may be limited records of historical settings.
	Resource manipulation	Many services provided by ISPs and particularly those provided by third parties can be configured by the user for example e-mail may be re-directed to another account, web space requests may be directed to another server or DNS settings may be rearranged. The system that is used for this configuration may keep logs that allow historic configurations to be reconstructed.
Billing		Many forms of access are paid for. Billing data may relate to an individual or could also be that of an organization. Some systems may be sub-let and billing records will relate to the 'letting company'. Many services are re-sold. Some systems may be insecure and used without permission.
	Name, Street, Address	A service provider does not necessarily verify a subscriber's name and address details. This is dependant upon the service a subscriber utilises whilst on the Internet. In many instances the subscriber will provide CLI (Caller Line Identifier) as a part of the authentication process prior to their use of that service. This CLI can often be mapped to a geographical location by the appropriate telco.
	Credit Card/ Bank Details	For accounts where payments are made, credit card, debit card, direct debit, cheques or standing order will provide traceability through the banking system. Where postal orders or cash payments are made or accepted, these will not always be verified. It should be noted that billing information may not be retained by the backbone ISP but by the Virtual ISP who has ownership of the customer.

Glossary of Terms Used Within This Document

Access	Data access is being able to get to (usually having permission to use) particular data on a computer
ADSL	Asymmetric Digital Subscriber Line is a technology for transmitting digital information at a high bandwidth on existing phone lines
ATM	Asynchronous Transfer Mode. A switching technology for transferring packets of data. ATM was originally developed for voice application, but is now used for Internet transports and underpins current broadband technologies.
Cable Modem	A cable modem is a device that enables you to hook up your PC to a local cable TV connection in order to send and receive data packets.
CGI	Common Gateway Interface. A method of providing dynamic content within "web pages".
Chat	Facility to talk with others whilst on line
CLI	Calling Line Identifier is the telephone number that a person has used to access their required service
DNS	Domain Name System. Protocol for providing mappings from domain names to resource identifiers such as IP addresses.
Domain name	A domain name is a user-friendly method of identifying the location of resources on the Internet.
E-mail	E-mail is the exchange of text-based electronic messages by telecommunication.
E-mail Subject Line	A conventional e-mail header that is intended to provide a brief description of the contents of an e-mail.
FTP	File Transfer Protocol. A standard TCP protocol for transferring files between machines. FTP is often used to upload the content of web sites onto servers.
HTML	Hypertext Markup Language. This is the language which is used for "web pages" to indicate the structure of documents so that browsers can display them in a standardised manner.
HTTP	Hypertext Transport Protocol. A standard TCP protocol for transferring "web pages" from one machine to another.
HTTPS	Secure HTTP. Transfer of "web pages" over an encrypted transport protocol.
ICF	Internet Crime Forum. A body formed "To promote, maintain and enhance an effective working relationship between industry and law enforcement to tackle crime and foster business and public confidence in the use of the Internet in ways that respect human rights and are sympathetic to the needs of industry."

Instant Messaging	A quick and easy way of exchanging messages with others who are also online.
IM	Instant Messaging
IMAP	Internet Message Access Protocol. A standard TCP protocol for accessing e-mail that is received, organised and stored on a remote server.
MX record	DNS record entry that indicates where e-mail for a domain is to be delivered.
IP	Internet Protocol. The basic protocol used for communication between computers on the Internet.
IP address	Internet Protocol Address. A numeric value that serves to uniquely identify an interface that is connected to the Internet.
ISP	Internet Service Provider. An organisation that makes Internet services to its customers. ISPs usually provide connectivity, often many other services as well.
Java applet	A way of providing "mobile code" on web pages so as to enable extra functionality on web pages.
Leased Line	A method of providing a fixed connection to the Internet.
Local Loop	In telephony, a local loop is the wired connection from a telephone company's central office in a locality to its customers' telephones at homes and businesses
NNTP	Network News Transfer Protocol. A standard TCP protocol for transferring Usenet articles over the Internet.
POP3	Post Office Protocol 3. A standard TCP protocol for collecting e-mail from a server.
RADIUS	Remote Authentication Dial-in User Service. A standard TCP protocol for communicating authentication information and establishing parameters for dial-up connections to the Internet.
SMTP	Simple Mail Transfer Protocol. A standard protocol used for the transport of e-mail over the Internet.
SPOC	Single Point of Contact. A scheme whereby requests from law enforcement organisations are funnelled through a single part of that organisation and passed to a single contact point within ISPs.
TCP	Transport Control Protocol. A protocol layered over IP that provides a reliable delivery service for data.
URL	Unique/Uniform Resource Locator. A stylised naming system for web resources.
VISP	Virtual ISP. An ISP whose infrastructure is completely provided by third parties.
Web Cache	A cache is a server that retains copies of web content so as to provide timely local delivery for repeat requests.

Web Proxy	A proxy is a server that acts as an intermediary between a workstation user and the Internet
WiFi	Wireless systems (such as 802.11) that provide Internet connectivity.